

## Minireview

## Initiation of DNA replication in eukaryotes: questioning the origin

Patricia Françon, Domenico Maiorano, Marcel Méchali\*

Institute of Human Genetics, CNRS, Genome Dynamics and Development, 141 Rue de la Cardonille, 34 396 Montpellier Cedex 05, France

Received 16 April 1999

**Abstract** Although proteins involved in DNA replication in yeast have counterparts in multicellular organisms, the definition of an origin of DNA replication and its control in higher eukaryotes might obey to different rules. Origins of DNA replication that are site-specific have been found, supporting the notion that specific DNA regions are used to initiate DNA synthesis along metazoan chromosomes. However, the notion that specific sequences will define origins is still being debated. The variety and complexity of transcriptional programs that have to be regulated in multicellular organisms may impose a plasticity that would not be compatible with a fixed origin simply defined at the sequence level. Such a plasticity would be essential to developmental programs where the control of DNA replication could be more integrated to the control of gene expression than in unicellular eukaryotes.

© 1999 Federation of European Biochemical Societies.

**Key words:** DNA replication; Cell cycle; Origin; Transcription

## 1. The origin of DNA replication

The initiation of genome replication in multicellular organisms remains an exciting and unresolved issue. Chromosomal replication requires the coordination of multiple initiations within the cell cycle and feedback mechanisms to ensure that DNA is replicated once and only once at each cell division. The replicon model aimed to explain the replication of the *Escherichia coli* chromosome, proposed a genetic element, the replicator, as a target sequence recognized by a positive specific protein, called the initiator [1]. 36 years later, investigations carried out on DNA replication of bacteria, phages, eukaryotic DNA viruses and, at least to some extent, on the *Saccharomyces cerevisiae* genome are in agreement with the original model. To which extent origins of DNA replication in multicellular eukaryotes are defined according to the replicon rules is still being debated.

## 2. Origin recognition sequences

The best eukaryotic origin characterized to date is the autonomously replicating sequence (ARS) of the yeast *S. cerevisiae*. ARSs were isolated for their ability to promote extrachromosomal maintenance of plasmids [2]. This 200 bp origin is made up of an A element and of two or three B adjacent elements [3]. The A element contains a conserved AT rich consensus sequence (thereby its name ACS, ARS

consensus sequence). Although B elements are not conserved, they are essential for the origin function and contain binding sites for a transcription factor or DNA unwinding elements (DUE).

Thanks to the identification of the consensus target sequence element, a second important discovery was the detection and identification of factors which interact with this sequence, the origin recognition complex (ORC) proteins, and allow the formation of the initiation complex [4]. Thus, at a first glance, initiation of DNA replication appeared to obey the rules conserved 'from bacteria to the elephant', like many basic cellular functions. However, explorations of the anatomy of an origin in multicellular organisms led to consider that things could not be so simple. Although the notion of site-specific origins has now emerged for several chromatin domains, numerous and intense studies trying to identify a clear consensus sequence in metazoan have failed up to date.

The definition of a specific sequence element acting as an origin of DNA replication, confirmed in *S. cerevisiae*, cannot yet be generalized in the unicellular world. Autonomously replicating elements can be found in *Schizosaccharomyces pombe*, as in *S. cerevisiae*. However, in *S. pombe*, autonomously replicating sequences have been found difficult to define at the nucleotide level, as found in *S. cerevisiae*. Origins in *S. pombe* appear to cover large initiation areas organized in clusters of A-T rich elements without a clear consensus sequence [5,6]. Several genetic and biochemical features of *S. pombe* are more similar to multicellular organisms than *S. cerevisiae* and the organization of the two genomes for their replication may be as divergent as these two organisms are in evolution.

