uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Julia Romero

AUTEUR DE LA THÈSE / AUTHOR OF THESIS

Ph.D. (Biochemistry)

GRADE / DEGREE

Department of Biochemistry, Microbiology and Immunology

FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Identification and Detailed Characterization of a Novel Origin of DNA Replication in the Transcriptional Promoter of the Human *DBF4* Gene

TITRE DE LA THÈSE / TITLE OF THESIS

Hoyun Lee

DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

Jonathan Lee

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

Marjorie Brand

Martin Holcik

Bernard Duncker

Ian Lorimer

Gary W. Slater

Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

IDENTIFICATION AND DETAILED CHARACTERIZATION OF A NOVEL ORIGIN OF DNA REPLICATION IN THE TRANSCRIPTIONAL PROMOTER OF THE HUMAN *DBF4* GENE

Julia Romero

Thesis submitted to the

Faculty of Graduate and Postdoctoral Studies

in partial fulfillment of the requirements

for the PhD degree in Biochemistry

Department of Biochemistry, Microbiology and Immunology

Faculty of Medicine

University of Ottawa

Your file Votre référence
ISBN: 978-0-494-52339-1
Our file Notre référence
ISBN: 978-0-494-52339-1

ABSTRACT

Origins of DNA replication (oris) are the genomic sites where DNA synthesis starts to duplicate a cell's genome. Although prokaryotic and yeast oris have been characterized in detail, the nature of mammalian oris is still ill-defined. In fact, of ~30,000 oris estimated to exist in mammalian genomes, only the one localized at the human *lamin B2* locus has been mapped with nucleotide resolution.

This thesis describes the identification of a novel human ori located around the transcriptional promoter of the *DBF4* gene. Using a modified version of the replication initiation point (RIP) mapping technique, precise leading strand initiation sites have been determined at this new ori. Additionally, chromatin immunoprecipitation (ChIP) assays have been employed to study the interaction of replication and transcription proteins within the ori region.

The data presented here demonstrate that, at the DBF4 ori, DNA replication can initiate from multiple potential start-sites which are distributed within two replication initiation zones. These zones are separated by ~400 bp and are localized near binding sites for the origin recognition complex (ORC) and the Sp1 transcription factor. Strikingly, initiation of DNA synthesis from the two zones occurs sequentially and not simultaneously: replication starts first from initiation zone I and proceeds in the direction of *DBF4* transcription, followed by activation of zone II and DNA synthesis from the complementary strand. This replication pattern, termed asymmetric bidirectional replication (ABR), is reminiscent of initiation at oriC in *Escherichia coli* but greatly differs from what is observed at the lamin B2 ori, where a single ORC-binding site and unique replication start-points have been reported.

Overall, the data suggest, for the first time, that eukaryotic replication initiation can take place through either the traditional origin of bidirectional replication (OBR) model, such as is the case in budding yeast and at the human lamin B2 ori, or through the asymmetric ABR model described here for the DBF4 ori.

ACKNOWLEDGMENTS

First and foremost I would like to thank my supervisor, Dr. Hoyun Lee, for giving me an amazing opportunity by accepting me in as a PhD student. Certainly, none of this work could have been possible without his scientific instinct, guidance, and help. I am also very grateful to the members of my Thesis Advisory Committee, Dr. Jonathan Lee and Dr. Bruce McKay, who endured long and less than perfect videoconferencing sessions to provide helpful input and orientation.

I would like to give special thanks to all members of the Research lab at the Sudbury Regional Hospital, both past and present. Whenever things got hard there was nothing like having a good cup of coffee in the lunch room to regroup and start again. Lisa, Stacey, Paul and Kevin, I will never forget all the good times and the many laughs we shared, both inside and outside the lab.

I would like to give thanks to the Faculty of Graduate and Postdoctoral Studies of the University of Ottawa and to the Northern Cancer Research Foundation for their continuous financial support over these years.

Finally, for helping me keep my sanity and for all their patience, support and attention I am immensely grateful to my husband, Andres, and to my parents, my brother and friends back home in Venezuela.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Abf1	ARS binding factor 1
ABR	Asymmetric bidirectional replication
ACE3	Amplification control element 3
ACS	ARS consensus sequence
AER	Amplification enhancing regions
ARS	Autonomously replication sequence
ATP	Adenosine triphosphate
bp	Basepairs
Cdc6	Cell division cycle protein 6
Cdc7	Cell division cycle protein 7
CDK	Cyclin-dependent kinase
ChIP	Chromatin immunoprecipitation
Dbf4	Dumbbell former protein 4
DHFR	Dihydrofolate reductase
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotides triphosphate
DUE	DNA unwinding element
EDTA	Ethylene diamine tetraacetic acid
EGTA	Ethylene glycol tetraacetic acid
GINS	Go, Ichi, Nii and San protein complex
IHF	Integration host factor
kb	Kilobases
LMNB2	Lamin B2
LM-PCR	Ligation-mediated PCR
MAR	Matrix attachment region
Mbp	Megabasepairs
MBT	Midblastula transition
MCB	*Mlu*I-cell cycle box
MCFP	Mitochondria carrier family protein

MCM	Minichromosome maintenance
MW	Molecular weight
NRF-1	Nuclear respiratory factor 1
nt	Nucleotides
OBR	Origin of bidirectional replication
ORB	Origin recognition boxes
ORC	Origin recognition complex
ORF	Open reading frame
ori	Origin of DNA replication
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PMSF	Phenylmethylsulphonyl fluoride
Post-RC	Post-replication complex
Pre-RC	Pre-replication complex
Q-PCR	Quantitative PCR
rDNA	Ribosomal DNA gene
RIP	Replication initiation point
RNA	Ribonucleic acid
RPA	Replication protein A
SDS	Sodium dodecyl sulfate
SSB	Single-stranded DNA binding protein
SV40	Simian virus 40
TFIIB	Transcription factor IIB
TP	Transition point

LIST OF FIGURES

LIST OF TABLES

Table **Page**

Chapter 1

Introduction

1.1 DNA replication

The timely and faithful replication of a cell's genome prior to cell division is, without a doubt, one of the most astounding cellular processes. On one hand, cells must ensure that **all** their genetic material, which can reach up to 4,000 megabasepairs (Mbp) in some mammalian cells, gets copied in a relatively short time. On the other, the replication process must be extremely accurate to prevent the incorporation of potentially deleterious mutations. This delicate balance between high speed and accuracy is achieved through a multitude of highly sophisticated mechanisms that control and regulate the replication of cellular DNA.

In general, the process of DNA replication can be divided into three stages: (1) initiation, (2) elongation, and (3) termination. During the initiation phase, the chromosomal sites where DNA synthesis starts are recognized and bound by specific protein complexes. This process leads to melting of the parental DNA duplex, recruitment of the DNA synthesis machinery and, eventually, the ongoing polymerization of nascent DNA strands during the elongation phase. Finally, DNA synthesis ends at the termination stage, mainly due to the convergence of opposing replication forks.

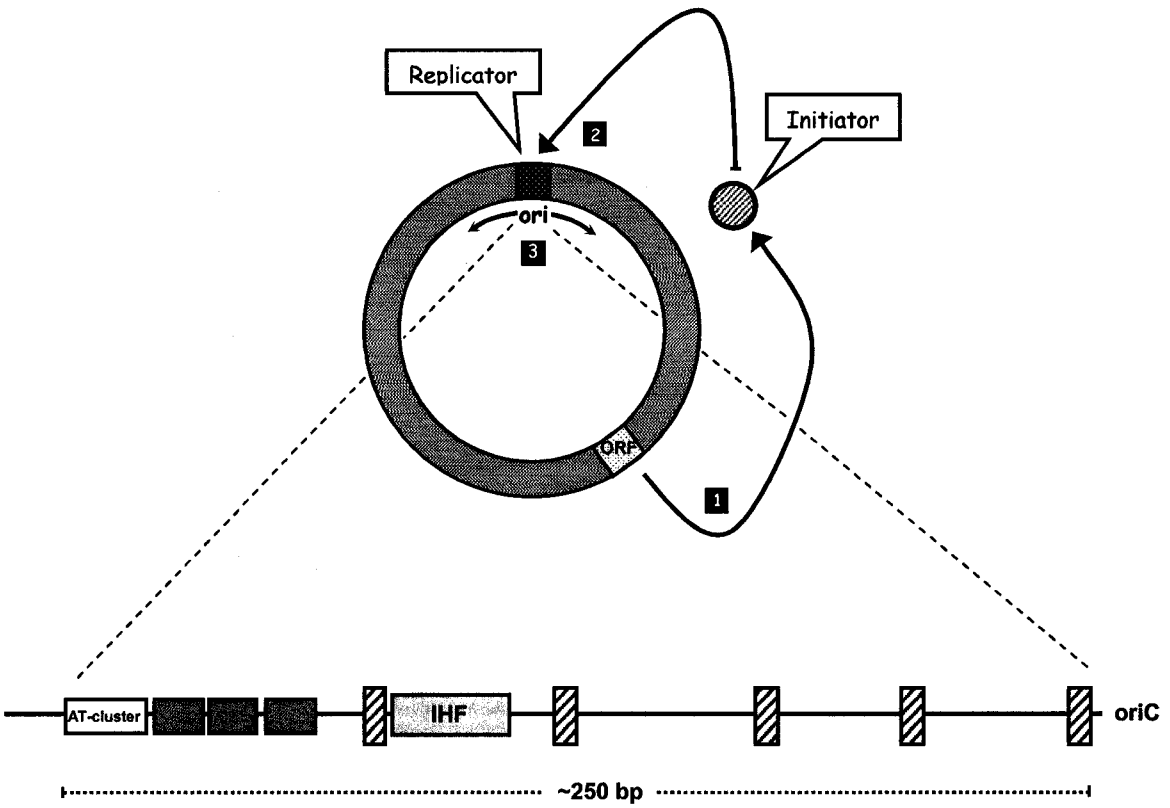
In all organisms, the initiation phase is, by far, the most intricate and heavily regulated stage of the replication process. The reasons for this are threefold: (1) DNA synthesis should only be initiated during a defined point in the cell cycle (S phase), (2) the entire genome must be replicated during S phase, which, particularly in eukaryotic organisms, requires the initiation of DNA synthesis from multiple chromosomal sites known as origins of DNA replication (oris), and (3) to prevent re-replication of genomic sequences, each ori must initiate DNA synthesis only once during a given S phase.

Given the tremendous complexity of the initiation process and the regulatory networks that affect it, it is not surprising that relatively little is known about it, particularly in higher organisms. The most relevant aspects of the initiation of DNA replication in different organisms will be discussed in the following sections.

1.2 Bacterial DNA replication: The replicon model

Due to their relative simplicity, bacteria and bacteriophages have provided us with the best characterized and most detailed models of DNA replication. In fact, a large part of our current knowledge regarding the enzymes involved in eukaryotic DNA metabolism has been fueled by the analysis of their bacterial counterparts.

In 1963, based on the transcriptional regulation models available at the time, Jacob, Brenner and Cuzin proposed a model to explain how DNA replication could be initiated and controlled in bacterial cells (Jacob et al., 1963). This model, known as the **replicon model**, postulated the existence of two essential components required for the initiation of DNA synthesis: (i) a trans-acting factor, termed **initiator**, which interacts with specific genomic sequences to activate replication of an entire genomic area or replicon, and (ii) a *cis*-acting element, the **replicator**, which corresponds to the DNA sequence recognized by the initiator protein (Figure 1). Binding of the initiator to the replicator promotes DNA synthesis initiation from an ori, which may or may not precisely coincide with the replicator element. Although mostly theoretical in its conception, the replicon model withstood subsequent experimental testing and was shown to be correct, particularly in prokaryotic cells.



1.2.1 Bacterial origins of DNA replication

Most bacteria contain a single, circular chromosome which is replicated from a unique ori known as oriC (Kornberg and Baker, 1992). In *Escherichia coli*, oriC is a ~250 bp region that contains all the essential sequences required for the initiation of DNA synthesis (Oka et al., 1980). These sequences include five copies of a 9 bp consensus element known as DnaA box, as well as three AT-rich 13-mer repeats and an AT cluster (Figure 1) (Messer, 2002). The DnaA boxes correspond to the binding sites for the initiator protein DnaA. DnaA-mediated opening of the easily unwound AT-rich area is a critical step in the initiation of DNA synthesis (Bramhill and Kornberg, 1988). The presence of multiple DnaA boxes and a conserved AT-rich region is a common theme in bacterial oris (Messer, 2002).

1.2.2 Bacterial initiator proteins and DNA replication machinery

DnaA is a 52 kDa polypeptide highly conserved in all eubacteria. The protein is a member of the AAA^+ family of ATPases and serves both as a replication initiator and a transcription factor. Four distinct functional domains have been identified in DnaA (Messer, 2002). The N-terminal domain I is involved in DnaA oligomerization and loading of the DnaB bacterial helicase. Domain II is highly variable among different bacterial species, has the structural features of a flexible linker, and includes regions that can be dispensable for DnaA function. Domain III contains ATP-binding and hydrolysis motifs, as well as additional sites for interaction with DnaB and DnaA oligomerization. Finally, the C-terminal domain IV includes a helix-turn-helix motif which mediates specific binding to DnaA boxes (Messer, 2002).

The first step in bacterial replication initiation requires high-affinity binding of DnaA monomers to each of the DnaA boxes present in the oriC. Following this, additional DnaA molecules bind to the region forming a central protein core around which the ori is wrapped. This structure appears to be stabilized through an intricate network of protein-protein and protein-DNA interactions and is believed to result in distortion of the duplex DNA around the ori's AT-rich area, forming a so-called "open complex" (Messer, 2002). In the next step, the DnaB and DnaC proteins are recruited to the ori. Hexameric DnaB constitutes the replicative DNA helicase, whereas DnaC serves both as a loading factor and a regulator of the DnaB polypeptide (Davey et al., 2002). Finally, additional proteins are loaded onto the ori to form functional replication forks, including the single-stranded DNA binding protein (SSB), the DnaG primase, and the bacterial DNA polymerase III holoenzyme.

1.3 Archaeal DNA replication

Similar to other prokaryotes, archaeal genomes usually correspond to single, circular chromosomes. In fact, based solely on their structural and metabolic features, archaeal cells seem to be closely related to eubacteria. However, when it comes to other cellular processes such as transcription, translation and DNA replication, they appear to resemble more strongly their eukaryotic counterparts (Olsen and Woese, 1997). Consequently, archaeal cells have slowly become good model systems to study and understand eukaryotic DNA replication.

1.3.1 Archaeal origins of DNA replication

The first putative archaeal oris were identified *in silico* using skew analysis (to detect G/C strand-specific biases) and computational searches for ori-specific sequence properties such as AT-content and inverted repeat (IR) elements (Salzberg et al., 1998; Lopez et al., 1999). Following this, clear experimental evidence has accumulated, particularly from studies using the archaeon *Pyrococcus abyssi*, to demonstrate that at least some of the initially identified regions function as bona fide oris *in vivo* (Myllykallio et al., 2000; Matsunaga et al., 2001, Matsunaga et al., 2003).

In the case of *P. abyssi*, the entire genome appears to be replicated from a single ori located within an 800-bp non-coding region immediately upstream of the *cdc6* gene (Matsunaga et al., 2001). Interestingly, this gene codes for a polypeptide similar to both the ORC1 (origin recognition complex subunit 1) and Cdc6 eukaryotic replication factors, two proteins directly involved in ori recognition and helicase loading in eukaryotic cells (section 1.4.6). Replication initiation point (RIP) mapping of the *P. abyssi* ori has subsequently allowed the identification of unique, well-defined leading strand initiation sites on both parental strands, located within a 7 nt segment near the left end of the ori region (Matsunaga et al., 2003). A single ori has also been found in the large chromosome of the halophilic archaeon *Halobacterium* strain NRC-1. In this case, the ori maps near the *orc7* gene, one of three functional chromosomal Orc1/Cdc6 homologs present in this species (Berquist and DasSarma, 2003).

A slightly different replication mechanism was observed in other species such as *Sulfolobus solfataricus* and *Sulfolobus acidocaldarius*, in which at least three chromosomal oris have been shown to exist (Lundgren et al., 2004). However, similar to *Halobacterium*

and *P. abyssi*, two of the identified oris are located in the upstream region of genes coding for Orc1/Cdc6 homologs (three such genes are present in both *Sulfolobus* species, but only two of them are physically linked to oris) (Lundgren et al., 2004; Robinson et al., 2004).

Based on the combined results of these studies, a somewhat unified, although still incomplete picture of archaeal oris has emerged. First, as is the case in many bacterial genomes, most archaeal oris appear to be located in close proximity to genes coding for proteins involved in the initiation of DNA replication. Second, sequence comparison suggests that archaeal oris might share conserved inverted repeat sequence elements termed origin recognition boxes (ORB), which are recognized by Orc1/Cdc6 homologs (Robinson et al., 2004). Finally, archaeal oris display other features common to both bacterial and eukaryotic oris such as being located close to AT-rich DNA regions.

1.3.2 Initiation of DNA replication in archaeal cells

With the exception of the methanogens *Methanococcus jannaschii, Methanococcus maripaludis* and *Methanopyrus kandleri*, sequences encoding Orc1/Cdc6 homologs have been identified in all archaeal genomes analyzed to date (Barry and Bell, 2006). Since known archaeal oris do not include motifs similar to DnaA boxes and sequences related to bacterial *dnaA* genes have not been identified in any archaeal genome, the mechanism of archaeal DNA replication initiation appears to be more closely related to that of eukaryotic cells.

Just like bacterial DnaA and eukaryotic ORC1 and Cdc6 proteins, archaeal Orc1/Cdc6 homologs belong to the family of AAA^+ATPases and include a C-terminal domain required for binding to double-stranded DNA (Kelman and Kelman, 2003). In fact,

archaeal Orc1/Cdc6 proteins have been shown to bind oris *in vivo*, supporting their role in the initiation of DNA replication (Matsunaga et al., 2001; Robinson et al., 2004). Thus, according to the currently accepted model, the archaeal Orc1/Cdc6 proteins serve the function of eukaryotic ORC and Cdc6/Cdt1, executing origin recognition and loading of the DNA helicase onto archaeal oris. It still remains unclear however, whether individual Orc1/Cdc6 homologs are able to perform both of these functions, as would be expected in those species that contain only one Orc1/Cdc6 gene. Alternatively, some Orc1/Cdc6 proteins could correspond to different "subunits" of the eukaryotic ORC, while others could function as the helicase loader (Robinson et al., 2004; Majerník et al., 2004).

