# Minireview

# Initiation of Chromosomal DNA Replication in Eukaryotes

## LESSONS FROM LAMBDA*

### Bruce Stillman

*From the Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724*

Compared with the exquisite understanding of the enzymology of DNA replication in bacteria, virtually nothing is known about the mechanism and regulation of chromosomal DNA replication in eukaryotes. Even a subject as fundamental as the sequence specificity of the initiation of eukaryotic DNA replication has been vigorously debated. This brief review explores selected aspects of the mechanism and regulation of the initiation of chromosomal DNA replication in eukaryotic cells. In particular, I will emphasize that lessons can be learned from comparing and contrasting the problem in eukaryotes with the well characterized prokaryotic replicons such as the *Escherichia coli* chromosome and bacteriophage lambda.

### The Lambda Lesson

The concept of positive control of the initiation of DNA replication emerged from the description of the replicon model by Jacob *et al.* (1). In this scheme, they proposed that a regulatory DNA element in the bacterial chromosome called the replicator would function as the target for a regulatory protein called the initiator and that this interaction would activate initiation of DNA replication of the entire chromosome. Many aspects of this model have been well demonstrated, not only for the initiation of bacterial chromosome replication but also for the replication of phage and plasmid DNAs and eukaryotic virus chromosomes (2–4). One noteworthy and particularly informative example is the assembly of a multiprotein complex at the bacteriophage lambda origin (oriλ) and the subsequent events that lead to initiation of DNA replication.

I have recently argued (5) that the genetically defined, cis-acting DNA sequences that regulate the initiation of DNA replication be called the replicator, whereas the location in the DNA of the actual start site(s) for initiation of DNA replication be termed the *ori*, for origin of DNA replication. This distinction is necessary because physical mapping methods used to locate the *ori* do not map the replicator and indeed, in some cases, the replicator may not overlap with the *ori*. The phage lambda case may be a good example. Electron microscopic mapping of DNA replication intermediates in phage DNAs from lambda-infected cells (6) or more recent mapping of the location of either bidirectional DNA unwinding or DNA replication *in vitro* (7) shows that the *ori* maps to a broad region in the general vicinity of the oriλ replicator, but clearly, in some cases the *ori* does not overlap with the replicator. These experimental observations could be due to technical artifacts such as branch migration during preparation of the DNA samples for microscopy or to different rates of fork progression from *ori*. Alternatively, it may be that initiation of DNA replication actually occurs in a broad region or initiation zone (8, 9) surrounding the replicator. I point this out because although there has been much discussion about the

nature of the DNA sequences required for initiation of DNA replication in eukaryotic chromosomes, it must be clear what has actually been measured. Physical mapping studies only locate the *ori*, whereas genetic experiments are needed to map the replicator.

In bacteriophage lambda, the phage-encoded initiator protein, the O protein, binds to the four O protein binding sites and forms a multiprotein complex at oriλ (Fig. 1 (10, 11)). Binding of O protein to oriλ in a supercoiled plasmid causes significant structural changes in the DNA, including the local unwinding of an A·T-rich region of the DNA that forms part of the replicator, adjacent to the O protein binding sites (12). This unwinding is ATP-independent for the lambda O protein, but ATP is required for the *E. coli* DnaA initiator protein binding to and causing structural changes in the DNA at oriC, the *E. coli* chromosomal replicator (13). In both cases, the assembled nucleoprotein complex then functions as a landing pad for a complex series of protein-protein and protein-DNA interactions.

Generally, the next critical stage in the initiation of DNA replication is the loading of a DNA helicase onto the DNA, which will eventually allow unwinding of the parental duplex DNA. The DnaB helicase in uninfected *E. coli* is normally bound to the DnaC protein, and these two proteins function in the replication of the bacterial chromosome (3). But in lambda phage-infected cells, the phage-encoded P protein hijacks the DnaB protein and facilitates the loading of the helicase onto oriλ via its interaction with the DNA-bound O protein (7, 10, 11, 14, 15). The oriλ·O·P·DnaB complex remains inactive for subsequent replication events because the P protein inhibits the activity of the DnaB helicase. Thus, it is necessary to either inactivate the P protein or remove it from the complex before the initiation of DNA replication can proceed.