## 3. Origins in multicellular organisms

Experimental assays aimed at identifying origins in human cells have been difficult to design. Extrachromosomal DNA elements replicate very poorly in human or vertebrate cells, precluding the identification of efficient autonomously replicating elements. In absence of a suitable experimental assay for origins, the development of new methods to detect and map origins of DNA replication has been crucial. These methods have been only recently improved to map initiation events at a few base pair level [7], but still require large cell populations as starting material. The first mammalian origin mapped was the origin found in the DHFR locus [8]. This origin has been an experimental field for each of the methods aiming to map origins, as soon as they were available. Two preferential sites were characterized, ori (origin of replication)  $\beta$ , studying the polarity of replication forks [9], and ori  $\gamma$  in quantifying nascent DNA strands [10], with 80% of initiation events emanating from a 500 bp region surrounding ori  $\beta$ . However, by

\*Corresponding author. Fax: (33) (04) 99 61 9920.  
E-mail: mechali@igh.cnrs.fr

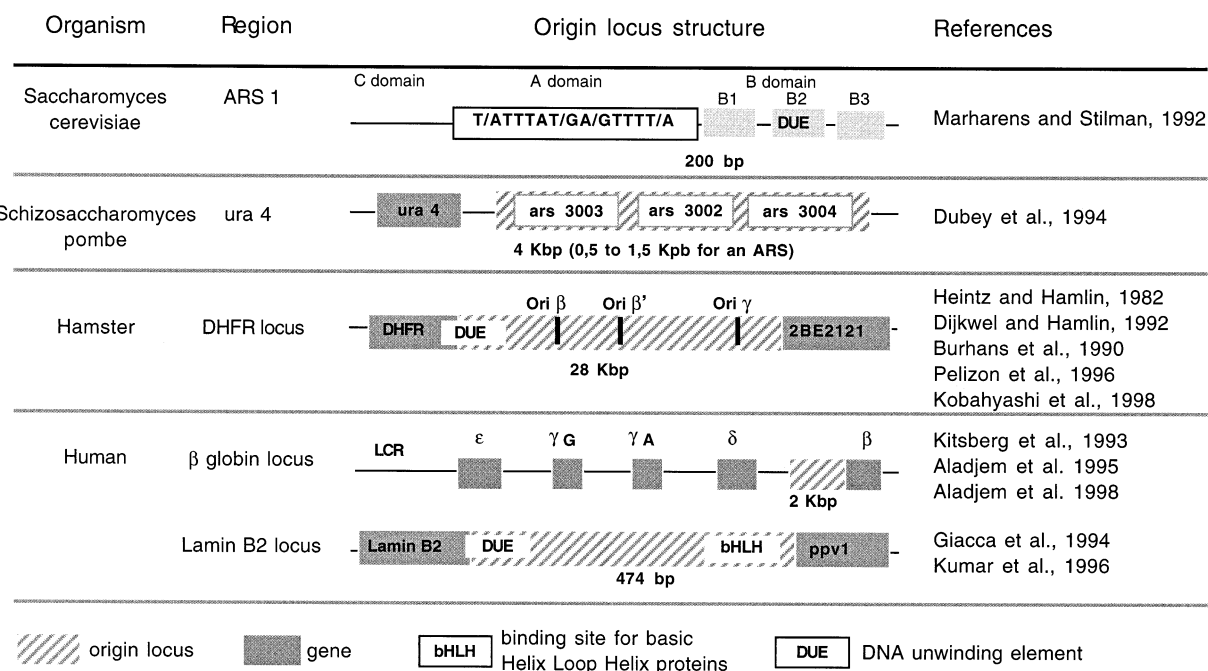


Fig. 1. Structure of eukaryotic origin loci. Four origin regions characterized in yeast or vertebrate cells are shown.

using neutral-neutral two-dimensional gel electrophoresis, initiation is observed over a 55 kbp region lying between the DHFR and 2BE2121 genes [11] and deletion of ori  $\beta$  resulted in an unchanged efficiency and timing of initiation in the corresponding domain [12]. Recently, a new origin was mapped in this locus, 5 kbp downstream of ori  $\beta$ , ori  $\beta'$  [13]. Despite these results that often appear controversial, it is clear that a specific region of initiation exists in this locus that is prone to select origins, some of them being more active than others. The chromatin organization in this region and its transcriptional status may influence the choice of origin.