In addition to Orc1/Cdc6 homologues, archaeal genomes also include at least one gene similar to the eukaryotic *m*ini*c*hromosome *m*aintenance complex (MCM) family of proteins. These proteins, as will be discussed later, form a heterohexameric complex believed to function as helicase during eukaryotic DNA replication. Although six different subunits take part in the assembly of the MCM complex in eukaryotic cells, most archaeal genomes contain a single *mcm* gene (Barry and Bell, 2006). Despite this difference, archaeal MCM proteins have been shown to oligomerize *in vitro*, forming homohexameric and double-hexameric ring-shaped complexes that display 3' to 5' DNA helicase activity (Chong et al., 2000). It is therefore hypothesized that, just like in eukaryotic cells, archaeal MCM complexes might function as helicases to unwind duplex DNA at replication forks during chain elongation.

Finally, it is worth noting that, despite the similarities between DNA replication initiation in archaeal and eukaryotic cells, several proteins required for the initiation process in eukaryotes do not appear to be present in archaeans. In addition, since some bacterial-like

replication proteins have been identified in archaeal genomes, the archaeal replication process likely combines features of both eukaryotic and bacterial cells.

1.4 Eukaryotic DNA replication

1.4.1 Budding yeast origins of DNA replication

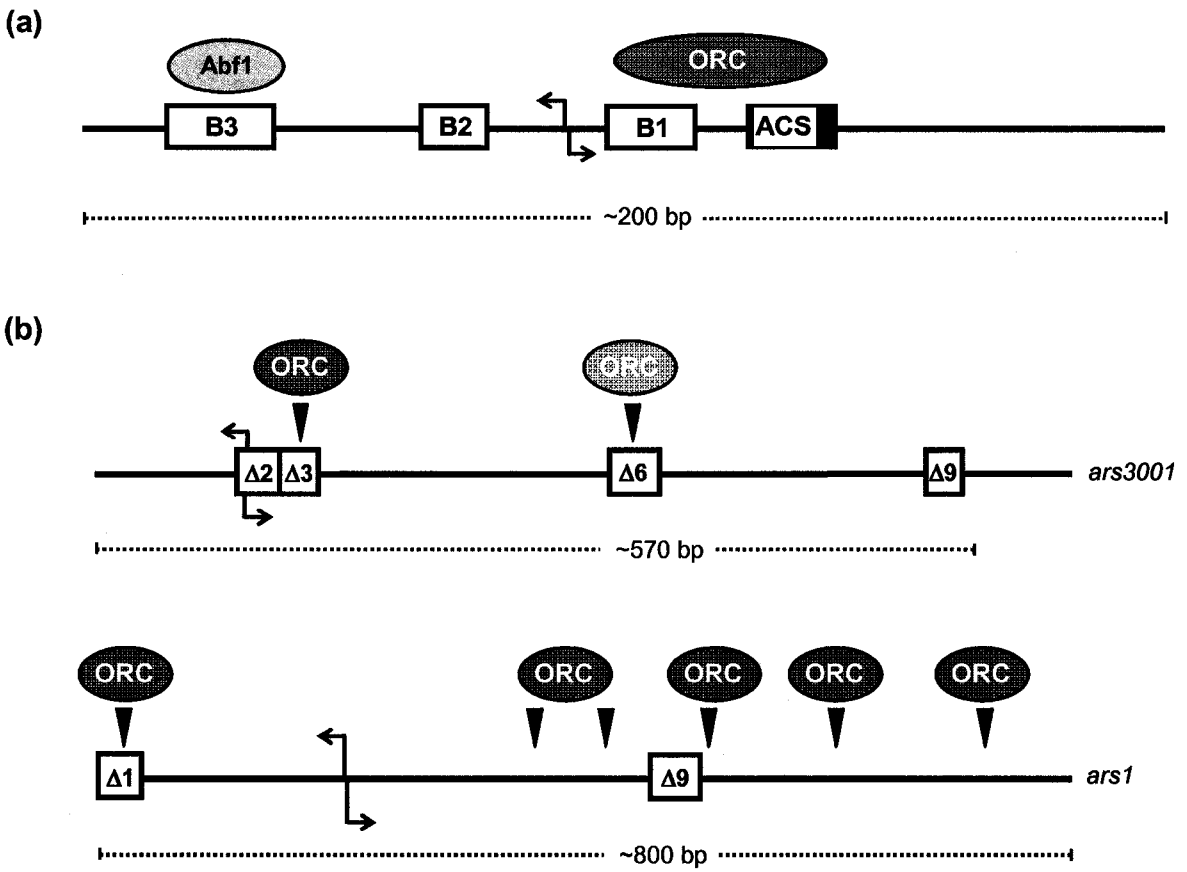
Saccharomyces cerevisiae represents the only eukaryotic organism in which the original replicon model proposed by Jacob and colleagues applies with only minor modifications. Obviously, since the budding yeast genome is composed of 16 individual chromosomes, replication from a single ori as it occurs in bacteria is not possible. Instead, multiple oris are required in order to copy all the genetic information of a yeast cell. In addition, when the size of the yeast genome (~13 Mbp versus ~4.5 Mbp of most bacteria), and the lower speed of eukaryotic DNA polymerases (~50 nt/second compared to ~500–1000 nt/second in prokaryotes) are taken into account, it becomes evident that more than one ori must be activated on each yeast chromosome in order to ensure that DNA replication can be completed within the duration of S phase (~40 min). Accordingly, it is estimated that between 300 and 400 oris exist throughout the *S. cerevisiae* genome (Wyrick et al., 2001; Raghuraman et al., 2001, Feng et al., 2006).

Still, in spite of their higher number, all budding yeast oris correspond to specific, well-defined genomic regions, just as is the case for bacterial replicators. These sequences were initially identified based on their ability to allow the extrachromosomal propagation of plasmids, and thus were originally named *a*utonomously *r*eplicating *s*equences or ARSs (Stinchcomb et al., 1979). All yeast ARSs identified so far are AT-rich sequences of approximately 100 to 200 nucleotides in length, which share a conserved 11 bp sequence

motif essential for ori function, known as the *A*RS *c*onsensus *s*equence or ACS (Newlon and Theis, 1993; reviewed by Aladjem et al., 2006).

Yeast oris are composed of several different functional domains, including one A-element (containing the ACS) and at least two less-conserved B-elements (B1 to B4), located to the 3' side of the T-rich strand of the ACS (Figure 2a) (Marahrens and Stillman, 1992; Theis and Newlon, 1994). Although individual B-elements are not essential for ori activity, together they are required for efficient replicator function (Marahrens and Stillman, 1992). The hexameric ORC (ORC1-6), which serves as the eukaryotic initiator, has been shown to directly interact *in vitro* and *in vivo* with specific nucleotide positions in both the ACS and the B1-element (Bell and Stillman, 1992; Rao and Stillman, 1995; Wyrick et al., 2001). The B2 region usually includes a DNA unwinding element (DUE) and might also serve as the interaction site for additional replication proteins (Aladjem et al., 2006). Finally, the B3 element has been shown to bind transcription factors (particularly the yeast protein Abf1) whereas the B4 region may be easily unwound as well as recognized by specific replication factors, similarly to the B2 domain (Aladjem et al., 2006).

Precise RIP mapping of leading strand start-sites has been reported for two budding yeast oris: ARS1 and ARS603. ARS1, arguably the most studied ARS in *S. cerevisiae*, is a ~ 200 bp sequence located on chromosome IV. At this ori, leading strand synthesis has been shown to initiate from unique nucleotide positions, displaced by only 2 bp on both parental strands (Figure 2a) (Bielinsky and Gerbi, 1999). The identified RIPs or transition points (TPs) between continuous and discontinuous DNA synthesis are located immediately adjacent to the ORC-binding region, between elements B1 and B2 (Figure 2a).



Similarly, a unique TP is also thought to exist in ARS603. In this case, initiation sites on both strands are located 9 bp apart from each other, coinciding within a 9/11 match to the ACS required for ORC binding (Bolon and Bielinsky, 2006).

Even though plasmid-based studies were instrumental in the discovery of yeast oris, not all ARSs identified through these assays act as oris in their native chromosomal locations. In addition, even though more than 10,000 ACS matches exist throughout the yeast genome, only some of them are bound by ORC and function as oris during S phase. Both of these observations suggest that, besides nucleotide sequence, ori activity in yeast cells might be modulated by other factors including chromatin structure and chromosomal context. This agrees with the preferential localization of yeast chromosomal oris in intergenic spacers or non-transcribed genomic segments (Wyrick et al., 2001). Furthermore, specific nucleosome arrangements have been shown to have a significant impact on ARS function (Simpson, 1990; Lipford and Bell, 2001).

1.4.2 Fission yeast origins of DNA replication

It is estimated that approximately 300 oris exist throughout the three chromosomes that constitute the ~14 Mbp genome of *Schizosaccharomyces pombe* (Feng et al., 2006). Most of these oris have been mapped within intergenic regions and gene promoters (Gómez and Antequera, 1999; Dai et al., 2005).

Surprisingly, fission yeast oris have turned out to be very different from those of the budding yeast. First, they are much larger, with the smallest one spanning at least 500 bp (Kim and Huberman, 1998) and the largest ones reaching up to 1.8 kb (Johnston and Barker, 1987). Second, although DNA replication has been shown to initiate from specific genomic

regions, no conserved sequence motif similar to the *S. cerevisiae* ACS has been identified amongst the currently known *S. pombe* oris (Robinson and Bell, 2005). Finally, and in contrast to budding yeast, DNA replication initiation in *S. pombe* has been hypothesized to be a stochastic process, with only a random subset of many potential oris being activated in a given S phase (Dai et al., 2005; Patel et al., 2006).

Importantly, *S. pombe* oris are extremely AT-rich and contain several regions critical for activity, which usually correspond to clusters of asymmetrically distributed A's and T's (Fig. 2b) (Zhu et al., 1994; Clyne and Kelly, 1995; Kim and Huberman, 1998; Okuno et al., 1999). Thus, even though no ori consensus sequence has been identified, origin recognition in fission yeast is believed to occur through cooperative binding of the ORC to all or some of the asymmetric AT-stretches present in the ori region (Lee et al., 2001; Kong and DePamphilis, 2002; Takahashi et al., 2003). Supporting this hypothesis, *in vitro* binding of fission yeast ORC to ori DNA has been shown to be mediated by a single ORC subunit, namely ORC4, which contains nine N-terminal AT-hooks (Chuang and Kelly, 1999; Lee et al., 2001).

Considering the large size of fission yeast replicators, it became important to determine whether leading strand synthesis initiated from single, precisely defined sites as reported in budding yeast, or if, instead, multiple delocalized initiation events occurred at these oris. RIP mapping of two different *S. pombe* oris has proved the former to be the case. For example, in the *ars1* ori, single transition points were detected on both parental strands, only 3 bp apart from each other (Gómez and Antequera, 1999). Similarly, in *ars3001*, unique initiation points are separated by 10 bp and are located immediately upstream of one of the two ORC-binding sites identified in this ori (Fig. 2b) (Kong and DePamphilis, 2002).

1.4.3 DNA replication in *Xenopus laevis*

Most of the initial studies aimed at elucidating the DNA replication mechanism of the frog *Xenopus laevis* were focused on understanding how plasmid DNA was copied in intact eggs or egg extracts. Data from these experiments suggested that any segment of the plasmid DNA molecule could act as an ori without particular sequence-specificity (Méchali and Kearsey, 1984; Hyrien and Méchali, 1992; Mahbubani et al., 1992).

A similar situation was also observed for replication at the chromosomal 40S rRNA precursor gene cluster in *X. laevis* embryos collected at a stage prior to the midblastula transition (MBT) point (Hyrien and Méchali, 1993). In this case, however, although any DNA segment of the rDNA cluster could serve as an ori, initiation events were usually regularly spaced at intervals of ~9- to 12 kb in individual chromosomes. This suggested that, while random in terms of nucleotide sequence, DNA replication in *X. laevis* early embryos starts from frequently spaced genomic sites which may be selected based on sequence-independent chromatin features (Hyrien and Méchali, 1993). Periodic spacing of initiation sites was believed to be crucial to ensure complete and rapid DNA replication during the short S phase that characterizes the early *X. laevis* developmental stages, a feat that would otherwise be compromised by having a completely arbitrary replication initiation system (an issue commonly referred to as the "random completion problem"). Although the establishment of regularly-spaced initiation sites is still considered an important contributing factor in addressing the random completion problem (Blow, 2001), additional solutions have been proposed. For example, it has been suggested that the frequency of replication initiation increases in late S phase in areas of unreplicated DNA (Herrick, 2000; Edwards et al., 2002) or that replication can initiate on individual chromosomes before the assembly of

complete nuclei, increasing the time during which DNA can be effectively replicated (Lemaitre et al., 1998).

Even though *Xenopus* eggs and early embryos constitute useful systems for the study of DNA replication, they do not necessarily reflect the situation encountered in somatic cells. One crucial difference is that the cell cycles of somatic cells include G1 and G2 phases, which are both absent during the early stages of *Xenopus* embryonic development. In addition, somatic cells display longer S phases and are proficient in gene transcription, which does not take place in *Xenopus* embryos until they reach the MBT.

Accordingly, when replication of the *Xenopus* rRNA cluster was analyzed at later developmental stages (i.e., after the MBT), the observed initiation pattern was dramatically different from the "random" system that seems to operate during early embryonic divisions (Hyrien et al., 1995; Bozzoni et al., 1981). Thus, after transcription activation occurs at the MBT, replication initiation was shown to be repressed within the *Xenopus* rRNA transcription units and to be mostly confined to the intergenic spacers, pointing to a close relationship between DNA replication and transcription (Hyrien et al., 1995). Moreover, although transcriptionally silent plasmids injected in *Xenopus* eggs replicate without sequence specificity, activation of transcription from inducible promoters results in site-specific initiation on the same DNA molecules. This event does not appear to depend on the transcription process itself, but instead seems to be correlated with transcription complex assembly and chromatin remodeling events (Danis et al., 2004).

Only one additional ori besides the one in the rDNA locus has been identified in late *Xenopus* embryos. This ori is located ~3 kb upstream of the *Xenopus c-myc* gene, close to a DNase I hypersensitive site and overlapping with a nuclear matrix attachment region (MAR)

(Girard-Reydet et al., 2004). Importantly, no conserved sequence is observed when the *c-myc* and rDNA *Xenopus* oris are compared, or when the *Xenopus c-myc* ori is analyzed side by side with oris located in the human and mouse *c-myc* domains (Girard-Reydet et al., 2004).

1.4.4. Origins of DNA replication in insect cells

Similar to what is observed in *X. laevis*, DNA replication initiation patterns in insect cells appear to be developmentally regulated. Thus, during very early embryonic stages, replication initiates almost randomly from frequently spaced sites, whereas at later developmental phases DNA replication becomes confined to specific genomic regions (Sasaki et al., 1999). At least four such regions have thus far been identified and characterized. Three of them correspond to amplified loci, with only one insect ori being mapped within a single-copy locus.

In cultured *Drosophila melanogaster* cells, DNA replication has been shown to initiate from multiple sites throughout the histone gene repeating unit, a ~5 kb DNA segment which contains five histone genes and is present approximately 110 times in the haploid genome (Shinomiya and Ina, 1993). Multiple initiation sites also appear to exist within a 10 kb DNA region located downstream of the *D. melanogaster* single-copy gene that codes for the 180 kDa subunit of DNA polymerase α (Shinomiya and Ina, 1994). However, the most studied oris in insect cells are probably the ones located in the third-chromosome chorion gene cluster of *D. melanogaster* and in the DNA puff II/9A of the salivary glands of the fly *Sciara coprophila*.

During *Drosophila* oogenesis, somatic ovarian follicle cells switch from regular genomic DNA replication to programmed over-replication of specific genomic regions. Among the sequences that undergo re-replication are two clusters of chorion genes located on the X- and third-chromosomes. Amplification of these genes by repeated ori firing during a single S phase ultimately allows the rapid synthesis of a large quantity of chorion proteins which are essential for eggshell formation (Gerbi et al., 2003). Several sequence elements have been shown to be important for amplification of the third-chromosome chorion gene locus. These include a 320-bp DNA segment termed *a*mplification *c*ontrol *e*lement (ACE3) and four additional sequences known as *a*mplification *e*nhancing *r*egions (AER-a to AER-d) (Delidakis and Kafatos, 1989). Two-dimensional gel electrophoresis studies have shown that DNA synthesis can potentially initiate from multiple sites spread along the ~12 kb that comprise the chorion gene cluster. However, the same studies have revealed that between seventy and eighty percent of all initiation events occur within an 884-bp region, termed ori-β, which partially coincides with AER-d (Delidakis and Kafatos, 1989; Heck and Spradling, 1990; Lu et al., 2001). Surprisingly, these results demonstrated that DNA replication is not frequently initiated from the ACE3 element, which is located ~1.5 kb away from AER-d and was believed to be essential for ori function and even sufficient for low levels of amplification. The role of ACE3 in chorion gene amplification has only been recently clarified by *in vitro* and *in vivo* assays which showed that the *Drosophila* ORC2 protein preferentially binds to several regions of ACE3 as well as to AER-d (Austin et al., 1999). Thus, according to the current model, ACE3 constitutes a "replicator" element as initially defined in the replicon model, but is not a strong ori. Instead, binding of the ORC to ACE3

is believed to be required for the activation of DNA replication from the region around the AER-d element (Zhang and Tower, 2004; Gilbert, 2001).

