Eukaryotic Chromosome DNA Replication: Where, When, and How?

Hisao Masai, Seiji Matsumoto, Zhiying You, Naoko Yoshizawa-Sugata, and Masako Oda

Genome Dynamics Project, Tokyo Metropolitan Institute of Medical Science, Tokyo 136-8506, Japan; email: masai-hs@igakuken.or.jp

Key Words
DNA helicase, prereplicative complex, replication fork, replication origin, replication stress, replication timing

Abstract
DNA replication is central to cell proliferation. Studies in the past six decades since the proposal of a semiconservative mode of DNA replication have confirmed the high degree of conservation of the basic machinery of DNA replication from prokaryotes to eukaryotes. However, the need for replication of a substantially longer segment of DNA in coordination with various internal and external signals in eukaryotic cells has led to more complex and versatile regulatory strategies. The replication program in higher eukaryotes is under a dynamic and plastic regulation within a single cell, or within the cell population, or during development. We review here various regulatory mechanisms that control the replication program in eukaryotes and discuss future directions in this dynamic field.
1. INTRODUCTION

Studies regarding the mechanisms of DNA replication for the past 57 years, since the discovery of double-stranded DNA by Watson and Crick, revealed that sequential assembly and reorganization of complex arrays of proteins are crucial for the coordinated execution of initiation, elongation, and termination processes of DNA replication. The progression of the replication fork is monitored strictly to ensure complete replication of the entire genome. The physiological significance of the execution of the entire process has been emphasized by the finding that defects in proteins for assembly and monitoring of the replication fork lead to genomic instability, resulting in carcinogenesis or a series of diseases collectively known as the “chromosome instability syndrome.”

The extensive conservation of replication factors strongly suggests that the basic mechanism of DNA replication is evolutionally conserved. However, accumulating data also indicate that the modes of recognition and regulation of its firing are significantly more complex in eukaryotes than in prokaryotes and may be substantially different even between lower and higher eukaryotes. One of the most striking features of DNA replication in higher eukaryotes is its plasticity. A best-known example of plasticity is the very rapid replication of the entire genome of the fertilized eggs of amphibians (within a matter of 20–30 min), compared with the 8–10 h required for the genome replication in somatic cells. This robust process is achieved by very frequent initiation through utilization of virtually all the potential origins in contrast to the spatially and temporally regulated origin firing that occurs in differentiated cells. In spite of this, how this rather dramatic transition of the replication mode takes place is still under debate.

The challenges facing our understanding of eukaryotic DNA replication can be summarized by three questions—where, when, and how: Where in the genome and within the nuclei does DNA replication take place? When within the S phase does each replication origin start firing? And how is the entire process of DNA replication regulated?

In this review, we summarize recent progress in eukaryotic DNA replication and discuss how it contributes to our understanding of the basic mechanisms that enable highly strict and
plastic regulation of this process. We also discuss how failures in the proper execution of replication program contribute to the etiology of various diseases including malignancy. Because the last review on eukaryotic DNA replication was published in the Annual Review of Biochemistry in 2002, we cite studies that were published mainly after that review. The readers are encouraged to refer to Reference 1 for the literature published before 2002.

2. WHERE?
Where on the genome and within the nuclei does DNA replication take place?

2.1. Assembly Sites of Replicative Complexes

2.1.1. Recognition by the initiator origin recognition complex. In bacteria, the initiator DnaA protein determines the site of initiation on the genome. Searches for the initiator of eukaryotic genomes led to the discovery of the origin recognition complex (ORC) (2). In the budding yeast, where ORC was first identified, it specifically recognizes the 11-hp or 17-bp conserved sequence present in the replicator sequences (3–5; see below). The ORC is evolutionarily conserved and plays important roles in DNA replication across species (1). However, the ORC from other species exhibits DNA-binding specificity that is much more relaxed than that of budding yeast. For example, the fission yeast ORC recognizes DNA through the AT-hook motif present on the ORC4 subunit and preferentially binds adenine/thymine stretches (6), and the Drosophila ORC also shows some preference for AT-rich sequences. In contrast, the human ORC binds to any DNA without apparent sequence specificity (7, 8). However, the sites of initiation of DNA replication are not random, at least in differentiated somatic cells, suggesting that the ORC binds to genomic DNA with certain specificity. The characterization of sequence elements required for ectopic replication from mammalian replicators has identified AT-rich sequences (9–11), dinucleotide repeats (10), asymmetrical purine-pyrimidine sequences (11), and matrix attachment region (MAR) sequences (8, 12) as being required for efficient initiation of DNA replication. These sequences in the chromatin context may facilitate recognition by the ORC. It should be noted that even the ORC in budding yeast could exhibit a relaxed mode of DNA binding in vivo and in vitro (13–15), and thus, DNA binding without significant specificity may be a conserved feature of the ORC. This is quite different from bacterial initiators, which are highly specific to cognate replication origin sequences.

Other factors may facilitate the selection of specific segments on the genome. The potential factors include the topology of DNA. Replication in vitro at the chromosomal origins of DNA replication of Escherichia coli depends on the negative superhelicity of the template DNA (16). In Drosophila melanogaster as well as in fission yeast, the ORC exhibits a preference for supercoiled DNA (17, 18). Topoisomerases were reported to associate with a human replication origin (312). Transcription factors may also play significant roles in the specific localization of the ORC. At the chorion loci of Drosophila follicle cells, transcription factors that contain the Myb protein seem to facilitate the site-specific DNA replication of the ORC at ACE3 (amplification control element 3) and Ori-β (19). This locus also interacts with retinoblastoma protein (Rb) (20). At the rat aldolase B gene promoter origin, binding of the ORC1-interacting factor AIF-C to a specific sequence in the origin is required for efficient initiation of replication (21). Recruitment of the ORC to the Epstein-Barr virus (EBV) replicator, oriP, was shown to be dependent on the RNA that links oriP-bound EBV-coded nuclear antigen-1 (EBNA-1) and the ORC (22). Prereplicative complex (pre-RC) factors may also contribute to origin recognition. In budding yeast, Cdc6 ATPase activity contributes to stable and specific binding of the ORC-Cdc6 complex to origin, because its activation by nonorigin DNA promotes dissociation of Cdc6 (15). Binding of fission yeast ORC to the origin DNA is facilitated by Cdt1 and Cdc18 proteins (18).
SUMMARY OF GENOME-WIDE INVESTIGATION OF DNA REPLICATION USING MICROARRAY ANALYSES

In the eight years since the last review, several studies on genome-wide analyses of DNA replication have been published. They can be categorized into two groups: (a) mapping of replication origins and factor binding sites and (b) mapping of replication timing. The methods used in the first group include (a) analyses of replicated heavy-light and unreplicated light-light fractions [after labeling with bromodeoxyuridine (BrdU)] isolated at different time points during S phase and separated by a density gradient centrifugation, (b) analyses of copy number changes, (c) analyses of the single-strand DNAs generated, (d) analyses of the ORC-MCM-binding sites by ChIP-Chip assays, and (e) analyses of BrdU incorporation by ChIP-Chip assays. The methods for the second group include (a) fractionation of cells into early and late S-phase populations and analyses of BrdU-incorporating DNA, and (b) analyses of round mitotic cells collected after “baby machine”-based cell cycle synchronization [in which cells were pulse-labeled, washed, and incubated for different time intervals (126)].

The following are conclusions on replication origins drawn from the above studies:

1. In budding yeast, pre-RC sites (ORC and MCM binding) coincide with actual initiation sites (origins) (23, 81).
2. In fission yeast, the pre-RC is present with an average interorigin length of 10 kb (25). About 40% of the pre-RC assembly sites are utilized for initiation in the presence of hydroxyurea (on chromosomes I and II) (Y. Kanoh, personal communication).
3. No consensus sequence was identified for origins. AT-rich (>70%) intergenic regions serve as potential origins in fission yeast (24, 25, 86, 307).
4. In Drosophila, the consensus sequences for origins and for ORC binding have not yet been identified. The ORC-binding sites are overlapped by RNA polymerase II–binding sites, suggesting that transcription factors could contribute to selective binding of the ORC (26).
5. In human, origins are abundant in the CpG island promoter. Half of the origins are localized within or near CpG islands (27).
6. In human, origins are strongly associated with transcriptional regulatory elements (e.g., c-Jun and c-Fos) (27).
7. In mammals, origin sequences are evolutionarily conserved between different species (27).

2.1.2. Origin recognition complex and minichromosome maintenance complex-binding sites as potential origins. The minichromosome maintenance (MCM) complex, a heterohexameric complex of MCM 2, 3, 4, 5, 6, and 7 subunits, is loaded onto the ORC bound to chromatin at late mitosis (M) to early G1 to generate a pre-RC. Genome-wide chromatin-immunoprecipitation assays in budding yeast showed that the ORC- and/or MCM-binding sites include 95% of the known autonomously replicating sequence (ARS), suggesting that the ORC- and/or MCM-binding sites can predict the potential origins with high accuracy (23; Supplemental Tables 1 and 2). Follow the Supplemental Material link from the Annual Reviews home page at http://www.annualreviews.org. See also the box titled Summary of Genome-Wide Investigation of DNA Replication Using Microarray Analyses.

In fission yeast, the ORC- and MCM-binding sites also correlate well with the replication origins predicted on the basis of AT-richness (24, 25; Y. Kanoh, personal communication). These potential origins are present with an average of 10-kb intervals, mainly at the intergenic segments in fission yeast. Similar assays with the Drosophila chromosome identified distinct ORC-binding sites colocalized with the actual initiation sites (26). Although similar ChIP-Chip assays have not been reported in human, mapping of the initiation sites resulted in the identification of 283 origins in 1% of the human genome, which largely colocalized with transcriptional regulatory elements (27–29). These results indicate that potential origins are assembled at specific sites even on the genome of higher eukaryotes. Please refer to Supplemental Table 3. Follow the Supplemental Material link from the Annual Reviews home page at http://www.annualreviews.org.

2.2. Selection of Replication Initiation Sites of DNA Synthesis

There is general agreement that cells are equipped with excess numbers of pre-RCs on the genome (potential origins) and that only a
subset of these pre-RCs is probably utilized for actual initiation (30, 31). Two of the outstanding questions are how this selection is regulated and how it is related to the plasticity of eukaryotic DNA replication, as discussed below (see Section 2.3.). Selection of the initiation sites may be regulated by various factors and may be achieved at the activation step as well as at the pre-RC assembly step.

2.2.1. Distal elements. The distal DNA elements can have a profound effect on the initiation site. In the Chinese hamster ovary cells, DNA replication at the DHFR (dihydrofolate reductase) locus (which is amplified) initiates within the 55-kb intergenic region (initiation zone) (32). When the transcription reads through into the intergenic region as a result of deletion of the 3′ end of DHFR, the initiation is confined to the far end of the intergenic region. Introduction of a small DNA segment derived from the 3′ end of the gene (containing transcription terminator activity) generates the boundary defining the end of the originless segment (33). In the human β-globin locus, replication starts at the initiation region located at the 5′ end through the coding region of the β-like β-globin gene (34). This initiation event depends on the presence of the locus control region, which is located further upstream of the globin gene clusters and serves as a master control element of β-globin gene expression. In the cytokine cluster segment of mouse chromosome 11, initiation of replication at the 3′ side of the IL-13 gene requires the presence of the conserved noncoding sequence in the 10-kb intergenic region between IL-13 and IL-4 (35). The conserved noncoding sequence, containing the DNaseI-hypersensitive sites, is required for cell-type-specific expression of Th2 cytokines. At present, the mechanism by which these distant elements affect the initiation events is unknown.