The origin found in the lamin B2 gene of HeLa cells was the first mammalian origin characterized using the competitive PCR method [14]. This 747 bp origin is functional in a variety of human cells [15] and appears therefore ubiquitous. It lies in a constitutively expressed gene domain, like many origins mapped so far, and protein complexes appear to associate with the origin in a cell cycle-dependent manner [16].

At present, some 20 origins have been identified in complex eukaryotes (Fig. 1). DNA unwinding elements, defined as sequences that can be easily opened, AT rich regions, are often found at the proximity of several origin loci. However, an identikit of an origin is still difficult to draw up and is certainly not sufficient to design a DNA element that can function as an origin. Methylation sites could be important for positive or feedback controls of initiation of DNA replication, but no clear common rules for their presence around origins has yet been defined.

#### 4. Initiation of DNA replication and the control of gene expression

The lack of a specific consensus sequence to define an origin is rather disconcerting. One possibility is that a simple consensus does not exist and that different consensus exist for different origin families, like different origin sequences exist

for different DNA phages or DNA viruses. A second possibility is that origins are dictated by the structural organization of chromatin domains. The regulation of cell division in multicellular organisms imposes an additional level of complexity. Cells divide not only to reproduce themselves but also to ensure the formation of tissues, e.g. specific and specialized cellular populations of defined sizes and forms.

Some evidence for a link between transcription and DNA replication in setting up an origin has been reported. In human cells, Kitsberg et al [17] first observed that DNA replication in the human  $\beta$  globin domain is initiated at a 2 kbp fixed origin locus in the  $\beta$  globin gene promoter, in different kind of cells. An upstream region containing an LCR that controls the transcription of the locus is responsible for the early replication of the gene. If this region is deleted, transcription of the locus is repressed as well as the origin of DNA replication [18]. The LCR may play a role on the decondensation of chromatin when transcription is activated, allowing the binding of initiator proteins. However, the  $\beta$  globin origin can function in an ectopic location without the LCR region. The initiation region, that contains an antibiotic resistance gene actively transcribed, may substitute the LCR function in the ectopic site [19]. These studies first emphasize that this origin contains a genetic determinant for origin function and after, suggest that the organization of the corresponding domain for transcription may be a major issue.

A regulated localization of DNA replication origins also appears at the onset of transcriptional programs during development. In *Xenopus*, initiation occurs at apparently random sites in the rDNA domain during early embryogenesis, when cells divide at a high rate without transcription. A major transition occurs after the 13th cell cycle (mid-blastula transition) when genes become transcriptionally active [20] and when origins of DNA replication become localized in the intergenic spacer [21]. A similar phenomenon is observed in the DNA polymerase- $\alpha$  locus of *Drosophila* embryos, when ori-

gins become localized in a specific manner only after the onset of transcriptional programs [22]. In the mouse embryo, the replication of *Polyoma* virus DNA does not require enhancers before the two-cell stage when transcription is repressed, whereas enhancers are required both for replication and transcription after the two-cell stage, when transcription resumes in the embryo ([23], for review).

Molecular mechanisms that connect transcription and replication remain an open issue. The link could be between the transcription and replication machinery. It may be direct or indirect through factors involved in chromatin decondensation, important for both the loading of transcription and replication proteins. How DNA replication initiates in inactive genomic domains, is still unknown.

## 5. The initiation complex: from yeast to metazoans

Proteins involved in the initiation of DNA replication have been, in most cases, isolated and first characterized in yeast thanks to autonomous replication of ARS containing plasmids and the use of genetics. DNA replication in *Xenopus* in vitro systems proved to be a useful tool to analyze up to which extent the *Xenopus* homologues of the yeast proteins play a similar role in vertebrates and to unravel their biochemical function. Proteins implicated in the early steps of the initiation of DNA replication in yeast appear to be conserved from yeast to humans, but the relatively low level of identity observed (20% for the ORC subunits and 25–30% for