Initial characterization of the ori from which DNA puff II/9A locus amplification takes place in *S. coprophila* was performed by two-dimensional electrophoresis (Liang et al., 1993). According to this study, DNA replication initiates from any of multiple potential sites distributed within a region of ~6 kb encompassing gene II/9-1. However, like in the chorion gene locus of *Drosophila*, most initiation events appeared to be confined to a small region of ~1 kb, centered ~2.5 kb upstream of the transcription start site of the *II/9-1* gene (Liang et al., 1993; Gerbi et al., 2003). An even more detailed analysis of replication initiation at this ori has been achieved in recent times by RIP mapping (Bielinsky et al., 2001). Notably, a precise and unique RIP was identified only for the top DNA strand, with multiple initiation sites but no precise TP being detected at or around the same position in the bottom strand (Bielisnky et al., 2001; Gerbi et al., 2003). This clearly contrasts to what has been observed in other eukaryotic oris mapped with similar resolution, such as the ARS1 and ars3001 oris of budding and fission yeast, respectively. In terms of protein-DNA interactions, ORC2 has been shown to preferentially interact *in vitro* and *in vivo* with a specific segment of the DNA puff II/9A locus located in the immediate vicinity of the top strand TP (Bielinsky et al., 2001).

Finally, similar to what is observed in *S. pombe* and in mammalian cells (section 1.4.5), no conserved sequence motif has been detected in insect oris. In fact, a recent genome-wide survey that allowed the identification of ~45 novel oris in the left arm of *Drosophila* chromosome 2 did not identify any consensus element amongst the ~491 ORC-binding sites found in this genomic region (MacAlpine et al., 2004).

1.4.5 Mammalian origins of DNA replication

Mammalian oris are highly heterogeneous and correspond to specific genomic regions that can vary in size from a few hundred base pairs to up to 55 kb. According to most estimates, between 30,000 and 60,000 oris exist in the genome of a typical mammalian cell (Hamlin, 1992; Aladjem et al., 2006). In light of these numbers, even if mammalian oris resembled budding yeast ARSs and contained an ori-specific consensus sequence motif, the process of identifying them all would still be an overwhelming task. However, like in fission yeast, insect cells and *X. laevis*, mammalian oris do not seem to share any sequence similarity. Therefore, it comes as no surprise that after more than twenty years of research only around 30 mammalian oris have been identified and characterized to different degrees (reviewed by Aladjem et al., 2006). This picture however, appears to be slowly changing in recent years, thanks to the complete sequencing of several mammalian genomes and the development of high-throughput, genome-wide ori mapping technologies (Todorovic et al., 2005; Touchon et al., 2005; Lucas et al., 2007).

Although the number of known mammalian oris might still be too small to constitute an unbiased sample and allow significant comparisons to be made, certain similarities have been observed amongst mammalian oris, despite their lack of sequence conservation (Table 1). For instance, most mammalian oris seem to be localized within intergenic regions, near transcriptional promoters, and in close to proximity to CpG islands (Delgado et al., 1998; Aladjem et al., 2006). They also appear to include AT-rich sequence elements and to map in the vicinity of MARs. Finally, several reports have raised the possibility that atypical DNA configurations such as bent DNA, cruciforms and triple-stranded DNA structures could be found near mammalian oris (Pearson et al., 1996; Kusic et al., 2005).

Table 1. Mammalian origins of DNA replication

Ori/Organism	Intergenic region-gene promoter	CpG island	AT-rich stretches
ALDOB/Rat	Y	N	Y
APRT/Hamster	N	Y	-
DHFR/Hamster	Y	N	Y
GADD45/Hamster	N	Y	-
TK1/Hamster	Y	Y	-
LMNB2/Human	Y	Y	Y
MCM4/Human	Y	Y	Y
HBB/Human	Y	N	Y
TOP1/Human	Y	Y	Y
MYC/Human	Y	Y	Y

Ten of the ~30 mammalian oris characterized to date are shown, along with the organisms in which they were identified. Presence (Y) or absence (N) of the indicated genomic features near the ori region is indicated. (-) denotes that information is not available for a particular ori.

These observations suggest that, similar to the situation in other organisms, chromatin context and organization might be more important than nucleotide sequence in the selection of mammalian oris.

In the following sections, three of the best characterized mammalian oris will be described, placing particular emphasis on the differences between them as well as on the features that they have in common.

1.4.5.1 Human c-myc ori

The existence of an ori in the vicinity of the human *c-myc* gene was first suggested when several cloned DNA segments mapping upstream of the *c-myc* transcriptional promoter were found to display strong ARS activity in human cells (Iguchi-Ariga et al., 1988; McWhinney and Leffak, 1990). Since then, numerous studies have demonstrated, using different techniques, that an ori is indeed localized in the human *c-myc* locus (Vassilev and Johnson, 1990; Waltz et al., 1996; Trivedi et al., 1998; Tao et al., 2000; Liu et al., 2003). Furthermore, oris have also been identified near the *c-myc* genes of other metazoans, including *Xenopus*, mouse and chicken (Phi-van et al., 1998; Girard-Reydet et al., 2004).

At the human *c-myc* locus, DNA replication has been shown to initiate from a broad region of ~12 kb that includes the area upstream of the *c-myc* transcriptional promoter, as well as the transcribed sequences and the 3'-end of the gene (Trivedi et al., 1998). Within this extended zone, initiation preferentially occurs at a ~2.4 kb DNA segment, known as the c-myc core origin. This region displays ori activity when moved to a different chromosomal location (Malott and Leffak, 1999), includes multiple potential replication initiation sites, and is localized immediately upstream of the *c-myc* P1 promoter (Waltz et al., 1996; Malott

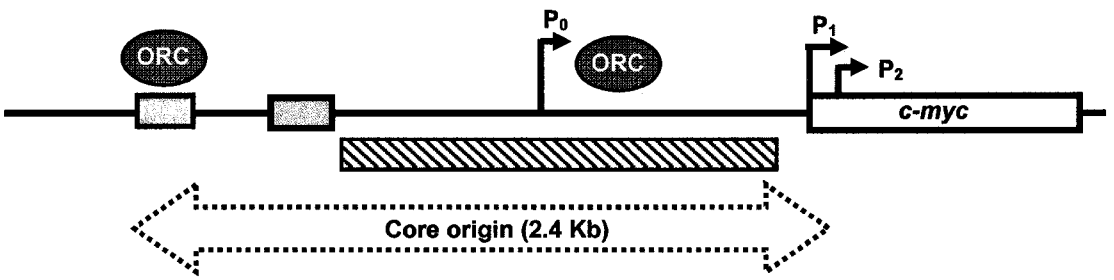
and Leffak, 1999; Tao et al., 2000) (Figure 3a). Other features of the human c-myc ori are shown in Figure 3a, including: (i) a DNA unwinding element (DUE) and a CpG island (Hay et al., 1987; Bazar et al., 1995); (ii) multiple transcription factor binding sites; and (iii) *in vivo* binding sites for replication proteins such as members of the human ORC (Ghosh et al., 2006).

1.4.5.2 Human lamin B2 ori

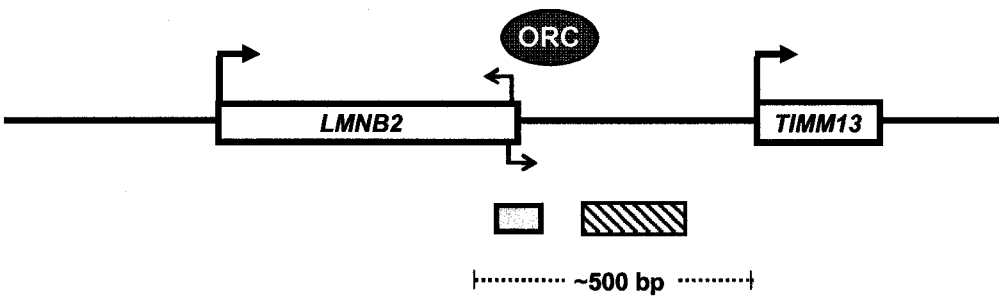
Originally identified through the isolation and cloning of DNA sequences replicated during early S phase (Tribioli et al., 1987), the ori located between the *lamin B2* and *TIMM13* human genes is, undoubtedly, the best characterized mammalian ori to date. Although this ori was initially shown to lie within a ~474 bp area spanning the end of the *lamin B2* gene and the promoter of the *TIMM13* gene (Giacca et al., 1994), subsequent studies identified the precise replication initiation sites with nucleotide resolution (Abdurashidova et al., 2000), making the lamin B2 ori the only mammalian ori for which such detailed information exists. Resembling the situation in yeast, unique initiation points were found by RIP mapping in the lamin B2 ori, in close proximity to each other on both parental DNA strands (Figure 3b) (Abdurashidova et al., 2000).

As expected, the lamin B2 ori interacts *in vivo* with several replication initiation proteins. Notably, a single binding site was reported for the ORC, close to the mapped RIPs (Figure 3b) (Abdurashidova et al., 2003). Finally, as is the case with many other mammalian oris, the lamin B2 ori is close to a CpG island, an asymmetric AT-rich stretch (Figure 3b), and includes binding sites for several transcription factors such as Sp1 and NRF-1 (Aladjem et al., 2006).

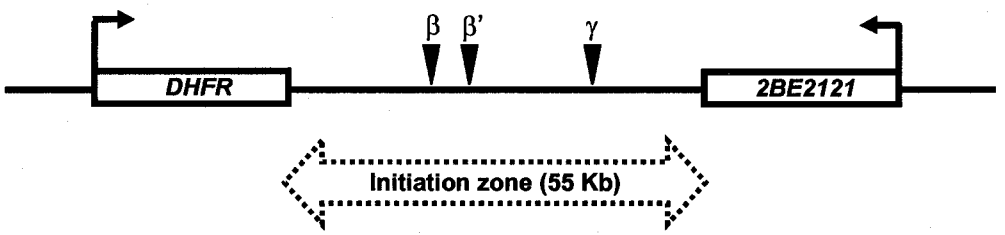
(a)



(b)



(c)



1.4.5.3 Hamster DHFR ori

The hamster dihydrofolate reductase (DHFR) ori is diametrically different from the human lamin B2 ori. In remarkable contrast to the precisely defined replication initiation sites identified in the latter, the DHFR ori spans ~55 kb and occupies the entire intergenic spacer between the *DHFR* and *2BE2121* hamster genes (Figure 3c).

Despite having been analyzed by all known ori mapping techniques, the exact nature and location of the DHFR ori is still controversial. On one hand, initiation events are detected throughout the 55 kb intergenic region by two-dimensional gel electrophoresis (Dijkwel et al., 1994). On the other hand, however, three preferred initiation sites (oriβ, β' and γ) are identified within this extended zone when nascent strand abundance analysis and the Okazaki fragment bias technique are used (Burhans et al., 1990; Kobayashi et al., 1998). Currently, the most accepted model proposes that the DHFR ori consists of a broad zone that includes multiple potential initiation sites, which are used with different efficiencies (Dijkwel et al., 2002). Broad initiation zones have also been reported for other metazoans oris including the Chinese hamster rhodopsin locus (Dijkwel et al., 2000) and the ribosomal DNA loci of *Xenopus* and human cells (Hyrien and Mechali, 1993; Little et al., 1993).

The above observations are clear evidence of the diverse nature of mammalian oris and suggest that they can correspond to either precisely localized origins of bidirectional replication (OBR), as exemplified by the lamin B2 ori, or to a diffused replication initiation zone like the oris at the *DHFR* and *c-myc* loci.

1.4.6 Eukaryotic initiation of DNA replication: the replication licensing model

The necessary increase in the number of oris in eukaryotic genomes carries with it several critical pitfalls. The most obvious of them is that cells must somehow develop regulatory mechanisms that would allow them to efficiently manage and control the activity of all these oris, ensuring that each one of them is activated **once and only once** during a given S phase. Failure of activating oris could potentially result in incomplete DNA replication, whereas repeated activation of an ori during a single S phase could generate harmful gene amplifications. Although, in theory, these events can also happen in cells that replicate their genome from a single ori, the probability of occurrence obviously increases when numerous initiation sites are used.

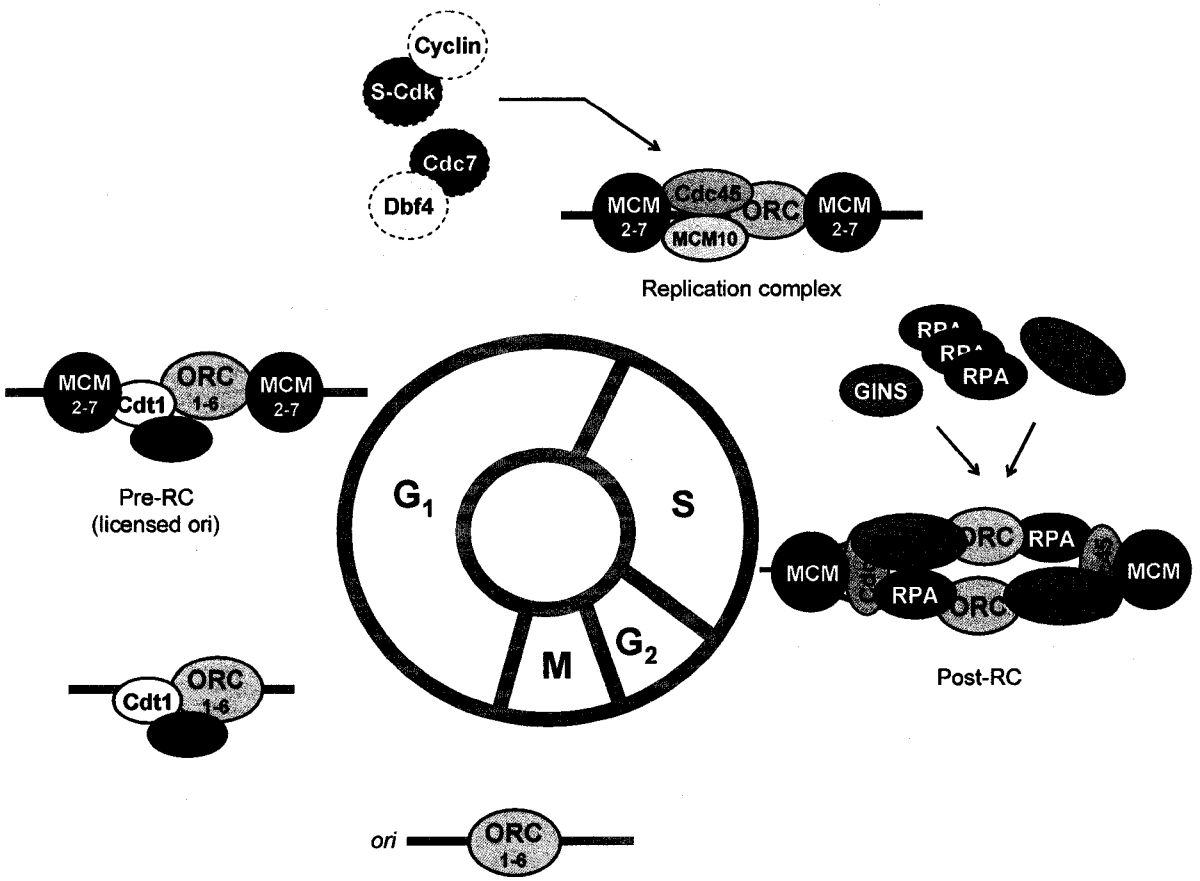
To circumvent this potential problem, eukaryotic cells have in place an intricate system that permits the regulation of DNA replication initiation at the level of individual oris. The heart of this control mechanism is the formation of a multiprotein complex known as the pre-replication complex (pre-RC). Assembly of this complex onto oris is an essential pre-requisite for the initiation of DNA synthesis, and involves orderly association of at least 14 different proteins (reviewed by Sivaprasad et al., 2006). Although most of the current knowledge regarding this process was initially derived from studies in the budding yeast, the same general principle seems to apply to higher eukaryotes, including human cells.

Pre-RC formation starts during G1 phase, and only occurs when the levels of S-phase-specific cyclin-dependent kinase (S-Cdk) activity are low (reviewed by Diffley, 2004). The first step in the assembly process is the ATP-dependent binding of the ORC to oris. In contrast to yeast, where ORC proteins form a hexameric complex that binds to oris throughout the cell cycle (Aparicio et al., 1997), human ORC subunits seem to be loaded

sequentially onto oris and to be partially disassembled during S phase (Siddiqui and Stillman, 2007). Once bound to DNA, ORC molecules recruit two additional initiator proteins, Cdc6 and Cdt1, which in turn promote binding of the hexameric MCM complex (MCM_{2-7}) to this region (Figure 4) (You et al., 1999; Takahashi et al., 2005). Following MCM binding, oris are said to be **licensed** or competent for DNA replication but are still in an inactive state. Origin firing and DNA synthesis initiation only occurs at the onset of S phase when CDK activity levels increase and the Cdc7-Dbf4 protein kinase is activated. Although the precise phosphorylation events that signal the initiation of DNA synthesis from individual oris are still largely unknown, the end result is the loading of additional protein factors onto oris, including the Cdc45 protein, Mcm10, GINS, RPA, and, eventually, DNA polymerases (Zou and Stillman, 2000; Labib and Gambus, 2007).

The entire process of pre-RC formation and activation is regulated by overlapping pathways at multiple stages. For instance, by only allowing pre-RC assembly to occur under conditions of low CDK activity, cells ensure that, once CDK activity rises and oris are fired during S phase, new pre-RCs cannot be assembled, thus preventing re-replication (reviewed by Diffley, 2004). Another example of the tight control exerted on pre-RC assembly is the existence of a Cdt1-specific inhibitor, called geminin, which is highly expressed throughout the S- and G2 phases in metazoan cells, also inhibiting unscheduled pre-RC assembly onto oris (Wohlschlegel et al., 2000).

Although the replication licensing system described above minimizes the possibility of re-replication, it does not address the issue of under-replication, which could be equally deleterious for a cell. In fact, since new pre-RCs cannot be assembled during an ongoing S



phase, if, for any reasons, replication forks are stalled or slowed down, entire genomic areas could remain unreplicated, leading to activation of checkpoint control mechanisms, cell cycle arrest or aberrant entry into mitosis (Bielinsky, 2003). To avoid this, metazoan origin selection is believed to follow the so-called **Jesuit model** (DePamphilis, 1993). This model suggests that an excess of potential oris exist in metazoan chromosomes, all of which assemble functional pre-RCs. However, during any given S phase, only a subset of these pre-RCs gets activated, with the rest being disassembled by the passage of moving replication forks. This ensures that the entire genome gets replicated by activating additional pre-RCs if necessary. Ori usage might be determined by epigenetic factors, with some oris being preferentially activated compared to others due to their chromatin context (i.e., nucleosome arrangement, binding of transcription factors, etc.) (Aladjem et al., 2006).