2.2.2. Chromatin structures. Chromatin structures are important factors not only for transcription but also for origin selection. In yeasts, mutations in histone deacetylases (HDACs), such as Sir2, facilitate the initiation events. The sir2 mutation rescues replication mutants involved in pre-RC assembly but not those involved in the following activation step, and in fact loading of the MCM complex onto origins is facilitated by the sir2 mutation at some origins (36, 37). These sensitive origins contain a sequence called I′, which mediates Sir2-dependent inhibition (38). In Xenopus egg extracts and human cells, HBO1, a histone acetyltransferase, is required for loading the MCM complex onto chromatin (39). HBO1 directly interacts with Cdt1 and enhances Cdt1-dependent rereplication (40). HBO1 interacts with the ORC and MCM2, suggesting that histone acetylation and deacetylation could affect the formation of the pre-RC. Thus, chromatin structures seem to affect the efficiency of pre-RC assembly, which may indirectly alter the initiation frequency at particular loci (41).

8. Origin specification and firing efficiency are generally maintained across different cell types in mouse (28).
9. Origins for re-replication are not identical to those for normal replication in budding yeast (308).

The conclusions on replication timing are as follows:
1. Genomes can be divided into different replication timing domains; each domain contains one to tens of replication origins that fire in a coordinated fashion (26, 108–111).
2. In budding yeast, there is no correlation between replication timing and transcription or GC (guanine or cytosine) content (81).
3. In metazoan, generally significant correlation was detected between replication timing and transcription or GC content (Supplemental Table 4).
4. Replication timing domains undergo substantial reorganization during development or differentiation (111, 127).
5. Timing transition regions are associated with the frequent occurrence of mutations responsible for various diseases and single-nucleotide polymorphism (134). (Note: This conclusion was drawn by regular analyses, not by microarray analyses.)

Helicase: an enzyme capable of separating the duplex DNA using the energy obtained by hydrolysis of ATP.

At the ACE3 and Ori-β origins of *Drosophila*, histone acetylation colocalizes with the ORC, and inhibition of HDAC results in hyperreplication concomitant with hyper-acetylation and redistribution of the ORC (42). Site-specific DNA replication was observed in *Xenopus* egg extracts, where initiation is normally random, by introducing a transcriptional complex at a specific site (43). This initiation at the targeted site correlates with local hyperacetylation of histones. Because the distribution of the ORC is not altered, transcription factor-induced localized chromatin remodeling or histone acetylation may facilitate the assembly of the pre-RC and/or activation of the pre-RC (43). Thus, chromatin structure, more specifically histone acetylation, appears to facilitate initiation, affecting the site of initiation. By contrast, at the β-globin locus, histone modifications did not precisely correlate with the activity of the replication origin. Thus, chromatin structures may play a pivotal role in origin selection but are probably not essential for firing (44).

2.2.3. Transcription, topology, and nucleosome formation. Replication origins are abundant in the intergenic regions. Generally, origins present downstream of a transcriptional unit are less active compared to those present upstream (45, 46), and this was thought to be caused by interference of the origin function by read-through transcription (47–49). Topology is another important factor in initiation of DNA replication. Duplex unwinding is highly stimulated by the negative superhelicity. Therefore, we also consider the possibility that transcription-induced topological changes could affect initiation. It is known that transcription induces negative supercoiling behind the transcription bubble (50). Thus, two divergent transcription events will generate strong negative supercoiling in the intergenic region, facilitating strand opening. In fact, genome-wide determination of pre-RC assembly sites and actual initiation events in fission yeast revealed that the initiation frequency in the presence of hydroxyurea is 50% and 25% among the pre-RC assembled at the intergenic region between divergent transcription and that between convergent transcription, respectively (25; Y. Kanoh, personal communication). Furthermore, recent findings indicate that divergent transcription occurs in most of the active genes, including the ones regarded as single transcription units. This will maintain the promoter segment between the two divergently moving RNA polymerases nucleosome free (51), which will permit more efficient assembly of the pre-RC.

2.2.4. Other factors regulating the site of initiation. It was reported that the mammalian MCM4-6-7 helicase is specifically activated by T-rich single-stranded DNA (ssDNA), including the DNA derived from known replication origin sequences (52). The ability to activate the MCM helicase correlated with the presence of stretches of thymine residues that are often found in the vicinity of eukaryotic replication origins. Thus, the chance of origin activation may be determined by availability of single-stranded thymine-rich DNA, which may be regulated by the intrinsic nature of the sequence as well as by the conformational state of the pre-RC, or chromatin, and/or topological context.

The initiation of DNA replication can influence that from adjacent loci. In budding yeast, a propagating fork from a neighboring origin suppresses the firing of a potential origin, probably by disrupting the preformed pre-RC (53). UV-footprinting analyses showed that the pre-RC to post-RC conversion is observed at suppressed late origins after passage of the replication fork from neighboring origins (54). By analyses of single DNA molecules, it was demonstrated that this so-called origin interference does occur in human cells as well, maintaining a relatively constant interorigin distance (55). The precise mechanisms of origin interference are not known, although passive replication may well play a significant role in the phenomenon.
2.2.5. Regulation at the origin decision point. Analyses of the replication pattern of nuclei isolated from differentiated mammalian cells in *Xenopus* egg extracts resulted in the discovery that initiation takes place randomly in nuclei isolated from the early G1 phase. However, beyond a certain point during the G1 phase, called the origin decision point (ODP), nuclei exhibit site-specific initiation, characteristic of differentiated cells. The ODP is temporally distal to the timing decision point (TDP) yet occurs 2–3 h prior to the restriction point, and its execution requires kinase activity, proteolysis, and transcription (56). Transcription may eliminate subsets of the assembled pre-RC but is not sufficient for origin specification.

2.3. Plastic Regulation of Origin Selection

2.3.1. Highly plastic mode of eukaryotic DNA replication. Eukaryotic DNA replication exhibits a great deal of plasticity (Figure 1). In fact, studies by Taylor in 1977 indicated that the cellular replication program is dynamically regulated (57, 58). In those studies, Chinese hamster ovary cells were depleted of thymidine by treatment with the drug fluorodeoxyuridine (FdU). DNA synthesis was then analyzed by labeling of the newly synthesized DNA with tritiated thymidine and visualized by DNA fiber autoradiography. These experiments indicated that increased duration of exposure to FdU (more starved for thymidine) resulted in reduced distances between the labeled segments on DNA fibers and reduced rates of fork movement. Apparently, therefore, thymidine-starved cells compensate for the reduced fork movement by increasing the number of initiation sites to complete the S phase within a given time.

Nearly 30 years later, this finding was confirmed and extended by using a novel fluorescence microscopy technique (59). Using DNA segments spanning the adenosine deaminase 2 (*AMPD2*) gene, which was amplified in a Chinese hamster lung fibroblast cell line, it was clearly shown that cells with slower fork movement activate additional origins on this DNA segment. Conversely, accelerating the

![Figure 1](https://www.annualreviews.org/doi/10.1146/annurev-biochem-050110-100333)

Plastic regulation of eukaryotic DNA replication. Eukaryotic genomes are equipped with potential replication origins (pre-RCs, blue circles) far exceeding in number what is actually required for the completion of genome replication. The initiation sites are selected from these potential origins. When the fork rate is fast, the initiation takes place less frequently (left), whereas initiation occurs more frequently when the fork rate is slower (right). When forks are stalled, the dormant (silent) origins fire (center).
fork movement by adding adenine and uridine reduced the interorigin distance and restored the preferential firing at an efficient origin, resulting in a lower density of replication origins. Furthermore, fork rates are coordinately regulated with interorigin distances in the unperturbed normal cells (60). This highly plastic, reversible nature of the replication program permits the cells to respond to various internal and external interferences as well as to developmental requirements. Potential mechanisms by which plasticity of DNA replication programs may be achieved are considered below.

2.3.2. Chromatin loops and origin selection.

It was reported that differentiated erythrocyte nuclei replicate only inefficiently (have fewer initiation events) in *Xenopus* egg extracts, but gain efficient replication ability (have high-frequency initiation) if they experience a single mitosis in egg extracts (61). This mitotic remodeling of chromatin is exemplified by the shortening of chromatin loops and recruitment of more initiator ORC proteins. Chromatin loops are known to regulate transcription by juxtaposing distal genomic loci and have been implicated in the determination of replicon sizes as well. The chromatin loop assays conducted with the above cells at different fork rates confirmed that loop sizes in G1 phase cells correlate with the interorigin distances in the preceding S phase. It was also shown that the loop sizes are established only after cells go through M phase. It was proposed that the “replication mark” left on the chromosome during the preceding S phase determines the loop sizes in the following M phase (62). In addition, the authors proposed that the mark may be related to the termination of replication forks present at the center of two active origins. Mitotic resetting in *Xenopus* egg extracts depends on topoisomerase II (61). Topoisomerase functions are essential for decatenation of two daughter strands at the location where two converging forks collide (63). Genome-wide mapping indicated the specific localization of topoisomerase II at the center of two origins (D. Fanchinetti & M. Foiani, personal communication). It is tempting to speculate that topoisomerase II involved in the decatenation of the daughter strands may contribute to the chromatin loop formation in the following M phase, keeping the memory of the origin density in the preceding S phase in the chromatin structures.

2.3.3. Dormant origins facilitate the maintenance of genomic integrity. The MCM complexes are present on chromatin in a number far in excess of the actual initiation sites (30, 31). What then are the roles of these seemingly surplus origins? It was reported in yeast that reduced numbers of pre-RCs can cause genomic instability (64, 65). Inhibition of pre-RC formation in human cancer cells by overexpression of the nondegradable form of geminin resulted in early or late S-phase arrest, eventually resulting in cell death (66). More recently, the effects of decreased pre-RC assembly on cell survival under normal and stress conditions were examined in *Xenopus* egg extracts and human cells (67–69). These studies concluded that these “dormant” origins are utilized when forks are stalled. This would facilitate the replication of those segments of the genome, which would otherwise stay unreplicated owing to flanking stalled forks on both sides. Cells with reduced levels of pre-RCs proceed through S phase without any apparent defect under an unperturbed condition but exhibit increased sensitivity to replication fork stress.

The adjustment of origin firing in response to fork progression rate appears to occur rapidly. The firing of additional origins takes place within a matter of minutes after the fork slows down, and the decrease in the overall number of active origins occurs within 2 h after a shift to “fast fork” conditions (62). The firing of dormant origins in response to fork stalling would also be a rapid cellular response. How does fork slowing or stalling induce activation of dormant or silent origins? One model hypothesizes that origin firing is a stochastic process and that the stalling fork increases the chance of dormant origins to be fired (see below). Under normal conditions, these origins
would be replicated passively, and are inactivated (70). Other models hypothesize that the stalling fork alters the local chromatin structure or topology of the template DNA, thus creating a permissive structure for activation of the nearby pre-RCs (69). Indeed, the introduction of a site-specific double-strand break activates initiation at silent origins nearby (71). The effect appears to be local, and activation seems to take place earlier than expected from random activation, which may support the second model.