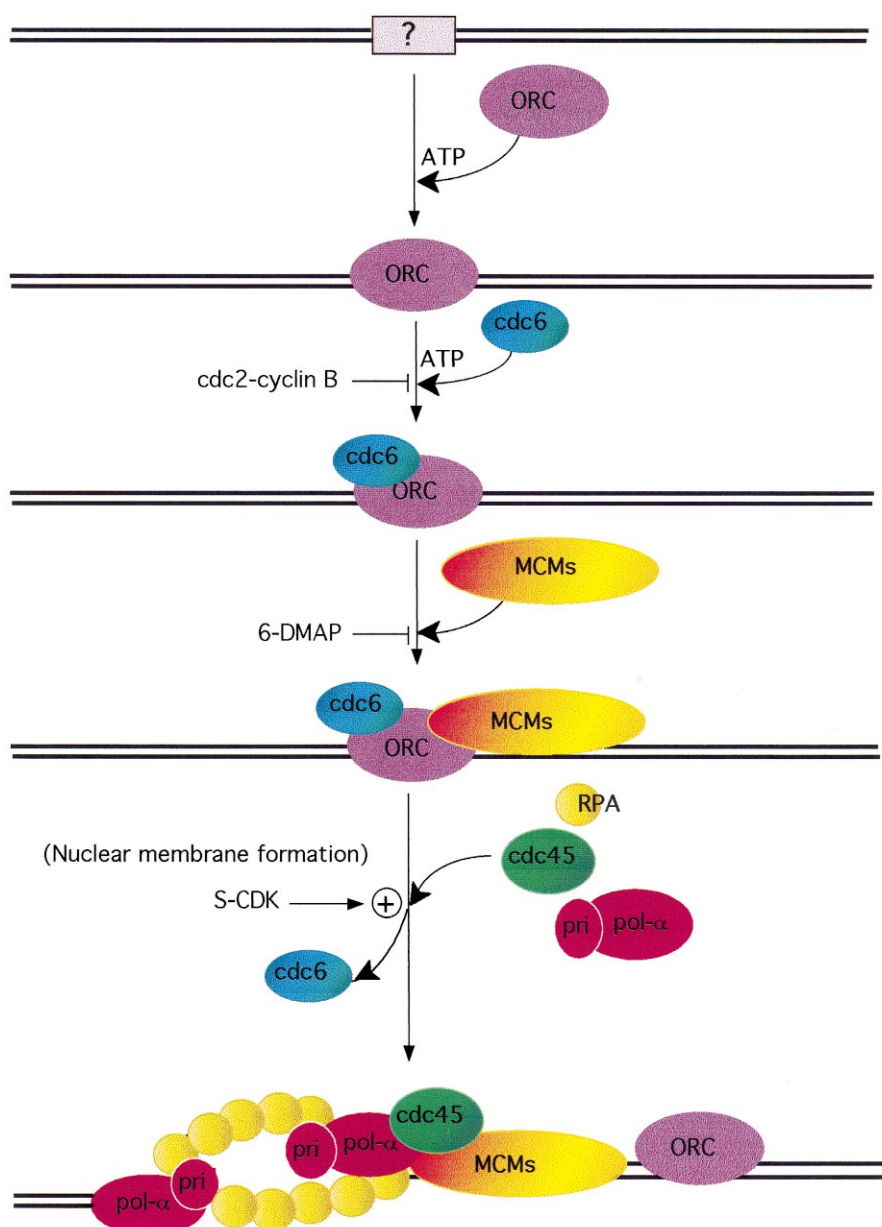


Fig. 2. Assembly of the pre-replication complex on a putative DNA replication origin. The assembly of the ORC with chromatin, in an ATP-dependent reaction, leads to the loading of the cdc6 protein. Once these proteins are present, MCM proteins are recruited to the origin. The cdk activation leads to the loading of the cdc45 protein which is responsible for the recruitment of the DNA polymerase- $\alpha$  on chromatin and the initiation of DNA replication. This highly simplified scheme is likely to expand to several discrete and limiting steps involving cyclin-cdk activities for both the activation and feedback control of origin firing.

the MCM subunits) makes the conservation of their functional mechanism in a complex eukaryote uncertain.