1.5 The human *DBF4* transcriptional promoter

As mentioned above, the Cdc7-Dbf4 kinase is essential for the activation of pre-RCs during S phase and for the initiation of DNA synthesis from individual oris (Bousset and Diffley, 1998; Donaldson et al., 1997). Thus, Cdc7-Dbf4 activity must be tightly regulated during cell cycle progression to prevent unscheduled ori firing. Since the level of the Cdc7 catalytic subunit remains largely constant throughout the cell cycle, regulation of kinase activity is believed to occur primarily through modulation of the DBF4 regulatory subunit (Jiang et al., 1999; Kumagai et al., 1999). To understand how DBF4 is regulated in mammalian cells, the transcriptional promoter of the human *DBF4* gene was previously identified and characterized (Wu and Lee, 2002; Yamada et al., 2002). According to these studies, a DNA segment located between positions -211 and -285 with respect to the DBF4

translation start-codon (taken as position +1), displays basal promoter activity. This region contains putative binding sites for several transcription factors including Sp1, TFIIB and *Mlu*I cell-cycle box (MCB) (Wu and Lee, 2002). Additional upstream sequences are required for full promoter activity and contain auxiliary and enhancer elements that remain to be identified. Transcription of the mitochondria carrier family protein (*MCFP*) gene, located ~20 kb upstream of the *DBF4* gene, also appears to utilize the *DBF4* promoter (Yamada et al., 2002). Finally, the promoter region includes a prominent CpG island and two stretches of AT-rich DNA. Considering that, as shown in Table 1, some of these features are characteristic of mammalian oris, it was initially hypothesized that the promoter region of the human *DBF4* gene might contain an ori.

1.6 Summary of work presented in this thesis

The development of recombinant DNA technologies during the late 1970s greatly expanded our knowledge of DNA replication, particularly in prokaryotes, eukaryotic viruses and yeast cells. However, despite significant improvements in the sensitivity of molecular techniques and almost 30 years of research, the most basic aspects of the DNA replication process in mammalian cells are still unclear. Most notably, it is still unknown what exactly constitutes an ori in higher eukaryotes, and how these sites are chosen from the hundreds of thousands of nucleotides that make up mammalian genomes. Certainly, one major reason for this lack of knowledge is the fact that very few mammalian oris (~0.1% of the estimated total) have actually been identified and characterized to date.

This thesis provides evidence for the existence of a novel ori around the transcriptional promoter of the human *DBF4* gene, further expanding the list of known

mammalian oris. More importantly, this novel ori has been mapped at the single nucleotide level and some of the protein-DNA interactions that take place at this locus during cell cycle progression have also been analyzed. This makes the DBF4 ori only the second human ori characterized in such detail with the other one being the lamin B2 ori, which, for almost 8 years, has been the paradigm for human oris.

Based on the available data, a model is presented which postulates that DNA replication from the DBF4 ori initiates asymmetrically from two separate "replication initiation zones". This mechanism is termed "asymmetric bidirectional replication" or ABR. Although previously reported for some bacterial and viral oris, this mode of initiation has never been described for a mammalian chromosomal ori. The ABR model is drastically different from what has been previously reported for the human lamin B2 ori (Abdurashidova et al., 2000). Thus, the results presented in this thesis provide an innovative and ground-breaking view of mammalian replication initiation and highlight the fact that, in order to fully understand this process and draw representative conclusions, it is absolutely necessary to detail the replication mechanism of many more mammalian oris.

Chapter 2

Materials and Methods

2.1 Cell culture and synchronization

HeLa cells were grown in RPMI-1640 medium containing 10% v/v fetal bovine serum (Hyclone), 2.05 mM L-glutamine, 100 μg/ml streptomycin and 100 units/ml penicillin. Human embryonic kidney cells (HEK293) were grown in Dulbecco's modified Eagle's medium (DMEM) containing 10% v/v fetal bovine serum and antibiotics. Human mammary epithelial 184B5 cells were cultured in serum-free, mammary epithelial basal medium (MEBM) supplemented with 1 ng/ml cholera toxin, 0.5 μg/ml hydrocortisone, 10 ng/ml human epidermal growth factor, 5 μg/ml insulin, 13 μg/ml bovine pituitary extract and 1% v/v GA-1000 (gentamicin-amphotericin mix, Clonetics). All cells were maintained at 37°C in a humidified atmosphere with 5% CO_2/95% air.

To synchronize cells in mitosis, exponentially growing cultures were incubated for 18 h in complete medium containing 50 ng/ml nocodazole (Cytoskeleton Inc.). Mitotic cells were collected by shake-off, washed with 1×PBS (2.7 mM KCl, 1 mM KH_2PO_4, 137 mM NaCl, 10 mM Na_2HPO_4, pH 7.5) and, if necessary, released back into the cell cycle by incubation in drug-free medium.

Synchronization of HeLa cells at the G1/S border was achieved by a double-thymidine block. Asynchronously growing cells were incubated for 18 h in complete medium containing 2 mM thymidine (Sigma). Cells were then washed once with 1×PBS, and maintained for 11 h in regular medium, followed by a second block with 2 mM thymidine for 14 h. Cells were finally washed with 1×PBS, and released into complete regular medium for the indicated times.

2.2 Flow cytometry

Cell cycle position following synchronization was confirmed by flow cytometric analysis of DNA content. To do this, cells were harvested by trypsinization, washed once with 1×PBS, and fixed in 70% v/v ethanol for at least 18 h at -20°C. The cells were then washed twice with 1×PBS and incubated in propidium iodide staining solution (0.1% w/v sodium citrate, 0.3% v/v NP-40, 100 μg/ml propidium iodide, 100 μg/ml RNase A) for 1 h at room temperature. Samples were analyzed using a Beckman Coulter Epics Elite flow cytometer.

2.3 Isolation of nascent DNA

Isolation of short single stranded nascent DNA by alkaline gel electrophoresis was carried out as previously described (Kamath and Leffak, 2001) with some modifications. Approximately 10^7 cells were collected by trypsinization, washed once with 1×PBS, and resuspended in 260 μl of 1×PBS containing 10% v/v glycerol. To minimize DNA damage during sample preparation, 85 μl of the cell suspension were directly loaded into the wells of a 1.25% w/v agarose gel containing 50 mM NaOH and 1 mM EDTA. Cells were immediately lysed by adding 25 μl of 5×SDS Lysis buffer (50 mM Tris-HCl, 5 mM EDTA, 0.5% SDS w/v, pH 8.0). Following incubation for 10 min at room temperature, EDTA and NaOH were added to each well (in that order) to increase their concentration to 20 mM and 50 mM, respectively. NaOH and EDTA were also added to DNA size markers of the appropriate range (New England Biolabs), which were loaded next to the cell samples in every gel. Electrophoresis was carried out at room temperature for >8 h at 15 volts in freshly made alkaline running buffer (50 mM NaOH, 1 mM EDTA). Gels were then

neutralized with 1×TAE (40 mM Tris, pH 8.0, 20 mM acetic acid, 1 mM EDTA) for 45 min at room temperature. Lanes containing DNA ladders were separated from the rest of the gel, and stained with 1×TAE buffer containing 0.6 μg/ml ethidium bromide for 15 min. The size of the nascent DNA in the non-stained gel was determined by comparison to the stained size markers. Nascent DNA of the desired size fractions was isolated from the gel using a gel extraction kit (QIAGEN), according to the manufacturer's instructions (final elution volume was 35 μl/column).

2.4 Nascent strand abundance assays by quantitative PCR (Q-PCR)

Q-PCR was carried out in a 7900HT ABI Prism system (Applied Biosystems). Reactions were performed in triplicate in a final volume of 25 μl containing 12.5 μl of 2× SYBR Green Master Mix (Applied Biosystems) and 300 nM of each primer (Integrated DNA Technologies, Coralville, IA). Primer sequences are shown in Appendix I. Standard curves were generated for each primer pair using serial dilutions of total genomic DNA (10 ng, 1 ng, 0.1 ng and 0.01 ng). An equal volume of gel-extracted nascent DNA (usually 2 or 3 μl) was used as template for amplification with every primer set. Following initial denaturation for 10 min at 95°C, amplification was carried out for 40 cycles as follows: 95°C for 30 s, 57°C for 30 s, and then 30 s at 72°C. Amplification of a single product of the correct size was verified by melting curve analysis and by agarose gel electrophoresis. Relative sequence abundance in a given nascent DNA preparation was estimated by calculating the ratio between the amount of each DNA segment (as estimated by the standard curves) and the amount amplified by the Prom3 primer set in the same DNA

sample. Primer pair Prom3 was chosen as a control because it consistently showed the lowest level of amplification.

2.5 *In vitro* DNA binding assays

2.5.1 Preparation of nuclear extracts

Nuclear extracts were prepared from asynchronous HeLa cells according to the procedure described by Schreiber et al. (1989), with some modifications. Approximately 10^6 cells growing in one 6-cm tissue culture plate were trypsinized, collected by centrifugation and washed twice with 1 ×TBS (50 mM Tris-HCl pH 7.5, 150 mM NaCl). The cell pellet was then gently resuspended in 400 µl of cold Buffer A (10 mM HEPES pH 7.9, 10 mM KCl, 0.1 mM EDTA pH 8.0, 0.1 mM EGTA pH 8.0, 1 mM DTT, and 0.5 mM PMSF) and incubated for 15 min on ice. Nonidet NP-40 was added to the cell suspension to a final concentration of 0.58% v/v, followed by vigorous vortexing for ~10 seconds and centrifugation for 3 min at 400 *g*. After discarding the supernatant, the pellet was resuspended in 50 µl of ice-cold Buffer C (20 mM HEPES pH 7.9, 0.4 M NaCl, 1 mM EDTA pH 8.0, 25% v/v glycerol, 1 mM DTT, and 1 mM PMSF), and incubated for 15 min at 4°C with shaking. Extracts were cleared by centrifugation at maximum speed in a microcentrifuge for 5 min at 4°C and the supernatant transferred to a fresh tube. Supernatants obtained from ten plates were pooled together to perform a desalting/buffer exchange step. For this, pooled extracts (final volume 500 µl) were loaded onto a Microcon YM-10 column (Millipore) and diluted by a final factor of at least 1/130 with Buffer I (22.5 mM HEPES pH 7.0, 1 mM Tris pH 7.5, 20 mM KCl, 5 mM $MgCl_2$, 0.2 mM EDTA, 9% v/v

glycerol, 1 mM DTT, and 0.5 mM PMSF) through a series of centrifugation and dilution cycles. Column centrifugation was performed at room temperature at 14,000 g.

2.5.2 DNA labeling and coupling to streptavidin-coated magnetic beads

Biotin-labeled DNA fragments were generated by PCR using the primers shown in Appendix III. All "forward" primers were labeled on their 5'-ends with biotin (Integrated DNA Technologies). PCR reactions were performed in a final volume of 50 µl and contained: 1 ×PCR Buffer (ID Labs Biotechnology), 2 mM $MgSO_4$, 400 nM of each oligonucleotide, 1 M betaine (Sigma), 200 µM of each dNTP (Invitrogen), 0.75 units of IDXtra high fidelity DNA polymerase (ID Labs Biotechnology), and 100 ng of sonicated HeLa genomic DNA as template. Amplification was carried out according to the following program: (i) 95°C for 4 min, (ii) 35 cycles consisting of denaturation at 95°C for 30 s, primer annealing for 30 s at 55°C and DNA extension for 1 min at 72°C, and (iii) final extension at 72°C for 7 min. PCR products were separated by agarose gel electrophoresis, purified by gel extraction using a commercially available kit (QIAGEN), and then quantified by spectrophotometry (Beckman DU-530).

Labeled PCR products were coupled to streptavidin-coated magnetic beads (Dynabeads M-280 Streptavidin, Invitrogen) following the supplier's recommendations. First, magnetic beads (250 µg/*in vitro* binding reaction) were washed twice in 1 ×BW buffer (5 mM Tris-HCl pH 7.5, 0.5 mM EDTA pH 8.0, 1 M NaCl). After the second wash, beads were resuspended in 2 ×BW buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA pH 8.0, 2 M NaCl) to a final concentration of 5 µg/µl. Biotinylated DNA was then added and the reaction volume adjusted with sterile water to make the final NaCl concentration in the binding

mixture 1 M. Samples were incubated for 15 min at room temperature with gentle rotation. Supernatants were then collected and stored to check binding efficiency by agarose gel electrophoresis. Beads were washed three times with 1 ×BW buffer and once with Buffer I before proceeding to the protein binding assays.

2.5.3 *In vitro* protein-DNA binding reactions

For every set of reactions the same nuclear extract preparation was utilized. Equal protein amounts were used in each reaction, ranging from 25 to 40 μg, depending on the quality of individual extracts. Protein extracts were pre-incubated for 1 h at room temperature with 1 mM ATP and the appropriate concentration (from 75 μg/ml to 7.5 μg/ml) of poly(dI-dC)•poly(dI-dC) non-specific competitor DNA (Amersham), all in a final volume of 40-50 μl (adjusted with Buffer I). Next, the extracts were added to tubes containing magnetic beads coated with each individual biotinylated PCR product. For the LB2 and B13 DNA segments, 500 ng of each PCR product were bound to streptavidin-beads. Based on the size of the fragments, these amounts correspond to 3.1 and 4.5 pmoles of DNA, respectively. In the case of the *DBF4* locus fragments, typical reactions included 750 ng of DNA, which is equivalent to ~1.5 pmoles. Poly(dI-dC)•poly(dI-dC) concentrations were adjusted in order to obtain similar biotin-DNA/non-specific DNA competitor **molar ratios** in all cases. Calculations were done assuming an average length of ~2.26 kb for the double-stranded competitor DNA (based on the information provided in the product's data sheet). Protein extracts and DNA-coated beads were incubated for 30 min at room temperature with rotation. Following this, supernatants were collected and stored to analyze protein binding efficiency. In order to cross-link bound proteins to DNA, beads

were resuspended in 50 μl of Buffer I containing 1% w/v formaldehyde (Sigma), and then incubated for 5 min at room temperature. Following two washes with Buffer I, beads were finally resuspended in 20 μl of Buffer I. Five micro liters of 5×SDS-PAGE loading buffer (0.1 M Tris-HCl pH 6.8, 40% v/v glycerol, 10% w/v lithium dodecyl sulfate, 2% w/v DTT, 0.05% w/v bromophenol blue) were added to each bead suspension, and samples were then boiled for 30 min to reverse the cross-links. Bound proteins were detected by Western blotting after SDS-PAGE and semi-dry transfer to Hybond-P PVDF membranes (Amersham). Antibodies used for Western blotting were: mouse monoclonal anti-ORC2 (KAM-CC235; StressGen Biotechnology), rabbit polyclonal anti-ORC4 (sc-19725, Santa Cruz Biotechnologies), and mouse monoclonal anti-MCM7 antibodies (sc-9966, Santa Cruz Biotechnologies).

2.6 Chromatin immunoprecipitation (ChIP) assays

ChIP assays were performed following Abcam's X-ChIP protocol (http://www.abcam.com/ps/pdf/protocols/x_chip_protocol.pdf) with some modifications. Exponentially growing HeLa cells on three 10-cm dishes were washed with 1×PBS and cross-linked for 10 min at room temperature in serum-free medium containing 1% w/v formaldehyde (Sigma). Cells were then washed and scraped into 2 ml of 1×PBS containing 500 μM PMSF. Cell suspensions were pooled together and centrifuged at 700 *g* for 10 min at 4°C. If necessary, protease inhibitors and PMSF were added to the cell pellets which were stored at -80°C. To lyse the cells, cell pellets were resuspended in 1.5 ml of X-ChIP lysis buffer (50 mM HEPES-KOH pH 7.5, 140 mM NaCl, 1 mM EDTA pH 8.0, 1% v/v Triton X-100, 0.1% w/v sodium deoxycholate, 0.1% w/v SDS, 1 mM PMSF and protease

inhibitors). DNA was fragmented by sonication to an average size of ~300 bp (6 pulses of 30 s each, at setting 5 in a Sonic Dismembrator Model 100, Fisher Scientific). Sonicated lysates were cleared by centrifugation at maximum speed in a microcentrifuge for 1 min at 4°C. The supernatant was transferred to a fresh tube and stored at -80°C until use. An aliquot of the sonicated cell lysate was used to determine the size of the fragmented chromatin and the final DNA concentration using the following procedure. 450 µl of fresh Elution buffer (1% w/v SDS, 100 mM $NaHCO_3$) and 5 µl of Proteinase K (20 mg/ml) were added to 50 µl of sonicated lysate, followed by overnight incubation at 65°C to reverse the cross-links. After phenol-chloroform extraction and ethanol precipitation, DNA was resuspended in 100 µl of sterile water and quantified by spectrophotometry (Beckman DU-530). DNA size was estimated by electrophoresis in 0.8% w/v agarose gels. This purified DNA preparation was also used for Q-PCR as the "Input" DNA sample.