It has been generally assumed that stalling of the replication fork by hydroxyurea or aphidicolin, the condition used to slow the fork rate in the above experiments, induces checkpoint reactions and inhibits further origin firing (58, 72). It is unclear how the same treatment can induce opposite outcome. The conditions used for activation of dormant origins are generally much milder than those used for checkpoint activation (68, 69). Cells may differentially respond to the changing environment to adopt the most suitable cellular responses. Alternatively, as discussed above, origin recruitment may be a localized effect, affecting only the origins in the same cluster. The fork rate may dynamically affect the chromatin loop size during the S phase, activating additional origins (small loops) or suppressing potential origins (larger loops), when it is slowed or accelerated, respectively.

2.4. Localization of Chromatin within the Nucleus

2.4.1. Nuclear positioning is important for regulation of chromosome functions. Eukaryotic chromosomes are nonrandomly positioned within the nucleus. Each chromosome occupies a separate nonoverlapping position (interphase chromosome territory) in the nucleus (reviewed in References 73 and 74). Within each chromosome, the condensed heterochromatin tends to localize at the nuclear periphery (reviewed in Reference 75), and specific positioning within the nucleus is also observed for the chromatin regions where efficient transcription or DNA replication occurs. In budding yeast, proteins involved in transcriptional silencing, such as Sir4p, are abundantly present in discrete regions at the nuclear periphery, and positioning of chromosomal loci within these regions appears to facilitate transcription repression. Both the mating-type loci and telomeres of *Saccharomyces cerevisiae* tend to localize at the nuclear periphery (76). The nuclear periphery has been proposed to consist of a mosaic of microenvironments, some of which favor silencing and others transcriptional activation (77–79). It has been proposed that the DNA replication machinery is anchored at specific nuclear structures called the “replication factories.” See Replication Factories: Site of DNA Replication? in the Supplemental Material for more details. Follow the Supplemental Material link from the Annual Reviews home page at http://www.annualreviews.org.

3. WHEN?

When within the S phase does each replication origin start firing?

3.1. Timing of DNA Replication within the S Phase

In normal human cells, the S phase lasts 6–8 h. The genomes are divided into distinct domains; some domains replicate early in the S phase, and others replicate later in the S phase (Figure 1). Replication timing appears to be predetermined in a given cell type. It has been known that transcriptionally inactive heterochromatin, contained in Giemsa dark chromosome bands or G bands, replicates late during the S phase, whereas Giemsa light or R bands containing euchromatin, where most transcription takes place, replicate early (80). However, as described in the following section, the timing seems to be related to the larger chromatin and nuclear architecture.

3.1.1. Determinants of timing regulation. In budding yeast, early- and late-firing origins

**Checkpoint:** a surveillance system that ensures ordered progression of the cell cycle

**Replication timing:** timing within the S phase when a segment of genome is replicated. Each chromosome consists of many segments of different replication timing
have been clearly identified (81). Although the pre-RC is formed at all the potential origins at the late M/early G1 phase, activation of the pre-RC occurs in a genetically determined order on a chromosome (82). Telomeres are major elements that confer a late-replicating property (83). The replication timing of the origins near telomeres (within 10 kb) is rendered late as a result of the telomeric chromatin structure because the mutation in Sir3, required for spreading the silent domain, causes them to replicate early (84). The search for determinants of replication timing in late-firing origins distal to telomeres also led to identification of specific sequences that render early replicating origins to become late replicating (85).

In fission yeast, the definition of early- and late-firing origins is less clear. There is certainly a set of efficient origins that fire in the presence of hydroxyurea, and other origins are inefficient and normally do not fire under the same conditions (25, 86). It is not known whether these inefficient origins actually fire during the normal course of S phase or are passively replicated, although Cdc45 loading may take place in late S phase at some of these inefficient origins (87). However, the presence of sequences that inhibit early firing has been suggested; a GC-rich sequence was identified, which enforces late replication on ARS-driven plasmid replication in fission yeast (88). Relocation of an inefficient origin to the early replicating segment and that of an efficient origin to the late-replicating segment led to reversal of the replication timing in each case (25).

The mechanisms that regulate replication timing in metazoans remain elusive. Replication in Xenopus egg extracts occurs continuously throughout the S phase without apparent timing regulation in the entire genome. Incubation of nuclei from differentiated cells at various cell cycle stages in Xenopus egg extracts indicated that nuclei from M or early G1 phases replicate in an “embryonic” manner, whereas those from after the TDP within the G1 phase replicate with a unique timing regulation (in a differentiated manner) (89). However, the nature of the transition that occurs at the TDP is unknown. It was proposed that nuclear structures may play a role in determining the timing of DNA replication. Distinct patterns of DNA replication foci have been observed for mammalian cells in early, mid-, and late S phase. It was proposed also that anchoring of DNA to specific nuclear structures may determine the time at which the segment replicates within the S phase (80, 90). In fact, the transcriptionally inactive β-globin locus in Chinese hamster ovary cells becomes localized to the nuclear periphery between one and two hours after mitosis and then replicates late during the S phase within the same time frame as peripheral heterochromatin. This defines the TDP for the β-globin locus and suggests that establishment of the replication timing program coincides with its association with the nuclear periphery (90).

In budding yeast, late-replicating origins tend to localize close to the periphery of the nucleus specifically during the G1 phase, whereas early replicating origins tend to show random localization. Excision of one of the late-replicating origins (the ARS501-containing segment) as a circular plasmid during mitosis results in early replication, and excision during the G1 phase does not affect late replication of the excised ARS501-containing plasmid (91). Hence, interaction of the segment adjacent to ARS501 with nuclear periphery during G1 is required for establishing the late-replicating properties of ARS501. Furthermore, peripheral localization is required for its establishment but not for its maintenance. Indeed, ARS501 is abundantly found in nuclear periphery in the early G1 phase, but this distribution is not maintained in S phase, whereas early-firing origins are randomly localized within the nucleus throughout the cell cycle (92). These results suggest that the spatial arrangement of nuclear positions during the G1 phase may play a crucial role in determination of the replication timing in yeast, consistent with the above TDP proposal in mammalian cells.
3.1.2. Chromatin structures and replication timing. Deletion of Rpd3, a histone deacetylase, causes earlier firing of origins in budding yeast (93, 94). This is accompanied by advanced binding of Cdc45, an essential factor for generation of a replisome, to origins, suggesting that the firing step is accelerated under the condition of increased histone acetylation. A genome-wide search of initiation events indicates that more than 100 late-firing origins are deregulated in the absence of Rpd3L, one of the Rpd3 complexes (95). The Rpd3S complex (which is functionally distinct, but retains histone deacetylation activity) also affects replication timing, suggesting that histone deacetylation directly influences the initiation timing. Furthermore, a recent report indicated that not only histone acetylation but also methylation can accelerate S-phase progression and loading of Cdc45 at origins (96). Accelerated progression of S phase and Cdc45 loading at origins in the absence of Rpd3, one of the Rpd3 complexes (95). The Rpd3S complex (which is functionally distinct, but retains histone deacetylation activity) also affects replication timing, suggesting that histone deacetylation directly influences the initiation timing. Furthermore, a recent report indicated that not only histone acetylation but also methylation can accelerate S-phase progression and loading of Cdc45 at origins (96).

3.1.3. The trans-acting factors determining replication timing. Firing of replication origin depends on the phosphorylation events mediated by two kinases, the cyclin-dependent kinase (CDK) and the Dbf4-dependent Cdc7 kinase (Cdc7-Dbf4 or DDK) (99, 100). In budding yeast, firing of late origins is specifically abrogated in cdc5 mutant cells (one of the B-type cyclins), indicating that Cdc28 kinase plays a crucial role in activation of late origins (101). It was recently shown that Cdc2-cyclinA may regulate the timing of DNA replication. Induced activation of Cdc2-cyclinA resulted in premature activation of late replication origins in human cells (102).

The Cdc7-Dbf4 kinase complex is required for activation of each origin in budding yeast (103, 104). In fission yeast, origin firing efficiency is accelerated by an increased level of the Hsk1 catalytic subunit (homolog of Cdc7) or Dfp1 protein, its activation subunit (homolog of Dbf4) (87, 105), suggesting that Hsk1 kinase may also be a critical and limiting regulator. Both CDK and Cdc7 are required for the loading of Cdc45 protein at the origins (106), and overproduction of Cdc45 also increased origin efficiency (87). In budding and fission yeasts, loading of Cdc45 at origins is regulated under a temporal program, i.e., Cdc45 is loaded at early-firing origins at the onset of S phase, whereas it is loaded at late-firing origins late in S phase (87, 107). Thus, differential regulation of origins by CDK and/or Cdc7 may explain their early or late firing. How can these kinases differentially regulate the early and late-firing origins? In view of the significant roles played by the chromatin structure in timing control, histone modification near the origin may regulate access to these kinases. Alternatively, a specific factor that preferentially associates with the pre-RC at the early-firing origins may serve as a mark for early activation by these kinases. Furthermore, a factor that selectively associates with late-firing origins could

Replisome: a putative protein complex responsible for the coordinated synthesis of leading and lagging strands at a replication fork

CDK: cyclin-dependent kinase
negatively regulate the firing of these origins, and phosphorylation events may counteract this inhibition. In any case, the epigenetic marks on the protein complex at origins are likely to play crucial roles in differential activation.

3.1.4. Global genome organization and replication timing. Replication timing is generally regulated at a global genome level, and loss of a single origin does not affect global timing regulation. This is especially true in higher eukaryotes. Thus, each replication domain may be a functionally compartmentalized unit that coordinately regulates the origins contained within the domain. At first glance, the replication origins in budding yeast may appear to be regulated individually. However, large domains (>130 kb) contain multiple origins that initiate coordinately (85). Recent genome-wide analyses of origin firing in ccl5Δ mutants indicated that the budding yeast genome can be divided into a Ccl5-independent early-firing domain and a Ccl5-dependent late-firing domain, suggesting that the replication domains may exist also in budding yeast (101).

In metazoan, replication timing is determined at the large chromosomal domain. A replication domain can be as long as several megabases, which may contain just a single replicon or more than 10 replicons (26, 108–111). Inactivation of any particular origin may not affect the overall timing of the segment (35). In fact, replication timing is independent of the site at which replication is initiated. The fact that the TDP occurs independent of the ODP supports this notion. Global genome organization may be coupled to subnuclear repositioning and may also be linked to specialized chromatin structures.

3.1.5. Checkpoint regulation of replication timing. In order to coordinate fork progression with progression of S phase and to maintain the integrity of the S-phase chromosome, eukaryotic cells monitor the progression of the replication fork and elicit the so-called intra-S-phase checkpoint, when they are threatened by replication stresses. This checkpoint response reduces the total rate of DNA replication, which can be achieved by inhibiting the firing of new origins and/or by slowing down ongoing replication forks (112, 113).