Activation of the oriλ is facilitated by the combined action of three *E. coli* heat shock proteins, DnaK, DnaJ, and GrpE (16–18). They cooperate to remove the P protein from the oriλ·O·P·DnaB complex and liberate the DnaB helicase to unwind DNA in cooperation with the *E. coli* single-stranded DNA binding protein (SSB). Thus, the heat shock proteins orchestrate the assembly and disassembly of proteins into and subsequently out of the initiation complex and thus provide a crucial activation function.

Following activation of the DnaB helicase, the next major step is to load DNA primase onto the DNA to generate an RNA primer for subsequent DNA synthesis by DNA polymerase (3, 11). In *E. coli*, the primase interacts with the helicase and is not initially associated with the DNA polymerase (polymerase III holoenzyme). The DNA polymerase is then loaded onto the template DNA via the assembly of another multiprotein complex, including the polymerase accessory factors (*e.g.*, β, γ, δ, δ', and τ) (3, 19).

The well characterized protein-protein and protein-DNA interactions that lead to the initiation of DNA replication at oriλ demonstrate some general principles that should guide efforts to understand the initiation of chromosomal DNA replication in eukaryotic cells. At the same time, there are significant differences between the replication of bacteriophage DNA in *E. coli* and the complex process of progression through the S phase of the eukaryotic cell cycle.

### Replicators and Origins in Eukaryotes

*The Early Embryo*—Early studies on the sequence requirements for DNA replication in activated *Xenopus* eggs showed

oriλ
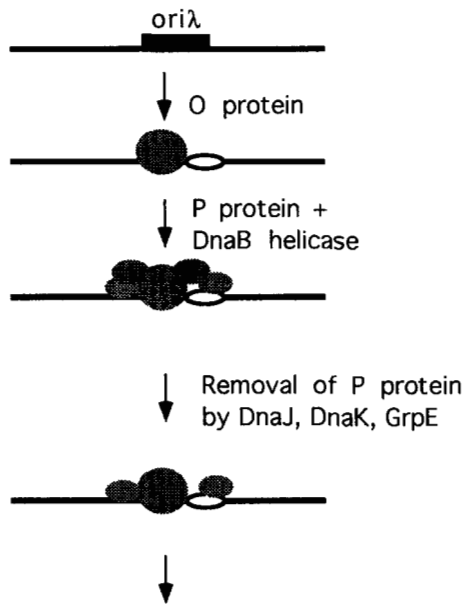


PRIMING AND INITIATION OF DNA REPLICATION

FIG. 1. **A simplified version of the initiation of DNA replication at the bacteriophage lambda origin.** The P protein, which is loaded onto the initiator protein and then is removed before priming of DNA replication, functions like a licensing factor.
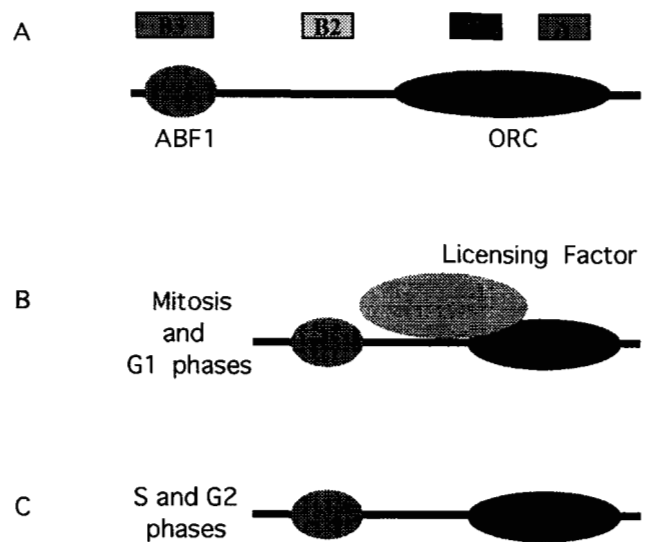