Vertebrate relatives of the yeast ORC complex, the *cdc6/cdc18* protein, the MCM protein family and the MCM-associated protein *cdc45/sna41* have been shown to be essential to initiate DNA synthesis (Fig. 2). Chromatin immunoprecipitation studies in yeast indicate that these proteins are targeted at DNA replication origins [24,25] in forming pre-replication complexes (pre-RCs). In metazoans, the chromosomal targets of these proteins are unknown, hence, studies on the assembly of pre-RCs have been limited to the association of these initiation factors with chromatin in the S phase.

**ORC proteins.** They represent the prototype of initiator proteins in *S. cerevisiae*, but how ORC proteins function in metazoans is a crucial point that remains to be elucidated. In yeast, the ORC complex specifically recognizes and binds the ARS consensus sequence in an ATP-dependent manner [26]. Recently, using a fine technique to map the sequence where the first deoxynucleotides are synthesized, the origin of DNA synthesis was shown to coincide with the binding site of the ORC complex in the A and partly B1 domains of ARS1 [7]. In the *Xenopus* in vitro systems, the assembly of pre-RCs can be uncoupled from the activation of DNA synthesis in several ways and discrete steps in the assembly of the initiation complex have been identified. The ORC complex is essential for the association of other initiation proteins, *cdc6* and the MCM proteins, indicating that the role of ORC may be to mark the site where pre-RCs assemble [27–30]. In metazoan, where no sequence consensus for an ori has been found, the ORC complex is still required to initiate DNA replication. It is then possible that the limited homology between yeast and mammalian ORC subunits is required for initiation functions other than the recognition of origin sequence motifs.

**Cdc6.** ORC is the target of the *cdc6* protein, whose major role appears to be the recruitment of MCM proteins onto chromatin [27,30,31]. It has been proposed that *cdc6* may accomplish this task in a complex with ORC1, 4 and 5 which are structurally related in a reaction that requires ATP [32,33]. In addition to ORC1, 4 and 5 [34], *cdc6* also shares sequence similarity with subunits of the RF-C complex, a PCNA loader [32,35]. *Cdc6* may then be a member of a prokaryotic and eukaryotic superfamily of nucleotide-dependent loading factors.

**MCMs.** MCM proteins comprise six related proteins which appear to have an important role in the regulation of origin firing. Little biochemical evidence for direct physical interactions between MCMs and the *cdc6*-ORC proteins has been reported [36]. The physical interactions between MCM proteins on chromatin are also not clear. Soluble MCM proteins associate in a stable complex, however, once they have gained access to chromatin, distinct MCM subunits may be re-distributed perhaps forming subcomplexes [37]. In support of the MCM subunit's distinct roles in DNA replication, a weak DNA helicase activity, which could be relevant to the unwinding of the DNA at replication origins, has been shown to be associated only with a subcomplex of MCM4, 6 and 7 proteins [38], whereas the association of the MCM2 protein with this subcomplex appears to have an inhibitory effect [39]. One of the MCM proteins, MCM4, shows distinct variations in its phosphorylation state that are cell cycle-dependent. One of the phosphorylation forms is present in the nucleus, specifically at the onset of DNA synthesis [40], suggesting a distinct

role of this form in the positive or negative control of initiation of DNA replication. Once MCM proteins have been loaded onto chromatin, ORC and *cdc6* may be dispensable for replication [30,31] and it is not yet clear if the loading site is the same for the three sets of proteins.