In budding yeast, inhibition of new origin firing appears to be a major pathway. The firing of late origins is restrained by checkpoint mechanisms, because a mutation in mec1 or rad53, sensor or effector kinase, respectively, leads to early activation of the late origins in the presence of hydroxyurea, which inhibits the ongoing DNA synthesis owing to depletion of nucleotide precursors (114, 115). In mammalian cells, both the initiation and elongation phases of DNA replication are inhibited in a manner dependent on the Chk1 effector kinase (72, 112, 116).

Checkpoint kinases regulate origin firing even in the absence of fork stress. Ataxia-telangiectasia (AT) cells lacking Ataxia-telangiectasia mutate (ATM) proteins were originally observed to replicate more rapidly than the wild-type cells (117). Likewise, the rate of bulk DNA synthesis increased in cells exposed to caffeine, which inhibits both ATM and ATR (ataxia-telangiectasia-mutated and Rad3-related) protein kinases (118). Furthermore, reduction of Chk1 protein levels caused elevated origin firing and increased DNA synthesis during an unperturbed S phase (119, 120). Similarly, deletion of Cds1, a checkpoint effector kinase of fission yeast, exhibits unusual replication foci, potentially reflecting the perturbed temporal replication program (121). These results indicate that ordered firing of replication origins during a normal S phase is under the regulation of checkpoint functions.

3.1.6. Is the timing of replication stochastic or predetermined? Analyses of origin activation in Xenopus egg extracts and more recently in fission yeast led to the proposal that origins are selected from preformed pre-RCs in a stochastic manner. Mathematical calculation predicts that large gaps may remain unreplicated in the stochastic model (the “replication gap” problem) (122). This could be
circumvented by the increased firing frequency in the late S phase (122, 123). In fission yeast, DNA fiber analyses of origin firing at specific loci during two consecutive rounds of S phase led to the conclusion that origin firing occurs stochastically at the preformed pre-RC sites (124). Chromosome-wide single-molecule analyses of replication kinetics in budding yeast showed that no two molecules have the same replication pattern, but averaging of all the patterns reproduced the results of bulk timing experiments (82), indicating that origin firing is probabilistic even in budding yeast. Firing of dormant origins in the presence of replication stress could be explained also by stochastic firing (70).

Genome-wide mapping of pre-RC assembly sites and initiation sites on the fission yeast genome indicates that replication is initiated only at the pre-RC sites (25). The location and extent of pre-RC assembly is highly reproducible, and pre-RCs are present on average once every 10 kb (25). Out of these origins, approximately 40% are utilized as efficient initiation sites in the presence of hydroxyurea (Y. Kanoh, personal communication). This selection is also highly reproducible and is not completely random. There is an obvious hierarchy in the selection of active initiation events. The data support the presence of preferred potential origins, which are selected for firing before other nonpreferred origins. Among the preferred origins that may exist in a cluster, the selection may be stochastic and exhibit variability among different cells or during the consecutive cell cycle. What then determines the hierarchy?

It was recently reported that loading of the ORC in fission yeast is cell cycle regulated, and it was proposed that the timing of ORC loading (thus pre-RC formation) during the late M phase may determine the initiation efficiency at the S phase (87). Prolongation of the M phase resulted in decreased initiation at “strong” origins and increased initiation at “weak” origins, and this may reflect decreased and increased ORC loading, respectively, at these origins during the M phase.

Overexpression of Hsk1, Dfp1, or Cdc45, but not Cdt1, resulted in a global increase of origin firing in fission yeast (87, 105). This suggests that the factors required for pre-RC activation are limiting. Because the levels of pre-RCs accumulating at early and late origins are roughly the same, it may not be the amount of a pre-RC that determines the preference. Advanced formation of the pre-RC at early origins may provide more time for the epigenetic or posttranslational modifications required for being preferentially targeted by Cdc7. Alternatively, there may be additional factors that are differentially recruited to early and late origins in a manner dependent on the timing of the pre-RC formation.

3.1.7. Transcription, chromatin, and replication timing: lessons from genome-wide analyses. A positive correlation between transcription and early replication has been well documented. Genome-wide analyses conducted in Drosophila and human cells support a positive correlation between transcription and replication timing (108, 110, 125, 126), whereas those in budding yeast revealed no significant correlation (81). The relationship in higher eukaryotes is indirect because 10% to 20% of late-replicating genes are expressed, and some genes change transcription without changes in replication timing and vice versa. The entire male X chromosome in C18 cells (isolated from wing imaginal discs) is early replicating, whereas the female X chromosome in Kc cells (derived from an embryo) contains late-replicating segments in spite of their similar transcription profiles. The enrichment of acetylated Lys16 of histone H4 may be responsible for early replication of the former chromosome (127). Thus, it would be accurate to state that replication timing is related to chromatin and nuclear architecture rather than transcription itself (109). (Refer to Supplemental Table 4.)

3.2. Developmental Regulation of Replication Timing

Differential regulation of replication timing in different cell types is most clearly illustrated
at the $\beta$-globin locus, which replicates early in erythroid cells and late in nonerythroid cells (128–131). An approximately 1-Mb segment surrounding the $\beta$-globin gene constitutes the acetylated, DNase I-sensitive domain in erythroid cells, whereas it is within the hypoacetylated, DNase I-resistant domain in nonerythroid cells. The naturally occurring deletion of the locus control region, located 50–75-kb upstream of the $\beta$-globin locus, renders the transcription inactive and also makes the entire globin locus late replicating (130).

Another well-established example of the developmental regulation of replication timing is X chromosome inactivation in female mammals, which accompanies a shift from early to late replication of the inactive X chromosome in the epiblast of $\sim$6.0 dpc (days postcoitum) mice (132). These facts suggest that replication timing is also under the control of development and differentiation. However, a limited number of studies in human transformed cells indicate that replication timing was unexpectedly similar between different cell types. In fact, a microarray-based comparison of the human chromosome 22 between fibroblast and lymphoblastoid cells revealed that only 1% of this chromosome differed in replication timing (see Supplemental Table 4) (108).

A dynamic change of replication timing was first demonstrated in timing analyses of mouse embryonic stem (ES) cells before and after induction of differentiation (111). Genome-wide replication timing in various ES cells as well as that in induced pluripotent stem cells from fibroblast cells was identical, indicating that replication timing could be viewed as a stable epigenetic mark specific to an individual cell type. However, upon induction of differentiation into neural precursor cells in vitro, a rather dramatic reorganization of replication domains (changes in up to 20% of the entire genome) was observed. Differentiation of ES cells lead to the consolidation of smaller differentially replicating domains into larger coordinately replicated domains (111). Extensive replication timing changes were also reported in Drosophila Kc (of embryonic origin) and C18 (from wing imaginal discs) cell lines or in human cell lines (127, 133). These results clearly indicate that replication timing undergoes major changes during the course of development.

Another determinant of replication timing is the boundary that defines the transition between early and late replication domains. The timing boundary segments may define the transition of the timing by two potential mechanisms. One may slow down the fork progression rate so that the early replication fork may not reach the late segment before the latter segment is fired for initiation. Alternatively, it may represent the chromosome domain, which might physically and spatially separate different replication timing domains within nuclei. A recent genome-wide microarray analysis has demonstrated the existence of molecular boundaries between the coordinately replicating units of chromosomes (109, 111). The results indicate that timing transition regions define long segments without origins of unidirectional replication.

It was reported that the timing transition seems to correlate with the junction between the syntenic segments. Accumulations of single-nucleotide polymorphisms and causative mutations for various diseases were reported in the timing transition segment (134). This may be due to the presence of replication fork blockages in these regions, impeding fork progression and eventually resulting in DNA damage. This would lead to an enhanced mutation rate owing to errors in the NHEJ (nonhomologous DNA end joining) or error-prone lesion bypass DNA synthesis. These fragile sites scattered on the genome have been implicated in chromosome breakage, leading to genome rearrangement (135). Fragility increases in a mutant of mec1 from budding yeast or in the ATR-deficient cells of human (136, 137). These so-called fragile sites are now regarded as replication slow zones or fork pause sites of the eukaryotic genome. A potential link between the timing transition regions and fragile sites has also been reported (138). It is important to identify more precisely the timing
transition segments and dissect the molecular events that take place in these regions.

4. HOW?
How is the entire process of DNA replication operated and regulated?

4.1. Initiation of Eukaryotic DNA Replication

4.1.1. Assembly of the prereplicative complex and its regulation. The initiation of eukaryotic DNA replication is a precisely regulated event that requires the ordered assembly of multiple protein complexes at replication origins (for reviews see References 1 and 139).

The processes of replication initiation are divided into two steps (Figure 2). In the first step, the ORC, Cdc6, Cdt1, and MCM2-7 proteins are sequen-tially assembled on the chromosome at the late M to early G1 phase, generating a pre-RC. In the second step, pre-RCs are activated to generate active replication forks. At the G1-S transition, the activities of two kinases, the Cdc7-Dbf4 and the CDK, facilitate the loading of other essential replication proteins onto the pre-RC to activate the replicative DNA helicase and initiate chain elongation by DNA polymerases (99, 140–144).

Loading of the MCM complex onto DNA is referred to as DNA replication licensing. Upon entry into the S phase, multiple mechanisms ensure that no new pre-RC is formed, so that rereplication of any portion of the genome does not take place (139, 145). A major pathway for prevention of relicensing in metazoans is through regulation of Cdt1 activity by degradation or by its specific inhibitor, geminin. The degradation of Orc1 (146) and inhibition of reassociation of the MCM complex to chromatin may also be a part of the mechanisms for inhibition of relicensing. In yeast, inhibition of relicensing is achieved mostly through CDK-dependent mechanisms, including rapid CDK-dependent destruction of the Cdc6/Cdc18 protein during the S phase. After the loading of MCM2-7, the putative MCM loading machinery, comprising the ORC, Cdc6, and Cdt1, is no longer required for initiation of DNA replication, indicating that its primary function in DNA replication is to deliver MCM2-7 to the origins (147–149).

The ORC is an ATP-regulated DNA-binding protein that was first discovered in budding yeast (1, 2). The mammalian ORC consists of a stable core complex, composed of subunits ORC2 through ORC5, that interacts weakly with ORC1 and ORC6 (150). In human cell lines, ORC1 is tightly bound to the G1-phase chromatin and selectively degraded during S phase by ubiquitin-dependent proteolysis, and then it is resynthesized during the M-to-G1-phase transition (151).

The ATP-bound ORC binds to origin DNA, and then Cdc6, in a complex with ATP, binds to the ORC on DNA. The MCM2-7 and Cdt1, likely in a complex, join the ORC-Cdc6-DNA complex (the MCM complex is “associated” with chromatin and salt sensitive). ATP bound to Cdc6 is now hydrolyzed, and Cdc6 dissociates from the complex. At this step, dissociation of Cdt1 is also stimulated, which triggers stable binding of the MCM complex with DNA (the MCM complex is “loaded” onto chromatin and salt resistant) (152). ATP hydrolysis by the ORC is required for reiterative loading of the MCM complex, and ORC mutants defective in ATP hydrolysis cannot support pre-RC formation in vivo and are inviable (153). Electron microscopy studies showed that the ORC-Cdc6 complex forms a ring-shaped structure on DNA with dimensions similar to those of the ring-shaped MCM helicase (154). This implies that the ORC-Cdc6 ring binds the MCM ring to facilitate MCM helicase loading onto the origin (152).