FIG. 2. *A,* the structure of the yeast chromosomal replicator *ARS1* with the essential A element and the three B elements. The regions bound by ABF1 and ORC are shown. *B* and *C,* speculative model for the regulation of DNA replication. A putative licensing factor, proposed by Blow and Laskey (51) and perhaps like the lambda P protein, binds to ORC in mitosis, remains there until it is removed at the G1 to S phase transition, and is reassembled at the next mitosis.

very convincingly that any DNA, even prokaryotic plasmids, could replicate in synchrony with the cellular chromosomes in the simple embryonic cell cycle (20). Equally importantly, the replication of the foreign DNA occurred in nuclear "factories" that resembled the sites of endogenous DNA replication (21). These experiments demonstrate that replication of extrachromosomal DNA in the activated egg does not require a replicator. They do not demonstrate that an *ori* is dispensable, and indeed, physical mapping studies have detected *oris* in plasmid DNAs replicated in a *Xenopus* egg extract (22). Thus, in the rapidly dividing cells in early embryos of *Xenopus* (and *Drosophila*), where the cell cycles lack the G1 and G2 phases, it is very likely that replicators are not required for DNA replication. This suggests that in the early embryo, the mechanism of assembling the DNA helicase and DNA primase onto the template DNA prior to initiation of DNA replication does not require an initiator protein.

*Yeast ARSs, Replicators, and oris*—In contrast to the studies in frog eggs, genetic studies in the yeast *Saccharomyces cerevisiae* demonstrated the existence of replicators in this eukaryote (reviewed in Refs. 8, 23, and 24). Specific DNA sequences derived from *S. cerevisiae* chromosomes, when cloned into plasmid vectors that contained a selectable marker, enabled the autonomous replication of the plasmid at each S phase of the cell cycle. Physical mapping techniques demonstrated that these genetically defined autonomously replicating sequences (ARS)[1] directed initiation of DNA replication, and in the cases studied, the *ori* mapped to the same general location on the chromosome as the replicator. Thus, initiation of DNA replication requires a replicator in the more conventional yeast cell cycle, which contains G1 and G2 phases and resembles the cell division cycle of mammalian somatic cells.

Only a subset of the ARS that support replication of plasmid DNAs in *S. cerevisiae* functions as replicators in their natural chromosomal location (*e.g.* Ref. 25). Thus, there are more ARS

elements in the chromosome than active chromosomal replicators. Furthermore, recent experiments have demonstrated that there is an excess of replicators in yeast chromosomes (26), a characteristic of eukaryotic chromosomes that was predicted long ago (27, 28). This suggests that there are more potential replicators in the chromosome than are actually needed, perhaps to allow for selected use of different replicators during development.

The structure of replicators in the yeast, *S. cerevisiae,* has been examined in several cases (24). Every ARS has a match to an 11-base pair, degenerate consensus sequence located in an essential element called the A element. Located adjacent to the A element is approximately 80 base pairs that constitute an essential B element. The B element contains DNA sequences that are easily unwound when present in a supercoiled plasmid DNA, and they may facilitate one stage of the initiation reaction (29, 30).