Twenty five micrograms of DNA from cell lysates (based on the concentration values determined above) were used for each immunoprecipitation reaction. Cell lysates were diluted 1/10 using ChIP dilution buffer (1% v/v Triton X-100, 2 mM EDTA pH 8.0, 150 mM NaCl, 20 mM Tris-HCl pH 8.0, 1 mM PMSF and protease inhibitors) and pre-cleared with 50 µl of a 1:1 mixture of Protein A-agarose beads (Upstate) and Protein G-agarose beads (Upstate) for 3 h at 4°C. Immunoprecipitation was carried out for 4 h at 4°C using 4 µg of the appropriate antibody or 4 µg of the corresponding normal IgG as negative control. Antibodies used were all from Santa Cruz Biotechnologies as follows: (i) rabbit polyclonal anti-ORC4, sc-20634; (ii) normal rabbit IgG, sc-2027; (iii) goat polyclonal anti-Sp1, sc-59-G; (iv) normal goat IgG, sc-2028; (v) goat polyclonal anti-MCM2, sc-9839; (vi) goat polyclonal anti-MCM3, sc-9850; (vii) mouse monoclonal anti-MCM7, sc-9966; and (viii)

goat polyclonal anti-MCM6 (sc-9845). Immunocomplexes were recovered by overnight incubation at 4°C with 50 µl of a 1:1 mixture of Protein A- and Protein G-agarose beads. Beads were washed four times with Washing buffer (0.1% w/v SDS, 1% v/v Triton X-100, 2 mM EDTA pH 8.0, 150 mM NaCl, 20 mM Tris-HCl pH 8.0) and once with Final washing buffer (0.1% w/v SDS, 1% v/v Triton X-100, 2 mM EDTA pH 8.0, 500 mM NaCl, 20 mM Tris-HCl pH 8.0). Elution was performed with 450 µl of freshly made Elution buffer for 15 min at room temperature. To reverse cross-links, eluted DNA was incubated with Proteinase K for 5 h at 65°C. Following phenol-chloroform extraction and ethanol precipitation, the recovered DNA was resuspended in 100 µl of sterile water. An equal volume of each DNA preparation was used as template for Q-PCR, as described for the nascent strand abundance assays (section 2.4). For each immunoprecipitation reaction, values generated by Q-PCR were normalized by subtracting the amount recovered with the appropriate pre-immune (normal) IgG control. Relative abundance was then calculated by comparing the value of each DNA segment recovered after immunoprecipitation to that in a sample taken prior to immunoprecipitation ("Input" DNA) and normalizing it to the value obtained for the Prom3 or Prom 2 primer pairs.

2.7 Replication Initiation Point (RIP) mapping

For RIP mapping, nascent DNA was isolated from HeLa cells as described in section 2.3. To analyze leading strand RIPs only, 2 µM emetine (Sigma) was added to the cell culture media for 1 h prior to DNA isolation. Where indicated, aphidicolin (Sigma) was also added. Equal volumes (usually 5 µl) of a single DNA preparation (or equal amounts, measured by Q-PCR, when comparing different DNA preparations) were used as template in

one-way PCR reactions with 5'-digoxigenin-labeled primers (Integrated DNA Technologies; primer sequences are shown in Appendix IV). Reactions were carried out in a final volume of 30 µl containing 1×PCR buffer (ID Labs Biotechnology), 200 µM dNTPs, 3 mM $MgSO_4$, 400 nM digoxigenin-labeled oligonucleotide, 1 M betaine and 1 unit of IDPol DNA polymerase (ID Labs Biotechnology). For the lamin B2 ori RIP mapping experiments, reactions were carried out in the absence of betaine. PCR conditions were as follows: initial incubation for 5 min at 95°C, 30 cycles of amplification (1 min at 94°C; 1 min at the appropriate annealing temperature as shown in Appendix IV; and 1.5 min at 72°C), a final extension step of 7 min at 72°C and incubation for 10 min at 4°C. As a negative control, a sample without template DNA was included in every experiment.

Following amplification, 20 µl of "sequencing" loading buffer (98% v/v formamide, 10 mM EDTA, 0.1% w/v xylene cyanol, 0.1% w/v bromophenol blue) were directly added to each sample. Samples were then heated for 4 min at 90°C, chilled quickly on ice and separated by PAGE in 5% w/v acrylamide gels containing 7.7 M urea. Gels were prepared in 0.5×TBE (44.5 mM Tris, 44.5 mM boric acid, 1 mM EDTA, pH 8.4), which was also used as running buffer. To estimate the size of the generated products, digoxigenin-labeled DNA size markers (DNA molecular weight marker V DIG-labeled, Roche Cat. No 1 082 170 5001) were run alongside RIP samples on every gel. After transferring onto Hybond N^+ nylon membranes (Amersham) for 1 h at 400 mA using a semi-dry transfer apparatus, the DNA was fixed with UV for 6 min. Detection of digoxigenin-labeled PCR products with anti-digoxigenin antibodies (Roche) was performed according to the manufacturer's instructions. First, membranes were blocked for 1 h at room temperature in 1×Blocking solution [1% w/v Blocking reagent (Roche) in Maleic Acid buffer (100 mM maleic acid,

150 mM NaCl, pH 7.5)]. Next, membranes were incubated overnight at 4°C with alkaline phosphatase-conjugated anti-digoxigenin antibodies (1/10,000 dilution in 1 ×Blocking solution). Membranes were then washed twice for 15 min at room temperature in Washing solution (0.03% v/v Tween-20 in Maleic Acid buffer) and once for 1 min in Detection buffer (100 mM Tris-HCl pH 9.5, 100 mM NaCl). Finally, membranes were incubated for 5 min at room temperature with Substrate solution [1/100 dilution of CDP-Star substrate (Roche) in Detection buffer] before X-ray film exposure.

2.8 DNase I-sensitivity assays

Nuclei isolation and DNase I digestion was performed according to the protocol described by Yan et al. (2006). Asynchronously growing HeLa cells in a 10-cm tissue culture dish ($\sim 10^7$) were trypsinized, resuspended in 1 ml RSB buffer (10 mM Tris, pH 7.5, 10 mM NaCl, 3 mM $MgCl_2$) containing 0.2% v/v NP-40, and incubated for 5 min on ice. Nuclei were then isolated by centrifugation for 5 min at 500 *g*. The nuclei pellet was resuspended in 600 µl of RSB buffer, and divided into six 100 µl aliquots. The appropriate amount of DNase I (Invitrogen) was added to each aliquot, followed by incubation for 5 min at 37°C. Digestion reactions were stopped by adding an equal volume of Stop solution (1% SDS w/v, 600 mM NaCl, 10 mM EDTA, 20 mM Tris-HCl, pH 8.0). Proteins in each sample were digested overnight with Proteinase K (Sigma), followed by phenol-chloroform extraction and ethanol precipitation. Finally, purified DNA was resuspended in 1 ×TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0), and quantified by spectrophotometry (Beckman DU-530).

Q-PCR was performed using 8 ng of DNA template for each sample. Standard curves were generated using serial dilutions of independently isolated, undigested total genomic DNA. Reaction mixtures (25 µl) contained 12.5 µl of 2×SYBR Green Master Mix (Applied Biosystems) and 300 nM of primers. Following amplification, the amount of each DNA segment remaining in the digested samples was determined using the standard curves. Results were compared to the amount detected in control samples not treated with DNase I.

2.9 RNA isolation and Northern blotting

Total RNA extraction and purification was performed using the RNeasy Mini kit from QIAGEN, according to the manufacturer's instructions. Twenty five micrograms of total RNA were loaded on 1.3% w/v agarose gels prepared in 1×MOPS buffer (20 mM MOPS pH 7.0, 5 mM sodium acetate, 1 mM EDTA pH 8.0) containing 1.85% w/v formaldehyde. Gels were run with 1×MOPS and transferred to Hybond N^+ membranes (Amersham) for 4 h with RNA transfer buffer (0.75 M NaCl, 75 mM sodium citrate, pH 7.0, 10 mM NaOH) using the downward capillary method. Membranes were briefly rinsed in 2× SSC (300 mM NaCl, 30 mM sodium citrate, pH 7.0), fixed by UV treatment for 5 min and pre-hybridized for 5 h at 50°C in High SDS buffer [7% SDS w/v, 50% v/v deionized formamide, 5×SSC, 2% w/v Blocking reagent (Roche), 0.1% w/v N-lauroylsarcosine, 50 mM sodium phosphate, pH 7.0]. Hybridization was performed overnight at 50°C using a digoxigenin-labeled probe corresponding to the coding region of the Chinese hamster *Dbf4* gene (Guo and Lee, 2001). Membranes were washed twice at room temperature in 2× Washing solution (2×SSC +0.1% w/v SDS) and twice at 68°C in 0.5×Washing buffer

(0.5 ×SSC +0.1 % w/v SDS). Finally, digoxigenin signal was detected using CDP-Star substrate (Roche) as described above (section 2.7).

Chapter 3

Results

3.1 Identification of an ori in the transcriptional promoter of the *DBF4* gene

3.1.1 Nascent strand abundance assays

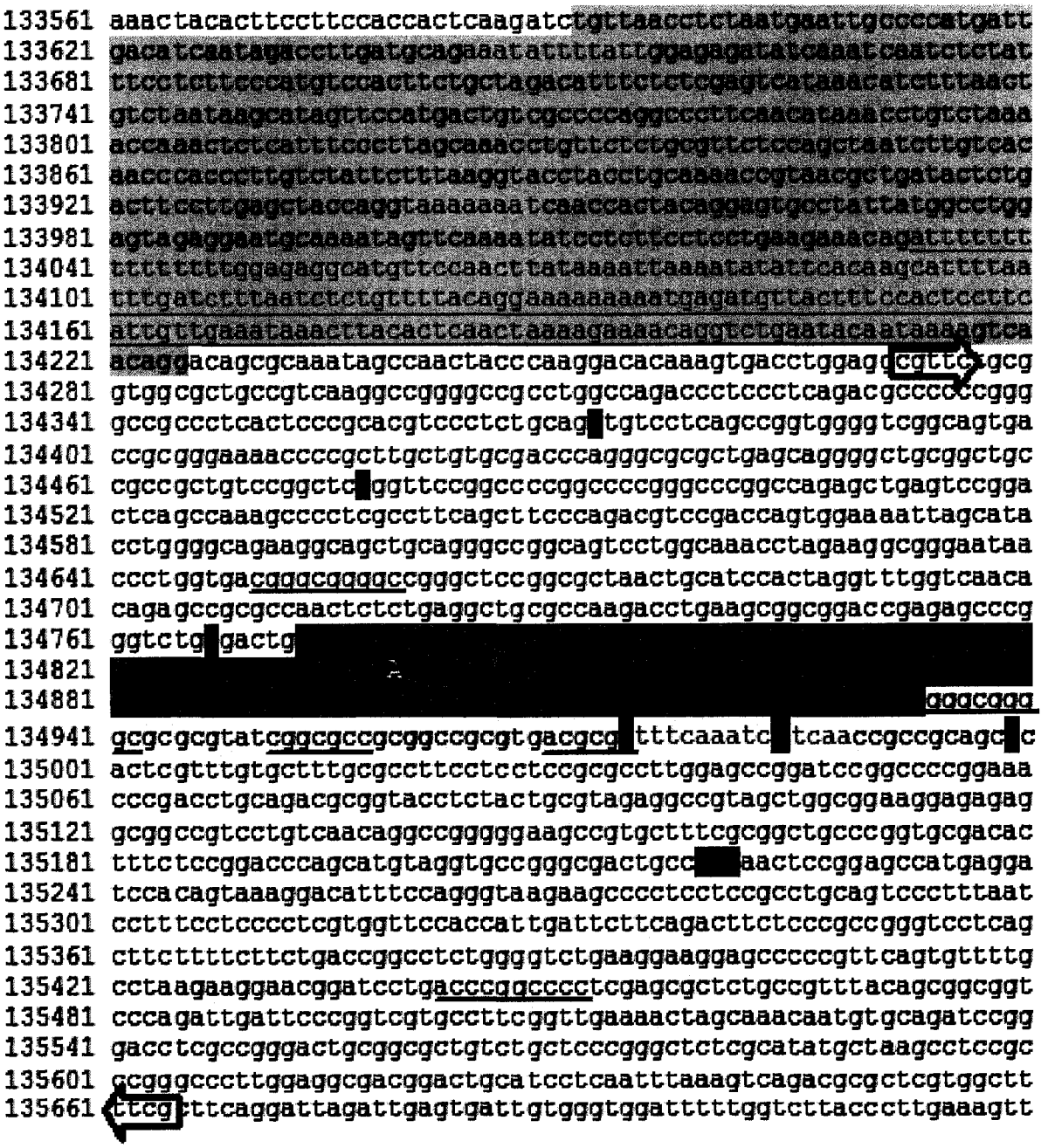
The nucleotide sequence and some of the most relevant features of the *DBF4* promoter region are shown in Figure 5. Since this genomic area includes many of the elements characteristic of mammalian oris (i.e., intergenic spacer, CpG island and AT-rich stretches), it was postulated that it could function *in vivo* as both a transcriptional promoter **and** a chromosomal ori.

To test this hypothesis, nascent strand abundance assays were performed. This technique, which is routinely used for the identification of oris (Yoon et al., 1995; Kobayashi et al., 1998; Keller et al., 2002), is based on the isolation of short, single-stranded, nascent DNA molecules (Giacca et al., 1997). Ideally, the nascent DNA should be short enough (1–2 kb) to allow high-resolution ori mapping, but should also be sufficiently long (at least 0.5 kb) to exclude Okazaki fragments which are expected to be evenly distributed throughout the genome in a population of exponentially growing cells and therefore could result in a low signal-to-noise ratio. The abundance of specific sequences in the nascent DNA sample is then analyzed using several different techniques, including Southern blotting, competitive PCR, and, more recently, quantitative PCR (Q-PCR). It is expected that DNA segments located close to an ori will be highly abundant in the nascent DNA, compared to those distant from replication initiation sites.

Short nascent DNA strands (1–2 kb) were isolated from **asynchronous** HeLa cells by denaturing gel electrophoresis as previously described (Kamath and Leffak, 2001). To minimize sample handling and prevent DNA fragmentation, cells were directly lysed in the

Figure 5. Nucleotide sequence and relevant features of the human *DBF4* promoter region
The sequence and the numbering to the left are from BAC clone CTB-60N22 (accession #AC003083). The DBF4 translation start codon (ATG) is shown in upper case letters with red highlight at positions 135219–135221. Numbers in parenthesis (see below) correspond to the relative position from the "A" of the ATG codon, which is taken as +1; "–" and "+" denote up- and downstream from the ATG, respectively.

- The core promoter of the *DBF4* gene is highlighted in yellow and extends from position 134934 (–285) to 135008 (–211) (Wu and Lee, 2002). Red letters within the core promoter mark a 15-mer AT-rich segment [AT-rich I; 134974 (–245) to 134988 (–231)].
- The major transcription initiation site for the *DBF4* mRNA is located within the AT-rich stretch of the core promoter, and is shown as a blue, upper case "T" with purple highlight at position 134984 (–235) (Wu and Lee, 2002). Additional transcription start-sites have been identified and are also shown as blue letters (lower case) with purple highlight ["c" at 134999 (–220); "t" at 134974 (–245); "a" at 134914 (–305); "c" at 134897 (–322); "c" at 134888 (–331); "a" at 134767 (–452); "c" at 134477 (–742); "c" at 134372 (–847)] (Wu and Lee, 2002).
- The auxiliary promoter element (blue highlight) extends from 134772 (–447) to 134934 (–286) (Wu and Lee, 2002).
- The 865-bp DNA segment between 133591 (–1628) and 134225 (–994) (highlighted in gray) contains a putative enhancer element (Wu and Lee, 2002). A 184-bp highly AT-rich segment (AT-rich II) spanning from 134033 (–1186) to 134216 (–1003) is located within the putative enhancer and is shown as red, underlined letters. Two additional AT-rich tracks within the enhancer are also marked by red letters.
- According to NCBI gene bank analysis, the DNA segment spanning from 134272 (–947) to 135664 (+445) is a putative CpG island (the start and end are marked by green arrows; % of G/C=66.5).
- The *MCFP* gene is transcribed in the opposite direction starting from the white, upper case "A" with black background shown at position 134839 (–380) (Haitina et al., 2006).
- Putative binding sites for the Sp1 transcription factor were identified using the TFSEARCH program (www.cbrc.jp/research/db/TFSEARCH.html; Heinemeyer et al., 1998) and are underlined in blue [–569 to –560; –430 to –421; –361 to –352; –315 to –306; –286 to –277; +224 to +233]. Sp1-binding has been confirmed for the site at –361 to –352 (Wu and Lee, 2002).
- Putative binding sites for TFIIB (–268 to –262) and *Mlu*I (–250 to –245) are underlined in green and black, respectively.



well of an alkaline agarose gel. Nascent DNA of the desired size was purified by gel extraction and used as template in Q-PCR reactions with primer pairs distributed throughout the *DBF4* locus (primers positions and sequences are shown in Fig. 6a and in Appendix I). As shown in Figure 6b (hatched bars), DNA segments located near the *DBF4* promoter (primers Prom7B and Prom8) were ~8- to 10-times more abundant than distal sequences in the nascent DNA preparations, indicating that they might lie in the vicinity of an active ori. Similar results were obtained with asynchronous HEK293 and 184B5 cells (Fig. 7), suggesting that this novel ori is used in human cells of different histological derivations. Enrichment values for the *DBF4* promoter were much greater than those obtained with the well-known lamin B2 and c-myc oris (Fig. 6b and Fig. 7), which indicates that the DBF4 ori is utilized by a considerable proportion of cells within a population.

To rule out that the results described above were due to artifacts generated during the DNA isolation procedure, the same methodology was applied to analyze "nascent strands" isolated from cells arrested in **mitosis** by nocodazole treatment (Fig. 6b, white bars). No substantial enrichment was observed in these samples with any of the primer sets tested, confirming that the nascent strand abundance data was not generated by artifacts such as DNA fragmentation. In addition, equal volumes of sonicated, total genomic DNA from asynchronous HeLa cells were also used as template in Q-PCR reactions to test the validity of the Q-PCR quantitation (Fig. 6b, black bars). Equal amounts were obtained with every primer pair in these experiments, supporting the reliability of the Q-PCR procedure.