Cdt1 plays a critical role in the licensing process. Loading of Cdt1 to chromatin is dependent on the ORC and Cdc6 (1, 155). Orc6 interacts with Cdt1 directly, thereby facilitating the loading of the MCM2-7 complex onto origin DNA (156). Cdt1 forms a complex with the MCM complex both in budding yeast and in mammalian cells (152, 157, 158) and plays a licensing role in the assembly
of the pre-RC. The level of Cdt1 protein oscillates during the cell cycle, from being high in the G1 phase, low in S phase, and high again at the M-G1 transition (159). Geminin is destabilized during the G1 phase but accumulates during the S, G2, and M phases of the cell cycle. Geminin directly binds to Cdt1 and prevents the loading of the MCM2-7 complex onto DNA (160–165). Cdt1 is degraded during the S phase by ubiquitination-dependent proteolysis in a CDK-dependent manner or through interaction with proliferating cell nuclear antigen (PCNA), and the remaining Cdt1 is inhibited by geminin. This double regulation ensures that Cdt1 is not present during the S phase, resulting in strict prevention of relicensing of replication origins (155, 166–171). Recent reports indicate a possible regulation of stability or cellular localization of Cdc6 and Cdt1 proteins by acetylation (172, 173). c-Myc, a frequently deregulated proto-oncogene, is required for DNA replication during the G1 phase after pre-RC formation and before firing, and its deregulation induces replication-dependent DNA damage (174). The precise roles of c-Myc in regulation of the pre-RC remain to be investigated. Phosphorylation of the N-terminal segment of MCM2 by the Cdc7 kinase was reported to be important for pre-RC formation during the quiescence to S phase transition (175), although it is not known whether this regulation operates during the G1 phase of proliferating cells.

Mechanistic studies regarding the initiation of DNA replication revealed striking conservation of the basic mechanisms through evolution (Supplemental Figure 1) (for more details, see the section Comparison of Initiation at Bacterial and Eukaryotic Origins in the Supplemental Material).

4.2. DNA Chain Elongation of Eukaryotic DNA Replication

4.2.1. Assembly of the replication fork and regulation. Conversion of the pre-RC to a replication fork complex (replisome) requires the association of additional factors. GINS (Go-Ichi-Ni-San (Go-Ichi-Ni-San represents the numbers 5, 1, 2, 3 in Japanese) for Sld5-Psf1-Psf2-Psf3) has been identified as a novel factor for the replisome (Supplemental Figure 2) (176–178). Its components were originally discovered through genetic material and Eukaryotic Origins in the section Comparison of Initiation at Bacterial and Eukaryotic Origins in the Supplemental Material).
screening for synthetic lethal mutations in combination with the dph11-1 mutant (dph11 is the Cut5-TopBP1 homolog) in budding yeast. They form a stable four-factor complex and are essential for DNA replication in both budding yeast and Xenopus egg extracts. GINS is necessary for stable engagement of Cdc45 with the nascent replisome and then associates stably with the MCM complex during S phase (179, 180). The crystal structure of a recombinant human GINS has been published, showing that it is a heterotetrameric complex (181–184). One report demonstrated that GINS does not contain a hole in the center and that the inner surface of the flat central cleft is unsuitable for DNA binding (181). Other reports showed that GINS has a ring-like structure with a small central channel (183, 184).

The replisome at the replication fork is assumed to be composed of many factors and to be involved not only in DNA replication but also in other chromosome transactions and chromatin regulation. This prediction was substantiated by the discovery of replisome progression complexes (RPCs) containing more than 20 replication-related proteins (185). In addition to the Cdc45, the MCM complex, and the GINS complex (referred to as CMG complex), the RPC contains Mrc1/Cla81p, Top1/Timelss, and Csm3/Swi3/Tipin (fork stabilizing factors). Sw1 and Sw3 from fission yeast were found to generate a stable complex termed FPC (fork protection complex) (186). Ctf1/Mc11/AND-1, and components of the histone chaperones including FACT, were also identified in the RPC. Other replication factors such as topoisomerase I and MCM10 are associated weakly with the RPC (185).

The human GINS complex interacts physically with DNA polymerase-primase α and markedly stimulates its polymerase function but not its priming function (187). In contrast, the archaeal GINS homolog (from Sulfolobus solfataricus) is composed of two subunits and forms a complex with the MCM helicase. It directly interacts with the heterodimeric core primase but has no effect on the primase activity. These results suggest that the GINS complex may be important in coordinating the progression of the MCM helicase and priming events on the replication fork (188). Also, GINS in fission yeast is necessary for chromatin binding of DNA polymerase ε but is not required for that of DNA polymerase α (189). The Pyrococcus furiosus GINS complex interacts with cognate MCM complex and stimulates its ATPase and DNA helicase activities in vitro (190). Recent reports indicate that the budding yeast Ctf1/Mc11/AND-1, required for efficient S-phase progression, interacts with GINS and DNA polymerase α and coordinates the progression of the MCM helicase and DNA polymerase α (191, 192). Thus, GINS-Ctf1/Mc11/AND-1 coordinates duplex unwinding and DNA chain elongation at the replication fork. For more recent information on DNA polymerases, the reader is encouraged to read the most recent reviews (193–195).

In budding yeast, Sld2 and Sld3, which are targets of CDK phosphorylation, are required for the initiation of DNA replication (196–198). Sld2 and Sld3 were isolated as sld mutants (199). The Dpb11/Cut5/TopBP1 was identified as a multicopy suppressor of mutations in DNA polymerase ε (200). Dpb11/Cut5/TopBP1 has two pairs of tandem BRCT domains and is required for checkpoint reactions (200–202). The C-terminal pair of the BRCT domains in Dpb11/Cut5/TopBP1 binds to the phosphorylated Sld2 (Thr 84), and the N-terminal pair binds to the phosphorylated Sld3 (Thr 600 and Ser 622) (197, 203). The preformed DNA polymerase ε-Sld2-Dpb11/Cut5/TopBP1-GINS complex may bind to the Sld3-Cdc45 complex on the pre-RC in a manner dependent on CDK-mediated Sld3 phosphorylation (H. Araki, personal communication). Furthermore, the phosphomimetic mutant of Sld2-T84D, in combination with a mutant Sld3 (S600A + S622A) fused to Dpb11/Cut5/TopBP1 or in combination with the cdc45/jet1 mutant, bypassed the CDK requirement for initiation of DNA replication, indicating that CDK-mediated phosphorylation of Sld2 and Sld3 is sufficient for initiation.

In budding yeast, Sld3 interacts with Cdc45, and both proteins associate with replication...
origins. Sld3 is required for interaction between the MCM complex and Cdc45 (204). Dpb11/Cut5/TopBP1, Sld3, Cdc45, and GINS assemble in a mutually dependent manner on replication origins to initiate DNA synthesis (176). After the establishment of DNA replication forks at early origins, Sld3 is no longer essential for the completion of chromosome replication, but Cdc45 and GINS associate stably with the MCM complex throughout the S phase (179). In contrast, in fission yeast, Sld3 functions in the initial phase of initiation complex assembly, followed by loading of GINS, Cut5/Dpb11, and then Cdc45 (142). The Hsk1 kinase, the fission yeast homolog of Cdc7, is required for chromatin loading of Sld3. In Xenopus extracts, Cdc7 must function before CDK for active DNA replication (205, 206). RecQ4L may be the metazoan homolog of Sld2, and it is also a target of CDK phosphorylation (207).

Although the human ortholog of Sld3 has not been identified yet, the interaction of replication proteins dependent on CDK-mediated phosphorylation might be well conserved. Recently, a vertebrate factor with functions potentially related to Sld3 were reported (208, 209).

4.2.2. Activation of replicative helicase.

Upon initiation of DNA replication in the S phase, the MCM complex moves away from replication origins as part of the DNA replication fork machinery (210–213). In the past decade, persuasive evidence has accumulated that suggests that the MCM2–7, a ring-shaped hexameric complex, is the replicative DNA helicase. Like most other replicative DNA helicas, the MCM2–7 complex consists of six subunits, containing highly conserved DNA-dependent ATPase motifs in their central regions (1, 214). The MCM proteins form several stable subcomplexes, including MCM2-3-4-5-6-7 (MCM2–7), MCM2-4-6-7, MCM4-6-7, and MCM3-5 (215–217). Biochemical characterization of the mouse MCM4-6-7 subcomplex revealed that it has intrinsic DNA-dependent ATPase and DNA helicase activities (215, 218). Later, helicase activity was reported in the MCM4-6-7 complex from other species as well (216, 217, 219). The activity and processivity of the MCM4-6-7 helicase can be highly stimulated by tailed substrate DNA (52, 216). The DNA helicase activity of the mouse MCM4-6-7 complex is highly enhanced also by the presence of thymine-rich ssDNA, which is often found near replication origins (52). Moreover, human FACT, a chromatin transcription factor, physically interacts with the MCM complex and promotes its DNA unwinding activity on forked nucleosomal templates in vitro (220).

Despite rather compelling genetic and biochemical evidence that the MCM complex is a replicative helicase at the fork, the isolated MCM2–7 complex does not show any detectable DNA helicase activity. Recently, however, such activity was detected in a purified recombinant budding yeast MCM2–7 complex, in a reaction mixture containing acetate and glutamate instead of chloride (221). It is yet to be determined whether the MCM2–7 complexes from other species exhibit DNA helicase activity under similar conditions.

As stated above, the CMG complex was identified (222), and evidence strongly suggests that both Cdc45 and GINS contribute to its helicase activity in vivo (210, 212, 213). Recently, the CMG helicase complex from Drosophila was reconstituted with purified components (223). CDK and Cdc7-Dbf4 are the two conserved protein kinases involved in activation of the MCM helicase in the pre-RC (99, 100, 214). Evidence indicates that the MCM2–7 complex is a target of phosphorylation by Cdc7, which facilitates the loading of the Cdc45 protein onto origins (140, 141). Cdc7 phosphorylates mainly the N-terminal tail segments of the MCM2, -4, or -6 proteins, and this phosphorylation is highly stimulated by prior phosphorylation by CDK in vitro (224). Cdc7 is an acidophilic kinase and favors the negatively charged segments generated by phosphorylation (225). Consistent with this finding, prior phosphorylation is required for Cdc7-mediated phosphorylation of in vitro assembled pre-RCs (144). The phosphorylated N-terminal tails of the MCM2, -4, or -6 proteins are recognized by Cdc45 or a complex containing Cdc45. These
phosphorylation events appear to be redundant, and combinations of different N-terminal tail mutations lead to defective growth in fission yeast (141). It was reported recently that phosphorylation of S170 of budding yeast MCM2, a target site of Cdc7, is required for cell growth (226). The formation of the pre-RC on a yeast replication origin in vitro leads to the assembly of a head-to-head double-hexameric MCM connected through the N-terminal rings on a duplex DNA (227, 228). It is tempting to speculate that Cdc7-mediated phosphorylation of the N-terminal tail region may separate the two hexamers and convert them into active helicase complexes.