The B element in *ARS1,* a chromosomal replicator located near the centromere on chromosome IV of *S. cerevisiae,* has been studied in greatest detail (31, 32) (Fig. 2). The B region contains three DNA elements (B1, B2, and B3) that are individually not essential for ARS activity, but any two of which are sufficient to cooperate with the essential A element to support *ori* activity (32). The B elements have recently been shown to function in the chromosome as important replicator elements, as has the A element in a number of chromosomal replicators (30, 33, 34).[2] The B3 element at *ARS1* corresponds to the binding site for a transcription factor called ABF1 (32). ABF1 functions in yeast both as a positive transcription factor when bound to promoters of many yeast genes and also to silence transcription in a special context when bound at the silent mating type loci on chromosome III. For *ARS1* activity in a plasmid context, the ABF1 DNA binding site can be replaced with the binding site for a number of different DNA binding proteins that contain a transcriptional activation domain. Thus there is a direct link between the regulation of initiation of DNA replication and transcription initiation in eukaryotic chromosomes, just as has been observed with eukaryotic virus

---

[1] The abbreviations used are: ARS, autonomously replicating sequence(s); LCR, locus controlling region; ORC, origin recognition complex.

[2] Y. Marahrens and B. Stillman, submitted for publication.

replicons (35). The functions of the B1 and B2 elements are not known, but they may bind to other proteins or they may function as DNA unwinding elements, or both.

*Mammalian Replicators and oris*—The nature of replicators in mammalian cells has been more elusive. The topic has been extensively reviewed recently (8, 9, 36, 37), so I will only point out a few new facts. Attempts to demonstrate the existence of replicators in mammalian cells have been difficult. Various efforts to detect ARS-like elements by transfection of cells with plasmids containing cloned mammalian DNA sequences have revealed little sequence specificity, although some fragments were more efficient than others (38–40). It is possible, however, that the sequence specificity for the autonomous replication of plasmids in mammalian species is relaxed and that chromosomal replicators may not function in the plasmid context.

Definitive evidence for a chromosomal replicator was recently reported (41). Using two different physical mapping methods, an *ori* was located in the region 5' to the human β-globin gene, within the chromosome 11 globin gene locus. However, in a cell line derived from a patient with β-thalassemia in which the region upstream of the β-globin gene was deleted, an *ori* was no longer present in this region; rather, replication proceeded through the globin locus from outside. These data argue quite strongly for the existence of a genetically defined element (*i.e.* a replicator) that determines the location of an *ori* upstream of the β-globin gene. It is probable that all initiation of DNA replication in mammalian cells will require a replicator.

The β-globin replicon is of interest for another reason. Transcription of the globin genes in this locus is dependent on the prior activation of the entire locus via the locus controlling region (LCR). The LCR lies upstream of all the globin genes in the locus and seems to be the target for the initial activation (determination) of the locus during development. Interestingly, the LCR also influences the time in S phase when the locus is replicated (42). Although the LCR can regulate the timing of the replication of the locus, it is not part of the replicator since DNA replication initiates at the β-globin *ori* whether or not the LCR is active. Similarly, in the yeast, *S. cerevisiae*, replication timing is not necessarily a property of the replicator but is influenced by DNA sequences in the vicinity of the replicator (27). For example, the telomeres of *S. cerevisiae* can silence transcription of genes located nearby and influence replication timing of flanking DNA (43).

### A Putative Initiator Protein in Eukaryotic Cells

In yeast and mammalian cells, but perhaps not in early embryo cells of *Xenopus*, the lambda example for the role of a replicator in initiation of DNA replication may apply. If a replicator exists, then the lambda and *E. coli* archetypes would suggest that an initiator protein would interact with the replicator and play a key role in the initiation of DNA replication. There have been many searches for the elusive initiator, but only one strong candidate protein has emerged (44). A multisubunit protein was identified and purified from extracts of *S. cerevisiae* cells that bound to the essential A element of the *ARS1* replicator. The activity, called the origin recognition complex (ORC), contains six different polypeptides of apparent molecular masses of 120, 72, 62, 56, 53, and 50 kDa that are bound together. The binding of ORC to all the ARS elements tested requires a functional A element but not specific DNA sequences in any of the B elements. ORC induces a periodic pattern of DNase I hyper- and hyposensitive sites in the B element region of ARS sequences that suggests that the protein may wrap the DNA around the ORC polypeptides, much like a histone octamer wraps DNA in the nucleosome. Most unusually, ORC binds to the DNA in an ATP-dependent manner. As noted above, ATP

is known to play a critical role in the function of the DnaA initiator protein at oriC in *E. coli* (3, 13). Thus ORC had certain characteristics of the prokaryotic initiators that suggested it as a strong candidate for the eukaryotic cell initiator protein.