## 6. Regulation of the initiation complex by phosphorylation

Protein kinase activity is necessary for the initiation of DNA replication and several S phase cyclin-dependent kinases required in the assembly of the initiation complex have been detected both in yeast and complex organisms [41–43]. In vitro studies using *Xenopus* egg extracts or human cell extracts showed that cyclin E-cdk<sub>2</sub> and cyclin A-cdk<sub>2</sub> are regulating early events in DNA replication [30,44–47]. One step involved in the regulation by phosphorylation is the loading of the MCM complex onto chromatin and its activation [40,48,49]. Both S-CDKs and M-CDKs may be involved in this regulation, facilitating the coordination of S and M phase during the cell cycle [40–42]. Recent evidence in yeast and *Xenopus* indicates that S-CDK activity is also required to load on chromatin an MCM interacting protein, *cdc45*, whose function would be to allow the loading of DNA polymerase- $\alpha$  onto chromatin [50,51].

Initiation of DNA replication does not simply require S-CDKs but also requires a fine tuning of the intracellular level of the kinase activity. Several analyses performed both in yeast and *Xenopus* stress the importance of the crucial level of cyclin CDKs for initiation of DNA replication and the control of the firing of origins only once per cell cycle (reviewed in [41,42,52,53]).

The recent observations that the temporal program of origin activation requires the participation of B cyclins in yeast [53,54] suggest a similar regulatory function in complex eukaryotes since the timing of early and late replicons is strongly linked to gene expression. Cyclin cdks may thus be involved at each step of the building of the initiation complex and also control the firing program of the 1000 origins that will be used during the cell cycle. Consequently, a clear picture of initiation of DNA replication will not be obtained until their molecular mechanisms of action as well as the discrete steps of phosphorylation involved in the assembly and activation of an origin will be dissected.

## 7. Replication protein A (RP-A)

RP-A was one of the first eukaryotic proteins shown to be involved in the initiation of DNA replication [55–58]. RP-A is a trimeric DNA binding protein that recognizes single-stranded DNA. Biochemical and immunofluorescence studies show that RP-A may act at both the initiation and the elongation step of DNA replication [58,59]. Its role in the elongation process may be in keeping single-stranded DNA regions opened at the replication fork, thus stabilizing and helping to unwind the DNA during replication. MCM proteins may help in this process [60]. However, in Hela cells, the majority of RP-A appears associated to double-stranded DNA regions in chromatin [61] and the precise role of RP-A in the initiation of DNA replication remains puzzling. RP-A binds and decorates the initiation centers defined as the first sites where DNA synthesis is observed. Yet, no interactions of RP-A with ORC, *cdc6*, and the MCM proteins have been reported. Moreover,

both ORC and MCM proteins do not co-localize with DNA replication centers. *Xenopus* RP-A associates with chromatin before nuclear membrane formation and its binding is not dependent on ORC subunits [27] but depends on another protein, FFA-1, an homologue of the Werner's syndrome protein [62]. In metazoans, RP-A may become redistributed at origin sites at the transition from the pre-initiation complex to the DNA synthesis initiation complex.

## 8. Perspectives

Despite a dramatic acceleration of our knowledge in this field during the last years, the DNA replication unit and the concept of DNA replication origin remain elusive in complex eukaryotes. Regulation of DNA replication could be integrated more in the control of gene expression than in simple organisms. In yeast, initiator proteins have been identified thanks to the use of ARS's. Maybe in mammals, things will work the other way around.