In budding yeast, the mcm5-bob1 (carrying a proline to leucine substitution at position 83) bypasses Cdc7-Dbf4 (229) probably by conferring conformational change within the MCM complex, which would mimic the effect of Cdc7-mediated phosphorylation. The bob1 mutation also reduces origin efficiency, and it was suggested that the mutant adopts several conformations, only one of which is active for origin activation (230–232). More recently, another bypass of Cdc7 was reported (233). Deletion of the N-terminal nonconserved segment of MCM4 containing the Cdc7 phosphorylation sites showed bypass of Cdc7 function. It was proposed that the function of Cdc7 is to antagonize the inhibitory effect of this segment on initiation. Several studies have provided information on the MCM structure, and several models for its mode of action have been proposed. Novel MCM-related proteins have been reported also. See the Structural Basis of MCM and Its Mode of Action and MCM-Related Proteins in the Supplemental Material for more details.

4.3. G1 Regulation and DNA Replication

Proliferation of mammalian cells is regulated by external growth signals, which induce a series of G1 signal transduction pathways ultimately leading to initiation of DNA replication. Key cell cycle regulators during G1 phase are Rb-E2F, Cdk4 and -6, and CDK inhibitors (Figure 3).

The best known link of these G1 regulators to DNA replication is the transcriptional induction of many replication factors, including licensing factors (Cdc6, Mcm7, and others) as well as Cdc45 by E2F at G1-S transition. Presumably E2F-mediated expression of these replication factors facilitates the licensing and initiation phases of DNA replication. These events are also subject to negative regulation. For example, induction of the CDK inhibitor p16 Ink4a disrupts not only PCNA but also pre-RC assembly via a mechanism involving reduction of Cdc6 and Cdt1 expression. Interestingly, Cdc6 was reported to regulate the INK4 locus, suggesting a feedback loop involving mutually coordinated regulation of INK4 and Cdc6 (234).

G1 CDKs also play a direct role in regulation of replication factors. Cdc6 is stabilized during the G1 phase by CDK-mediated phosphorylation of serine 54, allowing a time window for pre-RC assembly (235, 236). This pathway is subject to negative control by DNA damage acquired during the G1 phase, which results in p53 activation, induction of p21, and inhibition of Cdk2. The resulting decreases in Cdk2-mediated phosphorylation of serine 54 lead to destabilization of Cdc6 and inhibition of replication licensing (236).

It is well established that the Rb-CDK signaling axis regulates DNA replication via transcriptional control and posttranslational modification of licensing and initiation factors. However, several studies suggest that Rb may also regulate DNA synthesis directly, independently of its transcriptional repression activities. For example, Rb can interact directly with MCM7 (237), and the Rb-E2F complex colocalizes with replication foci in mammalian cells and with the ORC in Drosophila (20, 238). Ectopic activation of Rb leads to disruption of PCNA function but not other replication factors (239). This transient inhibition of replication may play an important role in response to DNA damage by facilitating the transfer of freed PCNA to repair foci (240).
Figure 3
G1 regulation and DNA replication. In the transition from quiescence to growth, Geminin and Cdc6 are both degraded by APC/C. Cdc6, required for prereplicative complex (pre-RC) assembly, is transiently protected from degradation by Cdk2-CyclinE-mediated phosphorylation at Ser54, permitting pre-RC formation. Expression of Cdt1 and Cdc6, which is inhibited by p16INK4a (through Cdk4-CyclinD), as well as Cdc7-mediated phosphorylation of N-terminal segments of MCM2 are required for pre-RC formation. Cdc6 also regulates p16INK4a expression by interacting with its promoter region. DNA damage such as ionized irradiation induces p53, leading to inhibition of Cdk2-CyclinE activity by p21 (not shown in the figure). This leads to reduced phosphorylation of Ser54 in Cdc6 and reduced DNA replication. Pre-RC formation is required for expression of CyclinD and for activation phosphorylation of Thr160 of Cdk2. In the absence of a pre-RC, CDK activity is inhibited by p53-dependent loss of Thr160 phosphorylation and induction of p21 and p27 (inset). This is called a licensing checkpoint. Rb inhibits not only G1 regulators but also the S-phase replication machinery. Rb downregulates Cdk2-CyclinA. This in turn disrupts the proliferating cell nuclear antigen (PCNA) through an unknown mechanism (dissociation from chromatin). Cdc7 plays a crucial role in initiation of DNA replication by phosphorylating the N-terminal tails of MCM2, -4, and -6. Orange arrows indicate the phosphorylation reactions. Blue arrows indicate the degradation of the target proteins. Abbreviations: APC/C, anaphase promoting complex/cyclosome (E3 ubiquitin ligase); Cdc6, -7, cell division cycle 6, -7; Cdk2, cyclin-dependent kinase 2; Cdt1, cdc10-dependent transcripts 1; CycD and -E, CyclinD and -E; Dbf4, dumbbell former 4; G0-G1, quiescent to G1 phase; MCM2, -4, -6, minichromosome maintenance complex subunit 2, 4, 6; ORC1-6, origin recognition complex subunit 1-6; Rb, retinoblastoma.
Although G1 signaling via Rb and CDKs influences DNA replication, several recent studies also suggest that appropriate and timely activation of G1 CDKs and G1/S-phase progression is critically dependent on replication licensing. It was shown that impaired pre-RC assembly delays G1 phase progression via multiple mechanisms, including reduced transcriptional induction of Cyclin D1 (the activating binding partner of Cdk4/6), induction of CDK inhibitors p21 and p27, and p53-dependent loss of Thr160 phosphorylation of Cdk2 required for its activation (241–243). Further studies are necessary to determine the precise mechanisms of these novel licensing checkpoints.

These various mechanisms likely serve to integrate DNA replication with G1 to S-phase progression and may provide opportunities for dynamic and versatile regulation, thereby preventing progression into an aberrant S phase and/or facilitating the repair of damaged DNA. Interestingly, several studies have shown that licensing checkpoints are impaired in cancer cells, perhaps indicating that these novel checkpoints contribute to tumor suppression.

4.4. Detection of Fork Stalling and Cellular Responses to Stalled Replication Forks

The process of DNA replication, involving the separation of two DNA strands, inherently provokes a risky situation for the genomic DNA. Therefore, the chromosomes of S-phase cells are protected by layers of cellular strategies. Failure of these systems would immediately lead to chromosome instability and genetic alterations that could cause various diseases including malignancy (244).

It is known that oncogenic signals induce hyperproliferation, which may result in the generation of aberrant replication forks by misregulation of the replication licensing system or by direct hyperactivation of replication origins. This would lead to stalled replication forks, which may be dealt with by the cellular DNA damage response systems, inducing senescence or apoptosis in some occasions. If the DNA damage response system is compromised, cells would suffer from major genomic instability and would be converted to malignant tumors (245–247). Thus, the cellular mechanisms that deal with unexpected stalling of replication forks are crucial in protecting the genome from undergoing potentially lethal transformation.

4.4.1. Recognition of stalled replication forks. The progression of replication forks can be interfered with by external DNA insults and also by numerous internal factors. These include unusual secondary structures (repetitive sequences such as dinucleotide or trinucleotide repeats, specific G-rich nucleotide repeats, hairpins, quadruplexes, and triplexes), stalled transcription machinery, DNA-binding proteins or chromatin proteins blocking fork progression, or a reduced supply of nucleotides. In response to these obstacles, forks are stalled, or their rate of movement is reduced (136). The first step of the cellular response to stalled replication forks requires the recognition of such an event (see the sidebar titled Recognition of Stalled Replication Forks in Bacteria).
In eukaryotes, ssDNA generated at a stalled fork plays crucial roles in its cellular recognition. When forks are stalled, for example, by hydroxyurea, an inhibitor of ribonucleotide reductase, or by aphidicolin, an inhibitor of DNA polymerases, uncoupling of replicative helicase and DNA polymerases takes place, generating a ssDNA of sufficient length (248, 249). RPA binds to these ssDNA, and ATRIP in a complex with ATR brings this sensor/master kinase to the site of the fork stall (250).

However, the ssDNA-RPA complex is not sufficient for checkpoint activation. It was shown that the DNA synthesis function of DNA polymerase \( \alpha \) is also required for checkpoint signaling. A primer-template structure (5'-recessed DNA), which may be synthesized on a lagging strand by DNA polymerase \( \alpha \), is required for loading of the Rad9-Rad1-Hus1 clamp, facilitated by the Rad17-RFC clamp loader (250–252). Rad17 and Rad9 are phosphorylated by ATR, and the phosphorylated Rad17 and Rad9 recruit Claspin and TopBP1, respectively (250). Subsequently, TopBP1 interacts with ATR and ATRIP, stimulating the kinase activity of ATR (253).

Using Xenopus egg extracts, various DNA structures were tested for their ability to activate an ATR-dependent checkpoint pathway. The results indicate that a primed ssDNA of sufficient length is required for efficient activation (254). These results are consistent with the requirement of DNA polymerase \( \alpha \), which generates DNA primer-template structures.

### 4.4.2. Regulation of the S-phase checkpoint

Replication fork stalling in response to various forms of DNA damage activates an S-phase checkpoint, which is mediated via the ATR-Chk1 signaling pathway. The S-phase checkpoint elicits the following reactions: (a) stabilization of stalled replication forks, (b) inhibition of initiation at unfi red origins of replication, (c) inhibition of entry into G2 or M phase (through inactivation of Cdc25), and (d) slowing of ongoing fork progression.

\(<br />\text{mecl} \text{ (an ATR homolog) or rad53 (a Chk2 homolog, functionally equivalent to Chk1) mutations in budding yeast repress late origins in the presence of hydroxyurea (114, 115). In mammalian cells, a known target of Chk1 is the tyrosine phosphatase Cdc25A, which activates Cdk2-CyclinE to facilitate the loading of Cdc45 onto origins (253). Chk1 phosphorylates serine 123 of Cdc25A, targeting the phosphatase for Skp1-Cullin-F-box/\( \beta \)-transducin repeat-containing protein (SCF\( ^{\beta T R CP} \))-mediated proteolysis, leading to inactivation of Cdk2 and unloading of Cdc45 in response to fork blocks. However, recent results indicate that Chk1 also induces dissociation of Cdc45 from chromatin in a Cdc25A- and Cdk2-independent manner (256). Several studies have implicated Cdc7 as a target of the S-phase checkpoint. For example, it was shown using Xenopus egg extracts that etoposide induces dissociation of Cdc7 from Dbf4, thus inhibiting further firing of origins. Importantly, in those experiments, addition of Dbf4 counteracted etoposide-induced inhibition of origin firing, suggesting that Dbf4 is a target of the S-phase checkpoint (257). However, the putative mechanisms by which ATR/Chk1 signaling may negatively regulate Cdc7-Dbf4 in Xenopus are not known. Moreover, recent studies using cultured mammalian cells have shown that Cdc7 remains active in response to replication stress-inducing agents and is required for viability after genotoxin treatments (258). Therefore, it is not yet clear whether Cdc7 is a target of the intra-S checkpoint in all experimental systems (258–260).