A role for ORC *in vivo* was also suggested by the finding of a DNase I protection pattern in the chromatin at the *ARS1* locus in permeabilized cells that was very similar to the pattern obtained with purified ORC *in vitro* (44, 45). The fact that a nuclease protection pattern was found in the chromatin isolated from asynchronously growing cells suggested that ORC might be bound to the replicator at most, and perhaps all, stages of the cell cycle. If this is true, then regulation of initiation of DNA replication at the G1 to S phase transition is not simply the sudden binding of the initiator protein, but rather, ORC probably acts as a landing pad (like lambda O protein) for other proteins that would cooperate to initiate DNA replication (46).

More direct evidence for the role of ORC as an initiator protein *in vivo* has emerged for recent genetic experiments (47–50). The *ORC2* and *ORC6* genes, encoding the 72- and 50-kDa ORC polypeptides, are essential for cell cycle progression, and in both cases, genetic studies of DNA binding have indicated that ORC binds to the A element of the replicator *in vivo* (49, 50). Analysis of a temperature-sensitive mutant of the *ORC2* gene has provided further evidence for a role for ORC in DNA replication *in vivo* (47, 48). Compared with wild type, strains carrying the ORC2$^{ts}$ protein lose plasmids at a high frequency, probably due to a failure of the plasmids to replicate. Similarly, analysis of the replication of the chromosomal DNA in cells grown at the non-permissive temperature demonstrated a defect in the entry of cells into S phase. These experiments, which demonstrate a positive role for ORC in the initiation of DNA replication, also suggest that ORC may play a role in preventing inappropriate DNA replication at the G1 to S phase transition.

### Licensing of DNA Replication

Based upon experiments on the replication of DNA using a cell-free extract prepared from *Xenopus* eggs, Blow and Laskey (51) proposed a model for the activation of DNA replication at the G1 to S phase transition, which also posits a mechanism for how DNA replication is normally restricted to once per cell cycle. They proposed that a protein, called the licensing factor, enters the nucleus during mitosis and functions to permit DNA replication at the appropriate time in the next S phase. They further suggested that the licensing factor would be degraded or inactivated in the S phase, thereby preventing re-replication until it was licensed at the next mitosis. This is an attractive model that has gained some experimental support (52, 53); however, the molecular nature of licensing factor is at present not understood.

In many ways, the phage lambda P protein fits the criteria for a DNA replication licensing factor. The protein is necessary for the initiation of DNA replication, and indeed it is loaded onto the lambda O initiator protein. But equally important is the fact that the P protein must be removed from the initiation complex before DNA replication can proceed. Thus the P protein would license DNA replication, but a new round of DNA replication would only occur if new P protein was available. Of course, in lambda-infected cells, the P protein is omnipresent, and the phage replicates its DNA many times during lytic infection. Nevertheless, if a protein like the P protein had restricted access to the initiator protein and there was a time delay between the loading of the protein onto the initiator protein and its removal, then a mechanism of temporal control of DNA replication could be envisioned.

One candidate protein for the eukaryotic cell licensing factor

is the CDC46 protein and its relatives in the yeast *S. cerevisiae* (54–57). This class of proteins, which also includes the MCM2 and MCM3 proteins, was identified genetically as participants in yeast DNA replication. Notably, the cell cycle location of the proteins varies throughout the cell cycle, consistent with an expected licensing factor. The yeast proteins enter the nucleus at mitosis, remain there through the G1 phase, and then disappear from the nucleus at the G1 to S phase transition. Recent genetic data (49) suggest that ORC and the CDC46 protein might interact, providing circumstantial support for this model.