## References

- [1] Jacob, F., Brenner, J., Cuzin, F. (1963) Cold Spring Harbor Symp. Quant. Biol. 28, pp. 329–348.
- [2] Stinchcomb, D.T., Struhl, K. and Davis, R.W. (1979) Nature 282, 39–43.
- [3] Marahrens, Y. and Stillman, B.A. (1992) Science 255, 817–823.
- [4] Bell, S.P. and Stillman, B. (1992) Nature 357, 128–134.
- [5] Dubey, D.D., Zhu, J., Carlson, D.L., Sharma, K. and Huberman, J.A. (1995) EMBO J. 13, 3638–3647.
- [6] Clyne, R.K. and Kelly, T.J. (1995) EMBO J. 14, 6348–6357.
- [7] Bielinsky, A.K. and Gerbi, S.A. (1998) Science 279, 95–98.
- [8] Heintz, N.H. and Hamlin, J.L. (1982) Proc. Natl. Acad. Sci. USA 79, 4083–4087.
- [9] Burhans, W.C., Vassilev, L.T., Caddle, M.S., Heintz, N.H. and DePamphilis, M.L. (1990) Cell 62, 955–965.
- [10] Pelizon, C., Diviacco, S., Falaschi, A. and Giacca, M. (1996) Mol. Cell Biol. 16, 5358–5364.
- [11] Dijkwel, P.A. and Hamlin, J.L. (1992) Mol. Cell Biol. 12, 3715–3722.
- [12] Kalejta, R.F., Li, X., Mesner, L.D., Dijkwel, P.A., Lin, H.B. and Hamlin, J.L. (1998) Mol. Cell. 2, 797–806.
- [13] Kobayashi, T., Rein, T. and DePamphilis, M.L. (1998) Mol. Cell Biol. 18, 3266–3277.
- [14] Giacca, M., Zentilin, L., Norio, P., Diviacco, S., Dimitrova, D., Contreas, G., Biamonti, G., Perini, G., Weighardt, F., Riva, S. and Falaschi, A. (1994) Proc. Natl. Acad. Sci. USA 91, 7119–7123.
- [15] Kumar, S., Giacca, M., Nori, X., Biamonti, G., Riva, S. and Falaschi, A. (1996) Nucleic Acids Res. 24, 3289–3294.
- [16] Abdurashidova, G., Riva, S., Biamonti, G., Giacca, M. and Falaschi, A. (1998) EMBO J. 17, 2961–2969.
- [17] Kitsberg, D., Selig, S., Keshet, I. and Cedar, H. (1993) Nature 366, 588–590.
- [18] Aladjem, M.I., Groudine, M., Brody, L.L., Dieken, E.S., Fournier, R.E., Wahl, G.M. and Epner, E.M. (1995) Science 270, 815–819.
- [19] Aladjem, M.I., Rodewald, L.W., Kolman, J.L. and Wahl, G.M. (1998) Science 281, 1005–1009.
- [20] Newport, J. and Kirschner, M. (1982) Cell 30, 687–696.
- [21] Hyrien, O., Maric, C. and Mech, M. (1995) Science 270, 994–997.
- [22] Sasaki, T., Sawado, T., Yamaguchi, M. and Shinomiya, T. (1999) Mol. Cell Biol. 19, 547–555.
- [23] Majumder, S. and DePamphilis, M.L. (1995) Bioessays 17, 879–889.
- [24] Aparicio, O.M., Weinstein, D.M. and Bell, S.P. (1997) Cell 91, 59–69.
- [25] Tanaka, T., Knapp, D. and Nasmyth, K. (1997) Cell 90, 649–660.
- [26] Klemm, R.D., Austin, R.J. and Bell, S.P. (1997) Cell 88, 493–502.
- [27] Coleman, T.R., Carpenter, P.B. and Dunphy, W.G. (1996) Cell 87, 53–63.
- [28] Romanowski, P., Madine, M.A., Rowles, A., Blow, J.J. and Laskey, R.A. (1996) Curr. Biol. 6, 1416–1425.
- [29] Rowles, A., Chong, J.P., Brown, L., Howell, M., Evan, G.I. and Blow, J.J. (1996) Cell 87, 287–296.
- [30] Hua, X.H. and Newport, J. (1998) J. Cell Biol. 140, 271–281.
- [31] Donovan, S., Harwood, J., Drury, L.S. and Diffley, J.F. (1997) Proc. Natl. Acad. Sci. USA 94, 5611–5616.