Recent reports indicate that stalled replication forks not only block the initiation of late origins but also slow down the rate of fork movements (112, 261). How does the activated Chk1 slow down the fork? The Tof1\(^{Swi1/Timless} \) and Csm3\(^{Swi3/Tipin} \), and Mrcl\(^{Claspin} \) proteins, which are conserved replication fork factors, are required for checkpoint kinase activation (186, 262, 263, 264, 265). Mrcl\(^{Claspin} \) is hyperphosphorylated in a manner dependent on the Rad3\(^{Mec1/ATR} \) kinase in response to fork arrest (266, 267). In the mutant of mrc1 or tof1, extensive unwinding of DNA occurs in the presence of hydroxyurea, which is mediated by the
helicase components uncoupled from DNA synthesis (268). A similar extensively unwound DNA with a ssDNA tract of 500 to 800 bases was observed also in a rad53 mutant (249). Thus, these proteins may antagonize the helicase action at the replication fork, contributing to the stabilization of stalled forks as well as to reduction of fork rate. Indeed, replication fork slowing in response to UV requires Tipin (261).

Another possible mechanism for fork pausing could involve recombination. Recombination may antagonize the slowing of replication forks by inducing template switching, causing a quick bypass reaction in yeasts. The rpb1Δ (encoding a RecQ-type DNA helicase) mutant, which displays hyperrecombination, does not slow forks in response to fork stress (269). Thus, suppression of recombination is important for the slowing of replication forks in response to fork stress. Indeed, Mrc1Claspin was recently reported to be required for suppression of homologous recombination (270). By contrast, homologous recombination functions are required for fork slowing in vertebrates (271).

4.4.3. Roles of conserved fork stabilization factors for normal fork progression. Mrc1Claspin is required for efficient fork movement under normal growth conditions both in yeasts and mammalian cells (272–275). Slow S phase is not observed in mrc1-AQ, which is specifically defective in checkpoint responses due to the alanine substitutions of Rad3Mec1/ATR-target SQ/TQ sequences, indicating that Mrc1Claspin stimulates fork progression in a checkpoint-independent manner. Tof1Swi1/Timeless is also required for a normal DNA replication rate in human cells (261). The requirement of budding yeast Tof1Swi1/Timeless for normal DNA replication is controversial. In experiments using BrdU incorporation, Tof1Swi1/Timeless was shown to be required for a normal replication rate (276), whereas it is not required for a normal S-phase progression rate in density-shift experiments (273).}

Mrc1Claspin is required for recovery of DNA synthesis from hydroxyurea arrest but not from methylmethane sulfate arrest in budding yeast (276). This role also appears to be independent of checkpoint function. Tof1Swi1/Timeless is also required for fork recovery, but bulk replication is more efficiently restored in tof1Δ than in mrc1Δ, presumably because more efficient initiation/elongation takes place from late origins in tof1Δ during recovery. Tof1Swi1/Timeless but not Mrc1Claspin is required for fork arrest at rDNA, at tRNA, or at centromeres (273). In fission yeast, Swi1 is required to pause at rDNA or at mating-type loci (277, 278). The current consensus is (a) Mrc1Claspin and Tof1Swi1/Timeless proteins play distinct roles in the normal maintenance of the replication fork. (b) Mrc1Claspin is required for efficient fork progression, and this function is independent of its checkpoint function. (c) Tof1Swi1/Timeless is required for fork arrest at specific loci. Fork stalling at artificial inverted repeat sequences was enhanced in mrc1 or tof1 mutants, suggesting that Mrc1Claspin and Tof1Swi1/Timeless may be required to overcome the fork barrier owing to a secondary structure of DNA (279).

How does Mrc1Claspin stimulate the progression of replication fork? Mrc1Claspin interacts with fork factors, including Cdc45 and MCM. It can also bind preferentially to branched DNA in vitro (280, 281). These data suggest that Mrc1Claspin is an integral fork factor linking MCM and other replication factors. Recently, Mrc1Claspin was found to interact with DNA polymerase ε through both its N- and C-terminal sequences (282). This interaction occurs throughout the cell cycle and may serve to link a polymerase component to the helicase component. Indeed, Mrc1Claspin was shown to interact with the C-terminal tail of MCM6 through its central domain. Thus, the physical link of a polymerase to the helicase may facilitate the progression rate of the replication fork (283). Stimulation of the fork movement by coupling of a replicative polymerase and helicase was reported previously in E. coli (284). Upon hyperphosphorylation of Mrc1Claspin in response to hydroxyurea, the Mrc1N-DNA
polymerase ε interaction is specifically lost. The loss of this interaction may cause a conformational change in DNA polymerase ε for checkpoint signaling. Alternatively, the released N-terminal portion of Mrc1Chaos may interact with other fork components to stabilize the stalled fork (282). Recently, a novel concept for intra-S-phase checkpoint was proposed (285). This model argues that the checkpoint slows down replication in all the origins. See the Novel Concept of Intra-S Checkpoint Regulation in the Supplemental Material for a more detailed discussion.

4.5. Chromatin Regulation of DNA Replication

Chromatin assembly is an integral part of eukaryotic DNA replication. See Chromatin Regulation of DNA Replication in the Supplemental Material for a more detailed discussion.

4.6. The Interplay between DNA Replication and other Chromosome Transactions

Chromosome dynamics during cell cycle progression is regulated in a coordinated manner. Recent reports emphasized the intricate coupling of the processes of DNA replication and other chromosome transactions. See Interplays Between DNA Replication and Other Chromosome Transactions in the Supplemental Material for additional discussion.

5. BEYOND DNA REPLICATION

5.1. DNA Replication and Diseases

Ample evidence indicates that defects in the process of DNA replication could result in genomic instability, which could lead to mutations and abnormal tissue growth as observed in cancer. It can also cause a group of diseases known as “chromosome instability syndromes.” These diseases are in most cases predisposed to cancer because of various gene rearrangements or mutations. A hypomorphic mutation in MCM4 (MCM4Chaos) is known to cause severe genomic instability, and 80% of Chaos3 females die of mammary adenocarcinoma (286). Mutant mice expressing low levels of MCM2 exhibited severe deficiencies in proliferative cell compartments of various tissues and died at an early age owing to the development of lymphoma (287). The above defect may be due to a deficiency in a pre-RC formation. In contrast, deregulation of replication licensing also leads to G2/M checkpoint activation owing to head-to-tail collision of replication forks and generation of short-rereplicated DNA fragments (288). Cdt1 overproduction in rat or human fibroblasts induced chromosome aberrations without inducing rereplication (289, 290). Thus, perturbation of pre-RC formation, either inhibition or stimulation, can lead to DNA damage or a chromatin aberration, which could eventually lead to carcinogenesis. RecQL4, a potential mammalian homolog of Sld2, is responsible for Rothmund-Thomson syndrome, a rare autosomal recessive genodermatosis with achromosome fragility characterized by high incidences of skin and bone cancers (207). Thus, a defect or aberration in the processes of DNA replication can profoundly affect the maintenance of genomic integrity in mammals.

By contrast, normal and cancer cells respond to depletion of replication factors in significantly different manners (See Supplemental Table 5. Follow the Supplemental Material link from the Annual Reviews homepage at http://www.annualreviews.org). Generally, cancer cells are very sensitive to depletion of replication factors and undergo cell death, which occurs in both p53-positive and -negative cells (291). For examples, Cdc7 depletion in cancer cells results in abortive S-phase progression and cell death after abnormal mitosis. In contrast, normal cells arrest in G1 phase in a p53-dependent manner (292, 293). Cdc6 depletion causes an S-phase defect and induces cell death in cancer cells (294). In normal cells, Cdc6 depletion induces an ATR-dependent checkpoint and blockade of DNA replication, thus escaping cell death (295). Downregulation of the ORC or depletion of an
MCM subunit may induce a licensing checkpoint, resulting in inhibition of CDK activity and delayed S-phase entry (see Figure 3) (241, 242, 296). In contrast, cancer cells progress into a lethal S phase. Thus, replication factors may be promising targets for cancer therapy. (Please refer to Supplemental Table 5 for effects of depletion of replication factors in various mammalian cell lines.) The use of replication factors as biomarkers has also been explored because many of the replication factors are overexpressed in tumors and cancer cells (297).

Increasing evidence points to the role of replication timing in the regulation of chromatin structures and nuclear positioning, which may affect not only transcription but also other chromosome transactions. Aberrations in replication timing are associated with various types of human cancers (298, 299). A change in replication timing, presumably a switch from early to late S phase, may contribute to carcinogenesis by inducing silencing of tumor suppressor genes. In addition, certain chromosomes carrying translocations have replication timing defects that are accompanied by delays in mitotic chromosome condensation and an increased frequency of chromosomal instability. Mutations in Drosophila ORC subunits result in death at late larval stages, with defects in replication timing and chromosome condensation, as well as in sister chromatid cohesion, indicating a link between replication timing and chromosome condensation and segregation (300, 301). The regions of hypomethylation in ICF (immunodeficiency, centromeric heterochromatin, facial anomalies) syndrome (caused by mutations in the DNMT3B DNA methyltransferase gene) are associated with advanced replication timing and nuclease hypersensitivity (302). Some X chromosomes are hypomethylated but still silenced; escape from silencing is only seen with early replicating X chromosomes, suggesting that replication timing is a major determinant of gene silencing. Histone deacetylase 2 (HDAC2) is recruited to replication sites through DNMT1 (DNA methyltransferase 1) during late, but not early, S phase, thus raising the possibility that late replication maintains the heterochromatin state through HDAC2 (303). These findings suggest that modification of replication timing could be a potentially effective strategy in developing efficient therapies through reversal of the gene silencing. In fact, incorporation of an active replication origin in an expression vector prohibited the silencing frequently observed in the original vector (304).

5.2. The Cell Cycle and Other Biological Cycles

Recent reports suggest that cell cycle and checkpoint control may be coupled to the regulation of other biological cycles involved in setting the circadian rhythm and metabolism. See the Cell Cycle and Other Biological Cycles in the Supplemental Material for a more detailed discussion.

6. CONCLUDING REMARKS

During the past several years, several unexpected findings were reported on the where and when problems of eukaryotic DNA replication. The initiation sites are specific but not in the sense that applies to bacterial replicons. Each origin is much more inefficient, but its specificity is more relaxed, and initiation sites are selected from the list of equally active pre-RG sites. The choice may be made when the chromatin loops are generated in the preceding M phase. Alternatively, it may be determined during the ODP in the G1 phase when the majority of potential origins are rendered inactive.

The timing may be determined on large chromosome domains. This may be closely linked to the nuclear repositioning that may occur during the early G1 phase. It is still not clear at this stage exactly what kind of mark is left on the chromatin or what kind of higher chromatin structures are formed to differentiate the initiation sites and timing of replication during the S phase.

A remarkable finding, reported in the past years, was the dynamic coregulation of fork
rate and initiation at new origins. This significant degree of robustness of the replication program was indeed first reported in 1977, but it was largely unnoticed (57). It is now clear that cells respond to perturbation of DNA replication (e.g., reduced or enhanced fork rate or fork blocks) by altering the selection of potential origins to be used for initiation. When forks are stalled or significantly slowed down, potential origins (not used under normal conditions) or dormant origins are used to compensate for the deficiency. Obviously, cells are equipped with a system that monitors the fork rates and rapidly adjusting the locations of new initiation within the ongoing S phase. This could be achieved by changing the chromatin loop sizes so that replication is initiated with less or more intervals. Alternatively, it could be explained simply by the stochastic nature of the selection of active origins. Elucidation of the precise mechanism of coupling between fork rate and initiation frequency is an important issue that needs to be tackled in the near future.