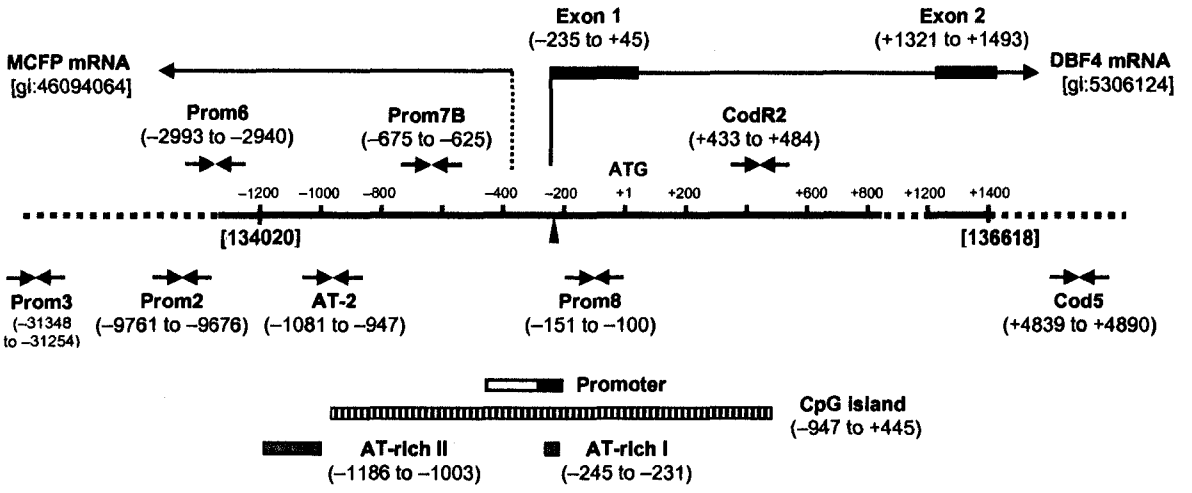
During the process of designing and selecting the Q-PCR primers used for the nascent strand abundance assays described above, it was noted that, besides the *DBF4* gene located on chromosome 7, the human genome contains a *DBF4* "pseudogene" that maps to

Figure 6. The human *DBF4* transcriptional promoter contains an ori

(a) Map of the human *DBF4* promoter region. Numbers in parenthesis correspond to relative positions with respect to the "A" of the DBF4 translation start codon (ATG), which is taken as +1. The major *DBF4* transcription initiation site at position -235 is indicated by a vertically elongated triangle (Wu and Lee, 2002). The putative transcription initiation site for the *MCFP* gene is shown by a vertical broken line (Haitina et al., 2006). Primers used to quantify nascent strand abundance are shown as converging arrows (primer details in Appendix I). White and black boxes at the bottom represent the *DBF4* auxiliary and core promoter elements, respectively. The hatched box marks a CpG island, while the gray boxes indicate the position of two prominent AT-rich segments (AT-rich I and AT-rich II).

(b) Nascent DNA was isolated from exponentially growing (hatched bars) or mitotic (white bars) HeLa cells. DNA abundance was quantified by Q-PCR with the indicated primer pairs (x-axis). In every experiment, values were normalized to the amount obtained with primer Prom3, which consistently displayed the lowest value (y-axis). Primers to the lamin B2 and c-myc oris were included as positive controls. Black bars, relative abundance of the indicated DNA segments compared to Prom3 in preparations of total sonicated genomic DNA from asynchronous HeLa cells. Error bars, standard error of the mean of at least two independent experiments.

(a)



(b)

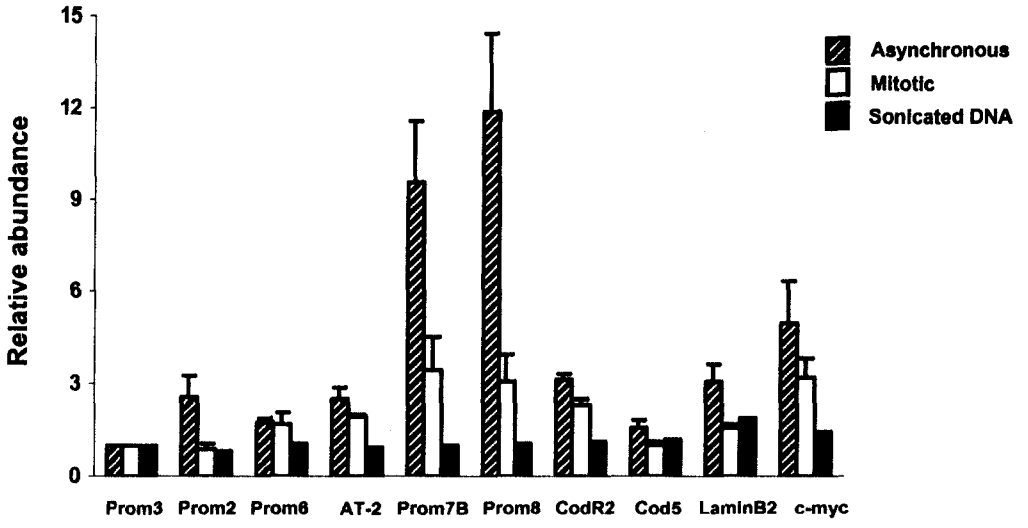
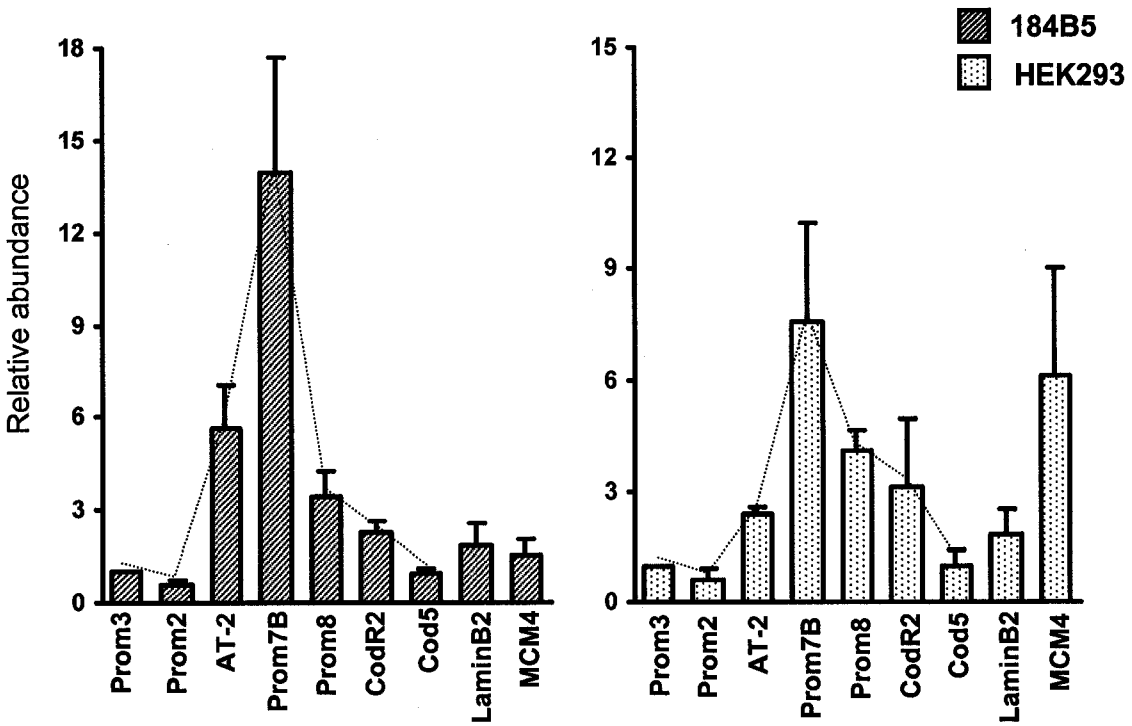


Figure 7. The DBF4 ori is active in HEK293 and 184B5 cells
Short strands of nascent DNA were isolated from asynchronous human embryonic kidney cells (HEK293) and immortalized normal breast cells (184B5). Nascent strand abundance assays were performed as described in the Figure 6b legend. Abundance of the indicated DNA segments was quantified by Q-PCR and is shown relative to the amount amplified by primer set Prom3. Primers for the human lamin B2 and MCM4 oris were included as positive controls. Error bars, standard error of the mean of three independent experiments.



chromosome 10. This locus includes all of the *DBF4* 12 exons, with the exception of exon 11, but lacks any of the intervening introns (Fig. 8a). The nucleotide sequence of the *DBF4* exons on chromosome 7 and chromosome 10 are 97% identical. In contrast, the flanking sequences immediately upstream of exon 1 are completely different between *DBF4* and its pseudogene. Since the Prom7B primer pair is specific for the *DBF4* locus on chromosome 7, it was safe to conclude from the experiments shown in Figure 6b that an ori exists at this genomic position. However, the Prom8 segment maps within the first exon of the *DBF4* gene, located on both chromosomes (Fig. 8a). Thus, the high enrichment values obtained with this oligonucleotide set could theoretically result from nascent strands emanating from both the chromosome 7 and chromosome 10 *DBF4* copies, raising the possibility that an ori also exists near the *DBF4* "pseudogene".

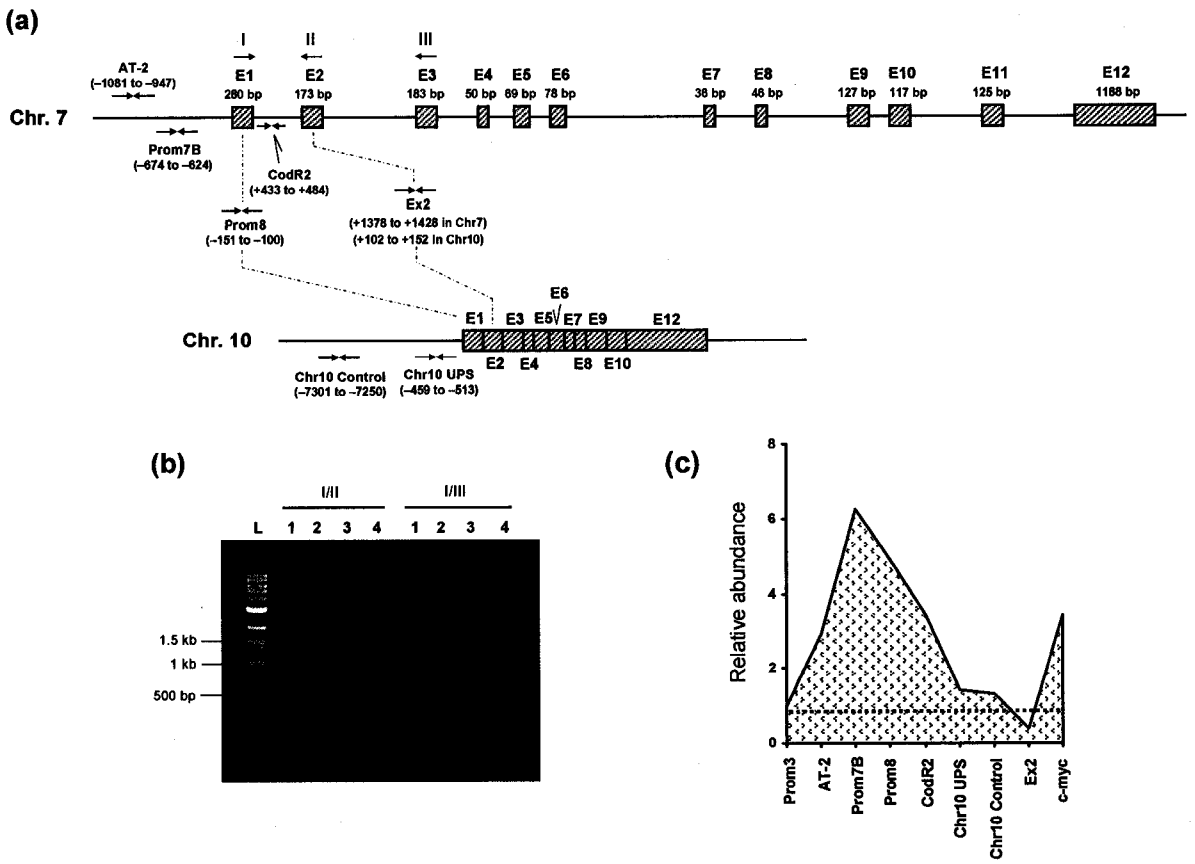
To address this issue, the presence of the *DBF4* "pseudogene" in the genome of HeLa, HEK293 and 293T cells was first verified by regular PCR with total genomic DNA as template and primers specific for *DBF4* exons 1, 2 and 3 (primer positions are shown in Fig. 8a and their sequences are given in Appendix II). Two amplification products were obtained in all cell lines when a combination of primers for exon 1 (forward) and exon 2 (reverse) was used in the PCR reaction (Fig. 8b, I/II). The longer product, of ~1.6 kb, corresponds to DNA amplified from the *DBF4* gene on chromosome 7. The short product, on the other hand, was from amplification of the chromosome 10 "pseudogene" in which the intron 1 has been deleted. Similarly, when primers for exon 1 and exon 3 were examined, a single product of ~500 bp was amplified from chromosome 10 (Figure 8b, I/III). Due to the size of the two intron segments, no product from chromosome 7 was expected in these reactions,

Figure 8. The *DBF4* pseudogene on chromosome 10 does not contain an ori

(a) Map of the human *DBF4* gene on chromosome 7 (Chr. 7) and the *DBF4* "pseudogene" located on chromosome 10 (Chr. 10). Hatched boxes represent exons (E1 to E12), with the size of each exon (in bp) also indicated (based on Accession #AF160249). The *DBF4* pseudogene includes all the *DBF4* exons, except exon 11, but no intron sequences. Convergent arrows represent the primer pairs used for Q-PCR-based nascent strand abundance analysis (primer sequences are shown in Appendix I). Red arrows are the primers used in regular PCR reactions (primer sequences are given in Appendix II).

(b) Presence of the *DBF4* pseudogene in the genome of HeLa, HEK293 and 293T cells was confirmed by PCR. Reactions were performed using the indicated primers [I to III; positions shown in (a)] and total genomic DNA isolated from each cell line. Reaction products were separated on an agarose gel and visualized with ethidium bromide. Lane 1, no template control; lane 2, HeLa DNA; lane 3, HEK293 DNA; lane 4, 293T DNA. L, DNA size marker (sizes shown to the left of the gel).

(c) The *DBF4* pseudogene locus does not contain an ori. Nascent strand abundance assays were performed using 1–2 kb nascent DNA from asynchronous HeLa cells as described in the legend to Figure 6b. Additional primer pairs were included in this experiment in order to study the *DBF4* pseudogene region (primer details are shown in Appendix I). Relative abundance was calculated using the *DBF4* gene Prom3 sequence as control.



since it would be too long (~8.5 kb) to be efficiently amplified under the PCR conditions used. Altogether, the results of these experiments corroborate the existence of a *DBF4* pseudogene in the human genome.

In order to determine if an ori is located in the vicinity of the *DBF4* "pseudogene", nascent strand abundance assays were performed using asynchronous HeLa cells, but this time including primers specific for the pseudogene on chromosome 10 as well as for exon 2 (on both chromosome 7 and chromosome 10). As expected, the Prom7B, Prom8 and c-myc DNA segments were enriched in the nascent DNA sample (Fig. 8c). However, no enrichment was observed with any of the primers designed to detect replication activity at the chromosome 10 location (i.e., Chr10 UPS and Ex2). This result is consistent with the notion that an active ori is present near the promoter of the *DBF4* gene in chromosome 7, but not in the *DBF4* "pseudogene" on chromosome 10.

3.1.2 Chromatin immunoprecipitation (ChIP) assays

Even though the nascent strand abundance assays described above clearly suggested the existence of an ori in the promoter region of the *DBF4* gene, it was still necessary to corroborate this result using a different technique. Additional methods to identify oris include leading and lagging strand polarity assays (Handeli et al., 1989; Burhans et al., 1990; Burhans et al, 1991), as well as two-dimensional gel electrophoresis (Huberman et al., 1987; Brewer and Fangman, 1987). However, while these approaches have been useful in the identification of oris in amplified genomic sequences, their use in the analysis of single-copy loci seems to be much more limited, mainly due to the low sensitivity associated with membrane-based hybridization techniques.

Although somewhat indirect, another strategy that can be used to establish if a particular DNA segment corresponds to an ori, is to determine if it is preferentially bound by replication proteins such as ORC and Cdc6, either *in vitro* or *in vivo.* Thus, chromatin immunoprecipitation (ChIP) assays were performed to examine if the human ORC complex was specifically bound to the DBF4 promoter region *in vivo*, as would be expected if the area contains an ori.

Cell extracts were prepared from formaldehyde-treated asynchronous HeLa cells and were sonicated to obtain chromatin fragments with an average size of ~300 bp. Immunoprecipitation was performed using antibodies against transcription factor Sp1 as a positive control, and against the subunit 4 of the human ORC complex (ORC4). Immunoprecipitated DNA was analyzed by Q-PCR using the primer pairs shown in Figure 9a and the resulting values were normalized to: (i) the amount of DNA recovered in control reactions using pre-immune antibodies ("normal" IgG), (ii) the amounts detected with each primer pair in a sample of total DNA before immunoprecipitation ("Input" DNA), and (iii) the normalized values obtained for a non-ori DNA sequence such as the segment amplified by the Prom3 primer pair.

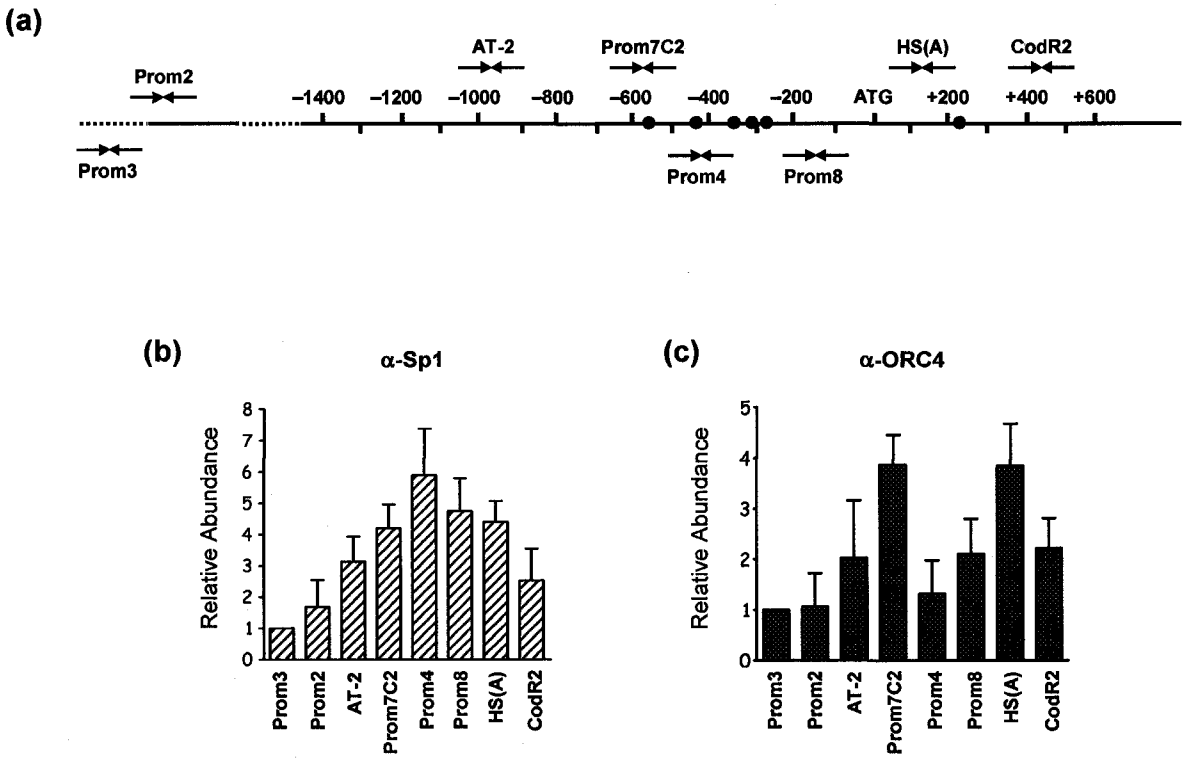
As shown in Figure 9b, promoter-proximal sequences, in particular the segment amplified by primer pair Prom4, were nearly 6-times more abundant than Prom3 DNA in anti-Sp1 immunoprecipitates. This suggests that at least some of the six putative Sp1-binding sites found in the *DBF4* promoter area (Fig. 9a) are actually occupied by this protein *in vivo*, in agreement with previous *in vitro* observations (Wu and Lee, 2002).