There are homologues of these proteins in other yeasts and in mammalian cells, but in the latter case, the protein appears to be constitutively nuclear throughout the cell cycle. It is possible to imagine, however, how the activity of a nuclear protein might be regulated in a cell cycle-dependent manner and satisfy the licensing factor model. Although this model is at the moment speculative, future work with eukaryotic cell initiator proteins and taking heed of the lessons from prokaryotes may resolve the complex process of initiation of chromosomal DNA replication.

## REFERENCES

1. Jacob, F., Brenner, S., and Cuzin, F. (1963) *Cold Spring Harbor Symp. Quant. Biol.* **28**, 329–348
2. Challberg, M. D., and Kelly, T. J. (1989) *Annu. Rev. Biochem.* **58**, 671–717
3. Kornberg, A., and Baker, T. A. (1991) *DNA Replication*, W. H. Freeman and Co., New York
4. Stillman, B. (1989) *Annu. Rev. Cell Biol.* **5**, 197–245
5. Stillman, B. (1993) *Nature* **366**, 506–507
6. Schnos, M., and Inman, R. B. (1970) *J. Mol. Biol.* **51**, 61–73
7. Mensa-Wilmot, K., Seaby, R., Alfano, C., Wold, M. S., Gomes, B., and McMacken, R. (1989) *J. Biol. Chem.* **264**, 2853–2861
8. DePamphilis, M. L. (1993) *Annu. Rev. Biochem.* **62**, 29–63
9. Hamlin, J. L. (1992) *BioEssays* **14**, 651–658
10. Echols, H. (1990) *J. Biol. Chem.* **265**, 14697–14700
11. McMacken, R. (1992) in *DNA Replication and the Cell Cycle* (Fanning, E., Knippers, R., and Winnacker, E., eds) pp. 35–49, Springer-Verlag, Berlin
12. Schnos, M., Zahn, K., Inman, R. B., and Blattner, F. R. (1988) *Cell* **52**, 385–395
13. Bramhill, D., and Kornberg, A. (1988) *Cell* **52**, 743–755
14. Dodson, M., McMacken, R., and Echols, H. (1989) *J. Biol. Chem.* **264**, 10719–10725
15. Liberek, K., Osipiuk, J., Zylicz, M., Ang, D., Skorko, J., and Georgopoulos, C. (1990) *J. Biol. Chem.* **265**, 3022–3029
16. Osipiuk, J., Georgopoulos, C., and Zylicz, M. (1993) *J. Biol. Chem.* **268**, 4821–4827
17. Alfano, C., and McMacken, R. (1989) *J. Biol. Chem.* **264**, 10709–10718
18. Wyman, C., Vasilikiotis, C., Ang, D., Georgopoulos, C., and Echols, H. (1993) *J. Biol. Chem.* **268**, 25192–25196
19. Friedman, J. M. (1992) *Curr. Biol.* **2**, 296–298
20. Harland, R. M., and Laskey, R. A. (1980) *Cell* **21**, 761–771
21. Cox, L. S., and Laskey, R. A. (1991) *Cell* **66**, 271–275
22. Mahbubani, H. M., Paull, T., Elder, J. K., and Blow, J. J. (1992) *Nucleic Acids Res.* **20**, 1457–1462
23. Fangman, W. L., and Brewer, B. J. (1991) *Annu. Rev. Cell Biol.* **7**, 375–402
24. Campbell, J. L., and Newlon, C. S. (1991) in *The Molecular and Cellular Biology of the Yeast Saccharomyces* (Broach, J. R., Pringle, J. R., and Jones, E. W., eds) pp. 41–146, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