Genome-wide analyses using microarrays provided the genomic landscape of replication origins and timing. In *Drosophila* and mammals, a strong correlation was observed between transcription and replication initiation sites. Most actively transcribed genes direct “divergent” transcripts, generating a nucleosome-free, negatively supercoiled state in the promoter region (305), which is favorable not only for transcription but also for DNA replication (pre-RC formation and initiation). This may be the reason why replication origins are abundant in the promoter segment. Localization of the pre-RC was examined genome wide in yeasts and *Drosophila*, and enrichment in the promoter region was observed, suggesting that the promoter regions provide a favorable chromatin environment for pre-RC assembly. It remains to be seen whether pre-RCs are assembled in a similar manner in mammals.

Investigation of replication timing with microarray analyses presented irrefutable evidence that replication timing is regulated during development. During the differentiation of ES cells, the timing changes (accompanied by nuclear repositioning) are observed within nearly 20% of the genome at the level of a large domain (600 kb) (111). Replication timing can be regarded at present as an epigenetic mark and might affect the expression profile (306).

Significant progress has been made in our understanding of the mechanisms of pre-RC formation. Currently, this process can be reconstituted with purified proteins both in *Xenopus* and budding yeast. The loading of the active MCM complex on chromatin requires the ORC, Cdc6, and Cdt1, and this process is under highly strict regulation to prevent reinitiation events within the single S phase. Efforts to reconstitute with purified proteins the steps following the pre-RC formation (i.e., loading of Cdc45 and initiation of DNA synthesis) on a defined template containing a specific origin sequence (e.g., budding or fission yeast origins or oriP from a model viral replicon) are in progress. The long-awaited origin-specific initiation of eukaryotic DNA replication in vitro is within our reach in the near future.

The replisome components are strikingly conserved across species. Newly conserved factors (GINS, Sld2, and Sld3) have been identified, and Sld2 and Sld3 have been shown to be critical targets of Cdk2 for assembly of the replisome. Isolation of the RPC indicates that many of these factors indeed generate a physically stable complex. Accumulating evidence suggests that the MCM complex is a central component of the replicative helicase. However, the MCM complex alone may not be sufficient to act as a replicative helicase. Identification of the helicase-active CMG complex suggests that the eukaryotic replicative helicase may be more complex than the prokaryotic counterpart. Additional biochemical studies of the helicase complex will not only clarify the mode of action of the MCM helicase at the fork but also eventually dissolve the MCM paradox. Other fork factors, such as Tof1/Timeless, Csm3/Tipin, and Mrc1/Claspin, may stabilize or slow down the fork upon encountering a replication stress signal...
by antagonizing the helicase action. Biochemical analyses of these proteins in the context of replication fork structures are needed.

A deficiency of replication factors can cause cancer or various other serious diseases collectively known as the chromosome instability syndrome. Yet, these factors could be cellular targets of novel cancer therapies. The unexpected link between cell cycle checkpoint and other biological cycles warrants further studies of replication factors at cellular and animal levels to elucidate the molecular basis of the potentially coordinated regulation of biological cycles.

**SUMMARY POINTS**

1. The ORC from metazoan (and even that from lower eukaryotes) binds to DNA with very little specificity, although it may generally prefer AT-rich sequences. The specificity may be provided by other factors, including transcription factors or chromatin structures.

2. Actual initiation sites are selected from many potential origins, determined by the pre-RC assembly present on the chromosomes, and the selection step is affected by many factors, including the distal regulatory elements, transcription state, topology, chromatin context, or the presence of nearby active origins.

3. The process of DNA replication shows high plasticity, with the fork rate and origin selection processes cross regulating each other, and dormant origins are activated when forks are stalled.

4. In metazoan, replication timing, although it correlates well with transcription, may be ultimately determined by the global chromatin context.

5. The two golden rules of eukaryotic DNA replication are “two-cycle engine” (pre-RC assembly and its activation occurring during mutually exclusive periods of cell cycle) and “once-per-cycle replication,” both of which are independent of where and when during the S phase replication is initiated.

6. Thus, stochastic or opportunistic initiation within a given replication domain is compatible with most of the known experimental data.

7. Replication stress induces a checkpoint, which represses both the fork progression rate and new origin firing.

8. A huge RPC containing the helicase, fork stabilizers, and histone chaperones can be isolated. The MCM complex is the central helicase component of RPC and may team up with GINS and Cdc45 to assemble an active helicase (the CMG complex).

**FUTURE ISSUES**

1. How is the site-specific initiation achieved? In other words, how are origins differentially selected from the list of pre-RCs? Alternatively, is this selection truly stochastic?

2. What are the exact mechanisms that determine the timing of origin firing? Are there any novel factors involved in defining the timing pattern of potential origins or setting the replication timing domains?
3. What are the mechanisms that coordinate fork rate and new origin firing during the S phase, and how are the stalled fork signals transmitted to activate dormant origins?

4. What is the nature of the active replicative DNA helicase at the eukaryotic replication forks and how does it operate?

5. What is the molecular architecture of the eukaryotic replication fork? Could assembly of a replication fork and initiation of DNA replication be reconstituted with purified proteins?

6. How is the replication fork reorganized upon encounter with fork blocks, and what are the exact mechanisms for inhibition of new firing, slowing the replication fork, and the eventual reinitiation of DNA replication?

7. How do aberrations in replication and checkpoint factors lead to various diseases, and what is the molecular basis of cancer cell–specific cell death induced by inhibition of replication factors?

8. How are the replication cycle and checkpoint regulation coordinated with the circadian cycle or metabolic cycle? Are there potential “extra-DNA replication” functions of replication factors?

**DISCLOSURE STATEMENT**

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

**ACKNOWLEDGMENTS**

The authors thank Dave Gilbert, Nick Rhind, Cyrus Vaziri, Anindya Dutta, and Hiroyuki Araki for critical reading of the manuscript. We also thank Yutaka Kanoh for the excellent artwork. We thank all the members of our laboratory for their hard work and helpful discussions. We also thank Ken-ichi Arai for guidance in the field of DNA replication and continued support throughout the course of the work conducted in the laboratory of H.M.

**LITERATURE CITED**


41. Burke TW, Cook JG, Asano M, Nevins JR. 2001. Replication factors MCM2 and ORC1 interact with the histone acetyltransferase HBO1. J. Biol. Chem. 276:15397–408


266. Yoo HY, Jeong SY, Dunphy WG. 2006. Site-specific phosphorylation of a checkpoint mediator protein controls its responses to different DNA structures. Genes Dev. 20:772–83


Contents

Preface

The Power of One
James E. Rothman ................................................................. v

Prefatory Article

Frontispiece
Aaron Klug ........................................................................... xiv

From Virus Structure to Chromatin: X-ray Diffraction
to Three-Dimensional Electron Microscopy
Aaron Klug ........................................................................... 1

Recent Advances in Biochemistry

Genomic Screening with RNAi: Results and Challenges
Stephanie Mohr, Chris Bakal, and Norbert Perrimon .................... 37

Nanomaterials Based on DNA
Nadrian C. Seeman ................................................................ 65

Eukaryotic Chromosome DNA Replication: Where, When, and How?
Hisao Masai, Seiji Matsumoto, Zhiying You, Naoko Yoshizawa-Sugata,
and Masako Oda .................................................................. 89

Regulators of the Cohesin Network
Bo Xiong and Jennifer L. Gerton ........................................... 131

Reversal of Histone Methylation: Biochemical and Molecular
Mechanisms of Histone Demethylases
Nima Mosammaparast and Yang Shi ........................................ 155

The Mechanism of Double-Strand DNA Break Repair by the
Nonhomologous DNA End-Joining Pathway
Michael R. Lieber .................................................................. 181

The Discovery of Zinc Fingers and Their Applications in Gene
Regulation and Genome Manipulation
Aaron Klug ........................................................................... 213
Origins of Specificity in Protein-DNA Recognition
Remo Rohs, Xiangshu Jin, Sean M. West, Rohit Joshi, Barry Honig, and Richard S. Mann .......................................................... 233

Transcript Elongation by RNA Polymerase II
Luke A. Selth, Stefan Sigurdsson, and Jesper Q. Svejstrup ........................................... 271

Biochemical Principles of Small RNA Pathways
Qinghua Liu and Zain Paroo ................................................................. 295

Functions and Regulation of RNA Editing by ADAR Deaminases
Kazuko Nishikura .................................................................................. 321

Regulation of mRNA Translation and Stability by microRNAs
Marc Robert Fabian, Nabam Sonenberg, and Witold Filipowicz .................................. 351

Structure and Dynamics of a Processive Brownian Motor:
The Translating Ribosome
Joachim Frank and Ruben L. Gonzalez, Jr. .................................................. 381

Adding New Chemistries to the Genetic Code
Chang C. Liu and Peter G. Schultz ................................................................ 413

Bacterial Nitric Oxide Synthases
Brian R. Crane, Jawahar Sudhamsu, and Bhumit A. Patel ......................................... 445

Enzyme Promiscuity: A Mechanistic and Evolutionary Perspective
Olga Khersonsky and Dan S. Tawfik ..................................................................... 471

Hydrogenases from Methanogenic Archaea, Nickel, a Novel Cofactor, and H2 Storage
Rudolf K. Thauer, Anne-Kristin Kaster, Meike Goenrich, Michael Schick, Takeshi Hiromoto, and Seigo Shima ........................................... 507

Copper Metallochaperones
Nigel J. Robinson and Dennis R. Winge .................................................................. 537

High-Throughput Metabolic Engineering: Advances in Small-Molecule Screening and Selection
Jeffrey A. Dietrich, Adrienne E. McKee, and Jay D. Keasling .................................... 563

Botulinum Neurotoxin: A Marvel of Protein Design
Mauricio Montal ....................................................................................... 591

Chemical Approaches to Glycobiology
Laura L. Kiessling and Rebecca A. Splain .................................................................. 619

Cellulosomes: Highly Efficient Nanomachines Designed to Deconstruct Plant Cell Wall Complex Carbohydrates
Carlos M.G.A. Fontes and Harry J. Gilbert ....................................................... 655
Somatic Mitochondrial DNA Mutations in Mammalian Aging
Nils-Göran Larsson ................................................................. 683

Physical Mechanisms of Signal Integration by WASP Family Proteins
Shae B. Padrick and Michael K. Rosen ....................................... 707

Amphipols, Nanodiscs, and Fluorinated Surfactants: Three
Nonconventional Approaches to Studying Membrane Proteins in
Aqueous Solutions
Jean-Luc Popot ........................................................................ 737

Protein Sorting Receptors in the Early Secretory Pathway
Julia Dancourt and Charles Barlowe ........................................ 777

Virus Entry by Endocytosis
Jason Mercer, Mario Schelhaas, and Ari Helenius ..................... 803

Indexes
Cumulative Index of Contributing Authors, Volumes 75–79 ........ 835
Cumulative Index of Chapter Titles, Volumes 75–79 ................. 839

Errata
An online log of corrections to Annual Review of Biochemistry articles may be found at
http://biochem.annualreviews.org