Figure 9. The *DBF4* ori locus contains two ORC4 binding sites *in vivo*

(a) Map of the human *DBF4* promoter region showing the position of the primers used for Q-PCR-based ChIP assays (primer sequences are shown in Appendix I). Red circles correspond to putative binding sites for the Sp1 transcription factor.

(b) ChIP assays were performed with anti-Sp1 antibodies (α-Sp1) and extracts from cross-linked, asynchronous HeLa cells. The abundance of the indicated DNA segments in the immunoprecipitated DNA was determined by Q-PCR. Resultant values were standardized by subtracting the amount recovered with pre-immune IgG. Relative abundance was calculated as % of Input DNA and normalized to the value obtained with the Prom3 primer set. Error bars, standard error of the mean of at least three experiments.

(c) ChIP assays were performed as described in (b) but with anti-ORC4 antibodies (α-ORC4) for immunoprecipitation.



Interestingly, when anti-ORC4 antibodies were used for immunoprecipitation, two separate regions of the *DBF4* locus were preferentially recovered (Fig. 9c). In this case, DNA segments corresponding to Prom7C2 and HS(A) primer sets were the most abundant in the immunoprecipitates, with the abundance of flanking regions gradually decreasing in both directions from the two peaks. This result confirms the presence of an ori in the *DBF4* locus, and suggests that it contains two distinct ORC-binding sites: one, around positions -608 to -529 (primer set Prom7C2), and the other at +131 to +182 [primer pair HS(A)].

3.1.3 *In vitro* DNA binding assays

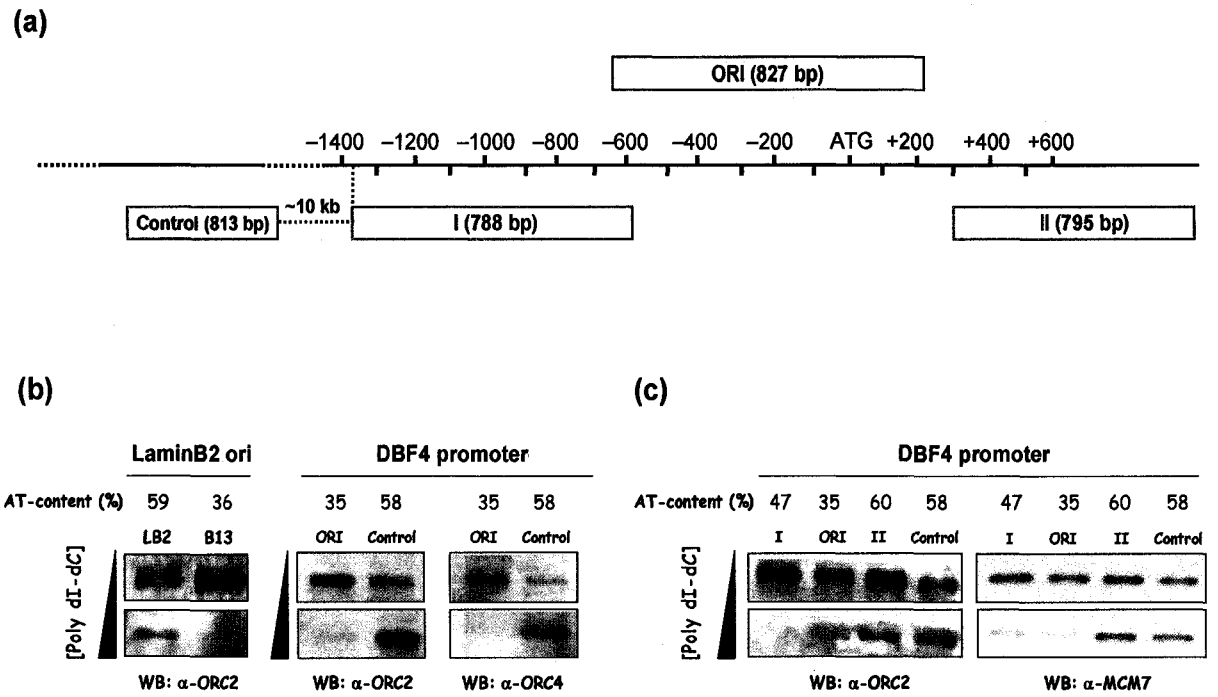
It has been previously shown that replication proteins such as ORC2 preferentially bind to oris *in vitro* (Stefanovic et al., 2003). However, other studies appear to suggest that these interactions are neither ori- nor sequence-specific and might depend instead on additional features such as AT-content (Kong et al., 2003). To analyze ORC binding to the DBF4 promoter *in vitro,* a biotinylated double-stranded DNA segment spanning the ori region (Fig. 10a, ORI) as well as a negative control fragment mapping ~10 kb upstream (Fig. 10a, Control) were generated by PCR using labeled oligonucleotides (primer details are shown in Appendix III). Purified PCR products were separately bound to streptavidin-coated magnetic beads and incubated with nuclear extracts from asynchronous HeLa cells in the presence of increasing concentrations of poly(dI-dC)-poly(dI-dC) non-specific competitor DNA. Following a series of washes and a final elution step, DNA-bound proteins in the beads were detected by Western blotting with antibodies against human ORC2, ORC4 and MCM7. To serve as control, similar experiments were performed using a biotinylated

Figure 10. *In vitro* binding of replication proteins to the DBF4 ori

(a) Biotinylated PCR products used for *in vitro* DNA binding assays. The size of each segment is indicated in parenthesis.

(b) Equimolar amounts of biotinylated DNA corresponding to the DBF4 ori region ("ORI") and to a control segment located ~9 kb upstream of the ori ("Control") were separately bound to streptavidin-coated magnetic beads and incubated with nuclear extracts from asynchronous HeLa cells in the presence of low and high concentrations of poly dI-dC non-specific competitor DNA. Beads were washed and DNA-bound proteins were detected by Western blotting with the indicated antibodies. As positive control, similar experiments were performed using a fragment from the human laminB2 ori (LB2) and a non-ori segment located ~5 kb downstream of LB2 (B13). The percentage AT content of each fragment is shown at the top of each panel.

(c) Experiments were performed as described in (b) but including two additional fragments from the DBF4 promoter region (I and II).



fragment from the human lamin B2 ori (LB2) and a segment that corresponds to a non-ori region located ~5 kb downstream of LB2 (B13).

As shown in Figure 10b, at low concentrations of competitor DNA, ORC2 bound to all sequences examined, including the B13 negative control. On the other hand, when the concentration of non-specific DNA was increased, protein binding was only detected in some of the fragments. For the lamin B2 ori, high concentrations of competitor DNA suppressed ORC2 binding to the B13 region but still allowed interaction with LB2. This suggests that ORC2 has a higher affinity for the LB2 sequence than for B13, as predicted since the LB2 segment coincides with an ori. The results obtained for the DBF4 ori, however, were completely opposite to what was anticipated. In fact, high concentrations of competitor DNA eliminated both ORC2 and ORC4 binding to the ORI fragment without affecting protein recovery with the negative control segment (Fig. 10b). Similar results were obtained when the experiments were repeated including two additional fragments from the *DBF4* promoter region (Fig. 10c).

Since it had been previously suggested that ***in vitro*** interactions between metazoan ORC and DNA might be influenced by the AT-content and the presence of asymmetric AT-rich tracts in the test sequences (Kong et al., 2003), the percentage of A+T in the fragments used in the experiments of Figure 10 was calculated. Interestingly, those sequences for which ORC proteins showed the highest affinity (i.e., LB2, Control and II) were also the ones that had the highest AT-content. For instance, in the case of the lamin B2 ori, the LB2 fragment had a much higher AT-content than the B13 control (59% versus 36%, respectively), and also displayed a higher affinity for ORC2, as described above. However, in the case of the DBF4 ori, the fragment spanning the ori region (ORI), which only showed

low-affinity ORC-binding, had a lower AT-content (35%) than the other segments, especially compared to the Control and II sequences. Considering that the ChIP assays shown in Figure 9 clearly indicate that ORC is bound *in vivo* to the DBF4 ori but not to distal sequences, these results support the increasingly accepted view that *in vitro* assays are not a reliable methodology to identify and characterize protein-DNA interactions taking place at oris, especially considering the critical role played by chromatin structure in origin function (Vashee et al., 2003).

3.1.4 Timing of activation of the DBF4 ori

Eukaryotic oris are activated following a specific temporal program, with some oris firing earlier than others during S phase progression (Goren and Cedar, 2003; Yompakdee and Huberman, 2004). In order to determine the approximate S phase period during which the DBF4 ori is activated, nascent strand abundance assays were performed using **synchronized** HeLa cells. First, cells were arrested in mitosis by nocodazole treatment followed by incubation in drug-free medium for different times. DNA content analysis by propidium iodide staining and flow cytometry indicated that most cells exited mitosis and entered the following G1 phase between 2 h and 3 h after release from the nocodazole block (Fig. 11a). Cells in early S phase were detected at 7 h post-release, whereas, at later times, a sub-population of cells transiting S phase was observed (Fig. 11a, 10 h). Under these conditions, the highest enrichment values for the DBF4 ori were obtained at the 7 h timepoint, suggesting that the ori is preferentially activated during early S phase (Fig. 11b).

Due to the relatively low synchronization efficiency obtained after nocodazole release (Fig. 11a), a second synchronization method was employed to confirm the activation

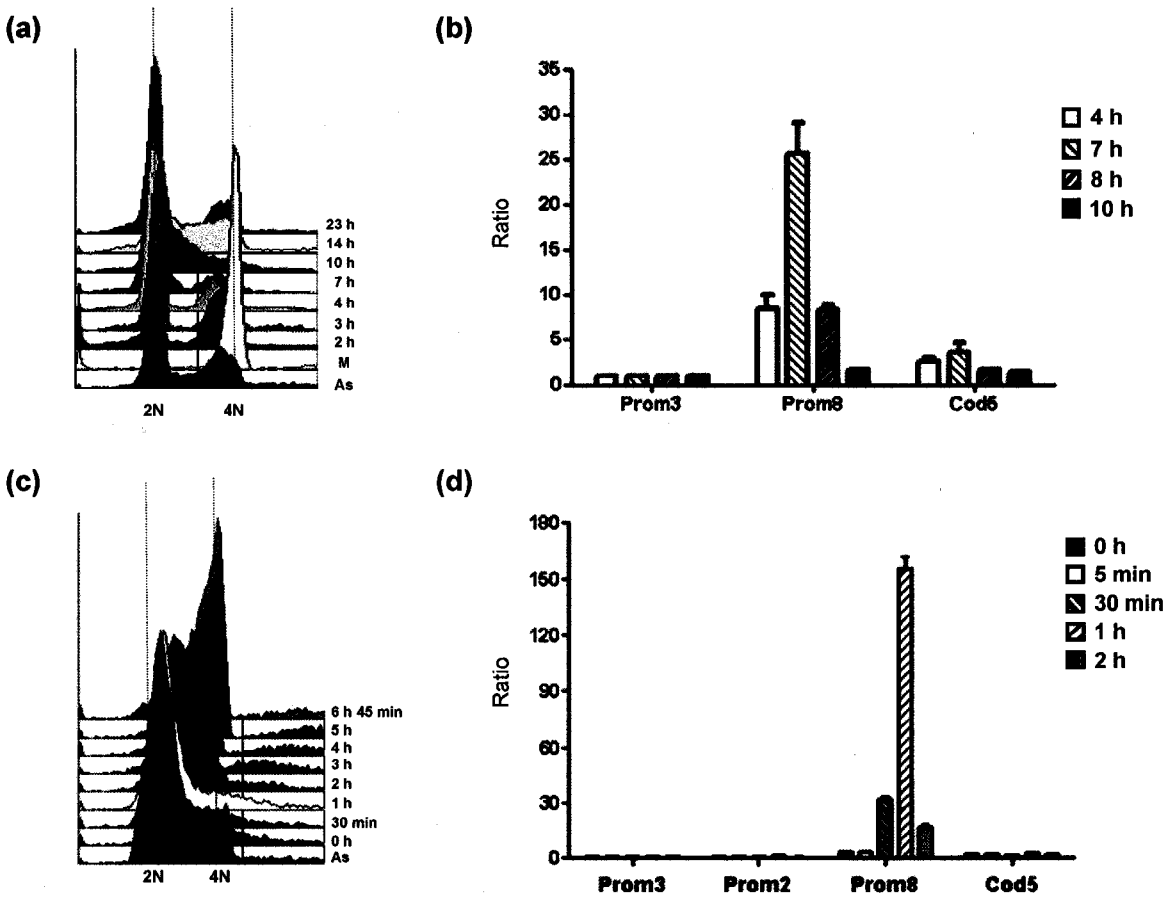
Figure 11. The DBF4 ori is preferentially activated during early S phase

(a) Flow cytometry profile of post-mitotic HeLa cells. As, asynchronous cells; M, cells arrested at mitosis by nocodazole (50 ng/ml for 18 h). Numbers to the right of the panel denote times (in hours) following release from the mitotic arrest.

(b) Short nascent DNA corresponding to Prom8 is most abundant at 7 h post-mitosis. Nascent strands (1–2 kb) were isolated at the indicated timepoints after HeLa cells were released from a nocodazole-based mitotic arrest. The abundance of DNA segments corresponding to each primer set was measured by Q-PCR and the resultant value was compared to that generated by the Prom3 primer pair (set as 1).

(c) Flow cytometry profile of post-G1/S HeLa cells. 0 h, cells arrested by double-thymidine treatment. Numbers to the right of the panel denote times (in hours) following release from the second thymidine block; As, asynchronous cells.

(d) Replication at the human DBF4 ori (corresponding to Prom8) is most active at 1 h post-G1/S. Nascent strand abundance assays were performed as described in (b) using short, single-stranded DNAs isolated at the indicated timepoints following the release from the double-thymidine block.



profile of the DBF4 ori. In this case, HeLa cells were arrested in G1/S to early S phase using a double-thymidine block, and then released into regular medium for different times. Flow cytometric analysis of DNA content demonstrated that almost 100% synchronization efficiency could be achieved with this approach (Fig. 11c). High levels of DBF4 ori activity were observed as early as 30 min after thymidine removal, reached a peak by the 1 h timepoint, and declined by 2 h post-thymidine release (Fig. 11d). Thus, the results from the two synchronization experiments collectively indicate that the DBF4 ori is preferentially activated during early S phase, similar to what has been reported for the lamin B2 and c-myc human oris (Biamonti et al., 1992; Waltz et al., 1996).

3.2 RIP mapping of the DBF4 ori

Several 5'-end-mapping procedures have been employed to identify the precise transition points (TP) between discontinuous (lagging) and continuous (leading) DNA synthesis in a number of viral and yeast oris (Hay and DePamphilis, 1982; Hendrickson et al., 1987; Bielinsky and Gerbi, 1999; Gómez and Antequera, 1999). These TPs correspond to the exact nucleotide position from which DNA synthesis is initiated, and thus serve to define an OBR. None of these techniques, however, were thought to be sensitive enough to allow high-resolution ori mapping in metazoan cells, mainly due to the enormous complexity of their genome (Abdurashidova et al., 2000). Consequently, most metazoan oris identified to date have only been mapped with low resolution methods such as the nascent strand abundance assay. Two notable exceptions to this are: (i) the II/9A ori of *S. coprophila* which, thanks to the fact that it is highly amplified in salivary gland cells, was mapped with single-nucleotide resolution using a common primer extension-based method (Bielinsky et

al., 2001), and (ii) the single-copy human lamin B2 ori, whose TP was determined using a cumbersome but highly sensitive ligation-mediated PCR (LM-PCR) procedure (Abdurashidova et al., 2000).

Since the DBF4 ori displayed very high activity levels compared to other human oris (Fig. 6b), it was hypothesized that a standard primer extension-based ori mapping procedure could be sensitive enough to determine precise replication start-sites at this locus without the need for more intricate methodologies such as LM-PCR.

To test this hypothesis, a modified version of the RIP mapping method described by Gerbi and Bielinsky was used (Gerbi and Bielinsky, 1997). Briefly, short strands of nascent DNA (isolated by alkaline gel electrophoresis as described in section 3.1) were utilized as template in one-way PCR amplification reactions using single, strand-specific primers, labeled on their 5'-ends with digoxigenin. Following sequencing gel electrophoresis, the position of the 5'-ends of the nascent strands was determined based on the location of the primer used and the size of the generated products. Theoretically, the shortest amplification product observed with a given primer marks the position of the RIP for the corresponding leading strand, with higher molecular weight products resulting from the successive ligation of Okazaki fragments (Gerbi and Bielinsky, 1997). One major drawback of this procedure is that it is not possible to determine if multiple potential start-sites for leading strand synthesis exist within an ori, since initiation sites located upstream of the shortest RIP will always be masked by the presence of Okazaki fragments. To circumvent this potential problem, cells may be treated with the drug **emetine** prior to the isolation of nascent DNA. This protein synthesis inhibitor has been shown to prevent the synthesis and ligation of Okazaki fragments, without significantly affecting leading strand progression (Burhans et al., 1991).