- [32] Perkins, G. and Diffley, J.F. (1998) Mol. Cell. 2, 23–32.
- [33] Weinreich, M., Liang, C. and Stillman, B. (1999) Proc. Natl. Acad. Sci. USA 96, 441–446.
- [34] Quintana, D.G., Hou, Z.H., Thome, K.C., Hendricks, M., Saha, P. and Dutta, A. (1997) J. Biol. Chem. 272, 28247–28251.
- [35] Tugal, T., Zou-Yang, X.H., Gavin, K., Pappin, D., Canas, B., Kobayashi, R., Hunt, T. and Stillman, B. (1998) J. Biol. Chem. 273, 32421–32429.
- [36] Grallert, B. and Nurse, P. (1996) Genes Dev. 10, 2644–2654.
- [37] Coue, M., Amariglio, F., Maiorano, D., Bocquet, S. and Mechali, M. (1998) Exp. Cell Res. 245, 282–289.
- [38] Ishimi, Y. (1997) J. Biol. Chem. 272, 24508–24513.
- [39] Ishimi, Y., Komamura, Y., You, Z. and Kimura, H. (1998) J. Biol. Chem. 273, 8369–8375.
- [40] Coue, M., Kearsy, S.E. and Mechali, M. (1996) EMBO J. 15, 1085–1097.
- [41] Dutta, A. and Bell, S.P. (1997) Annu. Rev. Cell Dev. Biol. 13, 293–332.
- [42] Wuari, J. and Nurse, P. (1996) Cell 85, 785–787.
- [43] Hengstschlager, M., Braun, K., Soucek, T., Miloloza, A. and Hengstschlager-Ottner, E. (1999) Mutat. Res. 436, 1–9.
- [44] Jackson, P.K., Chevalier, S., Philippe, M. and Kirschner, M.W. (1995) J. Cell Biol. 130, 755–769.
- [45] Strausfeld, U.P., Howell, M., Descombes, P., Chevalier, S., Rempel, R.E., Adamczewski, J., Maller, J.L., Hunt, T. and Blow, J.J. (1996) J. Cell Sci. 109, 1555–1563.
- [46] Chevalier, S., Couturier, A., Chartrain, I., Le Guelle, R., Beckhelling, C., Le Guellec, K., Philippe, M. and Ford, C.C. (1996) J. Cell Sci. 109, 1173–1184.
- [47] Krude, T., Jackman, M., Pines, J. and Laskey, R.A. (1997) Cell 88, 109–119.
- [48] Hendrickson, M., Madine, M., Dalton, S. and Gautier, J. (1996) Proc. Natl. Acad. Sci. USA 93, 12223–12228.
- [49] Todorov, I.T., Attaran, A. and Kearsy, S.E. (1995) J. Cell Biol. 129, 1433–1445.
- [50] Zou, L. and Stillman, B. (1998) Science 280, 593–596.
- [51] Mimura, S. and Takisawa, H. (1998) EMBO J. 17, 5699–5707.
- [52] Coverley, D., Wilkinson, H.R., Madine, M.A., Mills, A.D. and Laskey, R.A. (1998) Exp. Cell Res. 238, 63–69.
- [53] Donaldson, A.D. and Blow, J.J. (1999) Curr. Opin. Genet. Dev. 9, 62–68.
- [54] Schwob, E. and Nasmyth, K. (1993) Genes Dev. 7, 1160–1175.
- [55] Wobbe, C.R., Weissbach, L., Borowiec, J.A., Dean, F.B., Murakami, Y., Bullock, P. and Hurwitz, J. (1987) Proc. Natl. Acad. Sci. USA 84, 1834–1838.
- [56] Fairman, M.P. and Stillman, B. (1988) EMBO J. 7, 1211–1218.
- [57] Wold, M.S. and Kelly, T. (1988) Proc. Natl. Acad. Sci. USA 85, 2523–2527.
- [58] Adachi, Y. and Laemmli, U.K. (1994) EMBO J. 13, 4153–4164.
- [59] Yan, H. and Newport, J. (1995) J. Cell Biol. 129, 1–15.
- [60] Tanaka, T. and Nasmyth, K. (1998) EMBO J. 17, 5182–5191.
- [61] Treuner, K., Eckerich, C. and Knippers, R. (1998) J. Biol. Chem. 273, 31744–31750.
- [62] Yan, H., Chen, C.Y., Kobayashi, R. and Newport, J. (1998) Nat. Genet. 19, 375–378.