25. Dubey, D. D., Davis, L. R., Greenfeder, S. A., Ong, L. Y., Zhu, J., Broach, J. R., Newlon, C. S., and Huberman, J. A. (1991) *Mol. Cell. Biol.* **11**, 5346–5355
26. Dershowitz, A., and Newlon, C. S. (1993) *Mol. Cell. Biol.* **13**, 391–398
27. Fangman, W., and Brewer, B. (1992) *Cell* **71**, 363–366
28. Callan, H. G. (1973) *Cold Spring Harbor Symp. Quant. Biol.* **38**, 195–203
29. Natale, D. A., Umek, R. M., and Kowalski, D. (1993) *Nucleic Acids Res.* **21**, 555–560
30. Huang, R.-Y., and Kowalski, D. (1993) *EMBO J.* **12**, 4521–4531
31. Celniker, S. E., Sweder, K., Spienc, F., Bailey, J. E., and Campbell, J. L. (1984) *Mol. Cell. Biol.* **4**, 2455–2466
32. Marahrens, Y., and Stillman, B. (1992) *Science* **255**, 817–823
33. Rivier, D. H., and Rine, J. (1992) *Science* **256**, 659–663
34. Deshpande, A., and Newlon, C. S. (1992) *Mol. Cell. Biol.* **12**, 4305–4313
35. DePamphilis, M. L. (1993) *Trends Cell Biol.* **3**, 161–167
36. DePamphilis, M. L. (1993) *J. Biol. Chem.* **268**, 1–4
37. Benbow, R. M., Zhao, J., and Larson, D. D. (1992) *BioEssays* **14**, 661–669
38. Masukata, H., Satoh, H., Obuse, C., and Okazaki, T. (1993) *Mol. Biol. Cell* **4**, 1121–1132
39. Krysan, P. J., and Calos, M. P. (1991) *Mol. Cell. Biol.* **11**, 1464–1472
40. Krysan, P. J., Smith, J. G., and Calos, M. P. (1993) *Mol. Cell. Biol.* **13**, 2688–2696
41. Kitsberg, D., Selig, S., Keshet, I., and Cedar, H. (1993) *Nature* **366**, 588–590
42. Forrester, W. C., Epner, E., Driscoll, M. C., Enver, T., Brice, M., Papayannopoulou, T., and Groudine, M. (1990) *Genes & Dev.* **4**, 1637–1649
43. Ferguson, B. M., and Fangman, W. L. (1992) *Cell* **68**, 333–339
44. Bell, S. P., and Stillman, B. (1992) *Nature* **357**, 128–134
45. Diffley, J. F. X., and Cocker, J. H. (1992) *Nature* **357**, 169–172
46. Estes, H., Robinson, B., and Eisenberg, S. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 11156–11160
47. Bell, S. P., Kobayashi, R., and Stillman, B. (1993) *Science* **262**, 1844–1870
48. Foss, M., McNally, F. J., Laurenson, P., and Rine, J. (1993) *Science* **262**, 1838–1844
49. Li, J. J., and Herskowitz, I. (1993) *Science* **262**, 1870–1874
50. Micklem, G., Rowley, A., Harwood, J., Nasmyth, K., and Diffley, J. F. X. (1993) *Nature* **366**, 87–89
51. Blow, J. J., and Laskey, R. A. (1988) *Nature* **332**, 546–548
52. Blow, J. J. (1993) *J. Cell Biol.* **122**, 993–1002
53. Coverley, D., Downes, C. S., Romanowski, P., and Laskey, R. A. (1993) *J. Cell Biol.* **122**, 985–992
54. Chen, Y., Hennessy, K., Botstein, D., and Tye, B. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 10459–10463
55. Hennessy, K. M., Clark, C. D., and Botstein, D. (1990) *Genes & Dev.* **4**, 2252–2263
56. Hennessy, K. M., Lee, A., Chen, E., and Botstein, D. (1991) *Genes & Dev.* **5**, 958–969
57. Yan, H., Merchant, A. M., and Tye, B. K. (1993) *Genes & Dev.* **7**, 2149–2160