In order to validate the modified RIP mapping procedure, one-way PCR-based primer extension was used to map RIPs at the lamin B2 ori, the only human ori with known RIP positions. Since lamin B2 ori activity seemed to be quite low in exponentially growing HeLa cells (Fig. 6b), it was unlikely that lamin B2 RIP mapping by the one-way PCR approach could be achieved with these cells. Therefore, nascent strands were isolated from emetine-treated HeLa cells that were released for 4 h from mitotic arrest. Replication activity at the *lamin B2* locus reached the peak at this timepoint (Figure 12b). A band of the expected size (273 nucleotides; Abdurashidova et al., 2000) was obtained when nascent DNA from synchronized cells was amplified using a primer complementary to the "bottom" leading strands emanating from the lamin B2 ori (Figure 12c; primer position is indicated in Figure 12a and the primer sequence is shown in Appendix IV). Therefore, although the signal obtained was rather weak, this result confirmed the suitability of the RIP mapping methodology.

The one-way PCR method was then applied to determine RIPs at the DBF4 ori. Four digoxigenin-labeled primers specific to the *DBF4* locus (pA to pD) were used to amplify 1–2 kb nascent DNA from **emetine-treated asynchronous** HeLa cells (primer details and sequences are given in Appendix IV). Five different amplification products were obtained when "sense" (i.e. transcription direction) leading strands were amplified using primer A (Fig. 13a, pA, positions -106 to -125). Since these experiments were performed using cells pre-treated with emetine, this result suggested that DNA synthesis from the "bottom" parental DNA strand may initiate from several potential sites.

Based on the sizes of the observed bands, approximate replication start-sites were located at positions -390, -400, -485, -510 and -565 (Fig. 13c). No primer extension

Figure 12. RIP mapping at the human lamin B2 ori by one-way PCR

(a) Map of the human *lamin B2* locus. Numbering is from GenBank file M94363. Gray block arrows show the *lamin B2* and *TIMM13* genes (arrow directions indicate the direction of transcription). Convergent, horizontal arrows show the position of the primer pair used in Q-PCR to measure lamin B2 ori activity by nascent strand abundance analysis [see panel (b)]. Black vertical arrows at positions 3930 and 3933 correspond to the leading strand initiation sites determined by Abdurashidova et al (2000). The small gray arrow starting near position 4200 represents the digoxigenin-labeled primer used for RIP mapping (primer sequence is shown in Appendix IV and is the same as reported by Dimitrova et al., 1996). The size of the expected RIP product is 273 nucleotides (Abdurashidova et al., 2000).

(b) Replication activity at the *lamin B2* locus is maximal at 4 h post-mitosis. Short nascent DNA templates (1–2 kb) were isolated from HeLa cells at different timepoints following release from mitotic arrest. Analysis of DNA replication by nascent strand abundance assays and Q-PCR was carried out using primer pair lamin B2 (primer position shown in (a) and Appendix I). DNA abundance was compared to that of the Prom3 negative control primer set. As, asynchronous cells; 0, cells arrested at mitosis by nocodazole; 4 -10, 4 -10 h post-mitosis.

(c) One-way PCR-based RIP mapping was carried out using primer C1 to determine the leading strand RIP on the "bottom" strand of the lamin B2 ori. Nascent DNA (1–2 kb) was isolated from emetine-treated HeLa cells released for 4 h from nocodazole block. L, digoxigenin-labeled molecular weight ladder (sizes in bases are shown to the left of the panel); lane 1, no template control; lane 2, nascent DNA from cells 4 h post-mitosis; lane 3, nascent DNA from nocodazole-arrested cells (0 h). A band of the expected size (~273 bases) was observed only in the 4 h sample (lane 2).

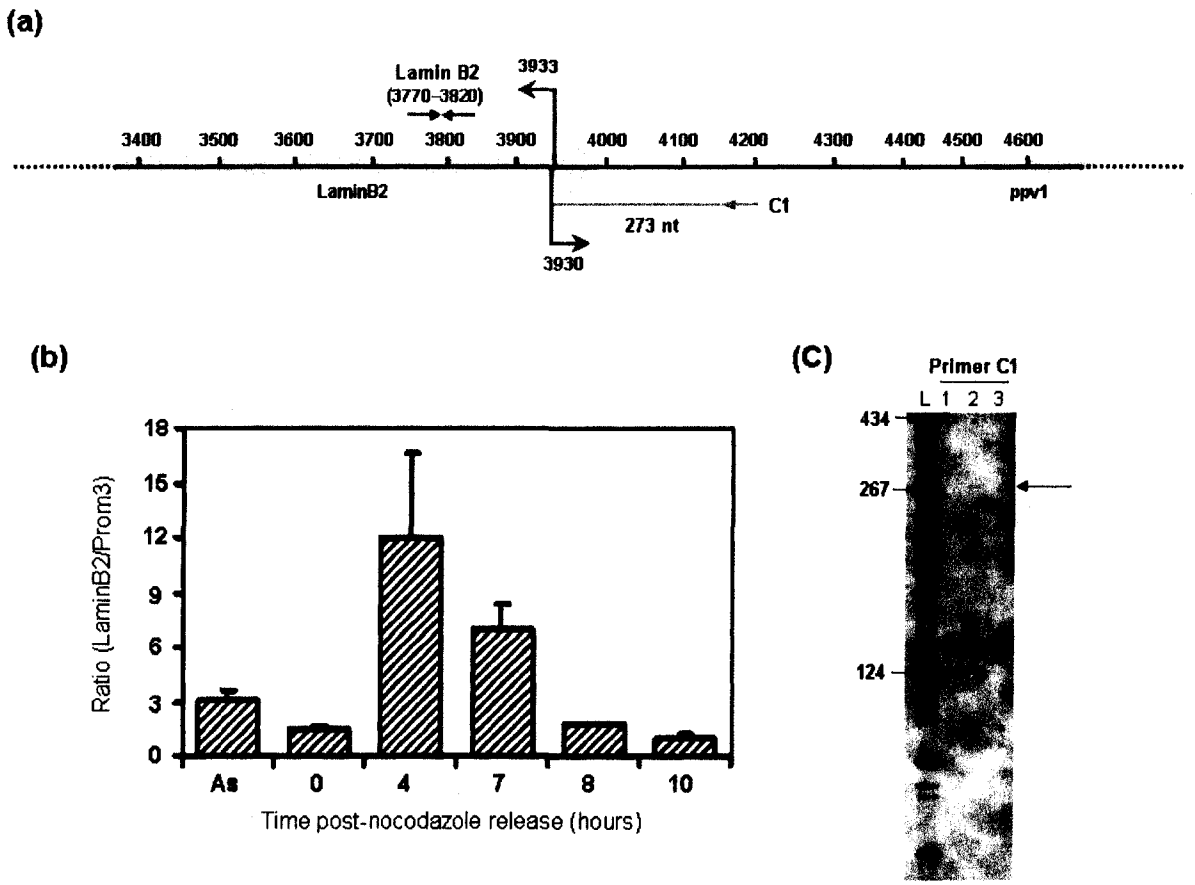
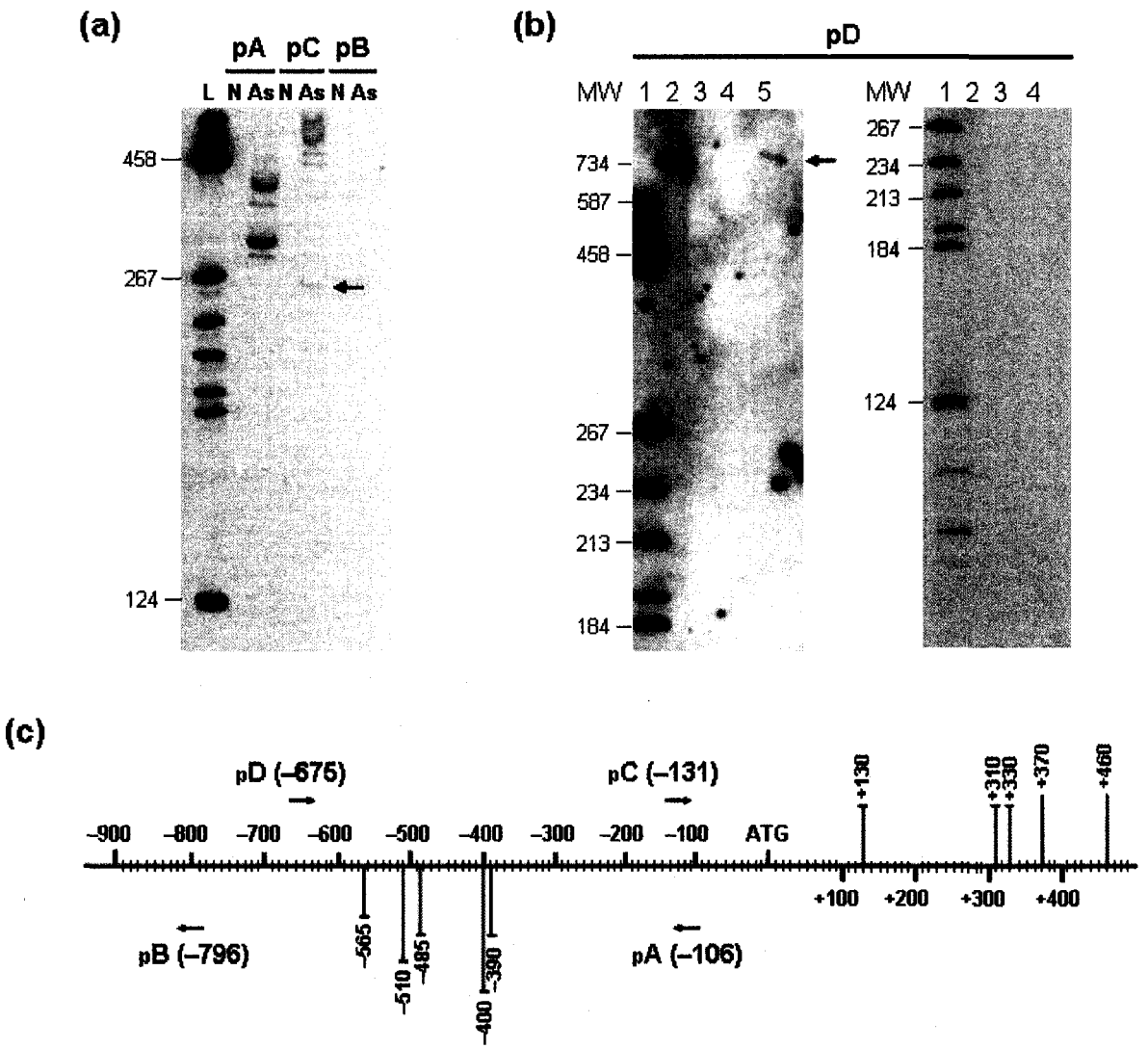


Figure 13. RIPs at the human DBF4 ori in asynchronous HeLa cells

(a) Primer extension was carried out using nascent DNA (1–2 kb) from asynchronously growing HeLa cells that had been pre-incubated for 1 h with 2 μM emetine (As). Primers used (pA, pC or pB) were 5'-end-labeled with digoxigenin (primer positions and sequences are shown in panel (c) and in Appendix IV). L, digoxigenin-labeled molecular weight ladder (sizes in bases are shown to the left); N, control reactions without template DNA. The approximate position of RIPs from the bottom (pA) and top (pC) DNA strands was estimated from the sizes of the obtained PCR products and is shown by the vertical lines in panel (c). The arrow points to the same RIP identified by pC and by pD in panel (b).

(b) *(Left panel)* The same nascent strand templates used in (a) were subjected to primer extension with primer pD (see panel (c) for primer position and Appendix IV for primer sequence). Reaction products were separated by gel electrophoresis to resolve high molecular weight fragments. Lane 1, molecular weight ladder (sizes shown to the left of the gel); Lane 2, 5'-digoxigenin-labeled PCR product included as an additional size marker (738 bases); Lane 3, no sample; Lane 4, sham reaction control including primer pD but no template DNA; Lane 5, primer extension using 1–2 kb nascent DNA templates from asynchronous HeLa cells. The arrow marks the same product amplified by primer pC in panel (a). *(Right panel)* The same as in the left panel, but electrophoresis was focused on resolving smaller DNA fragments. Lane 1, molecular weight ladder; lane 2, no sample; lane 3, sham reaction control; lane 4, the same PCR reaction as in lane 5 of the left panel.

(c) The human DBF4 ori does not follow the conventional OBR model. Summary of RIP sites identified at the DBF4 ori in asynchronous HeLa cells (vertical lines). Numbers indicate the position of each site with respect to the DBF4 translation-start codon (ATG). Labeled primers used for RIP mapping are shown as small horizontal arrows, with the number in parenthesis indicating the position of the 5'-end of each oligonucleotide.



product was obtained when primer B was used to amplify the same template DNA (pB, positions -796 to -815), indicating that DNA synthesis does not initiate frequently upstream of -815 on the sense strand (Fig. 13a, pB).

According to the generally accepted OBR model, it was expected that replication initiation sites on the antisense strand would be located very close to the positions mapped on the sense strand using pA. However, primer extension with primer D (pD, positions -675 to -657) generated a single high molecular weight fragment, which corresponded to strands initiated somewhere around position +100 on the antisense strand (Fig. 13b, pD, left panel). No shorter products were detected with this primer, even when they were actively looked for (Fig. 13b, pD, right panel). This suggested that initiation at the *DBF4* locus does not follow the conventional OBR model, since no initiation was observed on the opposite strand from RIPs detected with pA. Supporting this conclusion, amplification with primer C (pC, positions -130 to -111) showed that RIPs on the antisense strand were in fact located around positions +130, +310, +330, +370 and +460 (Fig. 13a, pC), of which the RIP at +130 most likely corresponds to the band observed with pD. Other RIPs detected with pC were not detected using pD, presumably because one-way PCR is not optimal to efficiently amplify such long templates. Finally, and similar to what was observed with pA, the multiple bands obtained with pC indicated that DNA synthesis from the antisense strand can also start from multiple potential sites.

RIP mapping at the DBF4 ori was next performed using **synchronized** HeLa cells 1 h after release from a double-thymidine block. As described in section 3.1, activity of the DBF4 ori was notably high at this timepoint, with promoter-proximal sequences being enriched up to ~150-fold over ori-distal DNA segments (Fig. 11d). In agreement with the

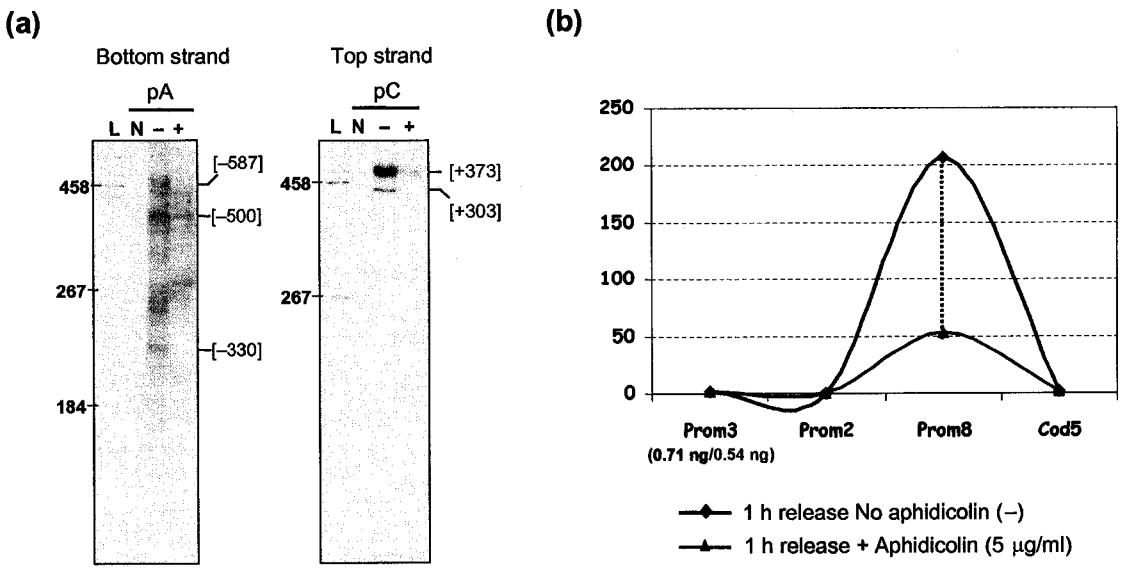
results obtained with the asynchronous cell population, multiple initiation sites were detected when nascent strands from emetine-treated, synchronized HeLa cells were used as template in primer extension reactions with pA. The major bands were distributed within a region spanning positions from -330 to -610 on the sense strand (Fig. 14a, pA, lane "-"). In the case of antisense leading strands, amplification with primer pC generated two major products at approximately +303 and +373 (Fig. 14a, pC, lane "-"). Although the RIPs determined in synchronized HeLa cells did not match exactly those observed in the asynchronous population, DNA synthesis at the DBF4 ori consistently initiated from any of several potential sites clustered within two "replication initiation zones" which are approximately 400-bp apart from each other.

In order to prove that the products observed in the RIP mapping experiments were the result of amplification of nascent strands and not contaminating randomly nicked DNA, primer extension reactions were performed using double-thymidine arrested cells that were released for 1 h in the absence or presence of the DNA replication inhibitor aphidicolin (and emetine). As shown by nascent strand abundance analysis, a DBF4 ori segment (Prom8) was almost 200-times more abundant than non-ori fragments (Prom3, Prom2 and Cod5) in the absence of aphidicolin (Fig. 14b, blue line). This value is in agreement with those shown in Figure 11d. As expected, a considerable reduction in the relative abundance of ori-proximal versus ori-distal sequences was observed in aphidicolin-treated cells, demonstrating the inhibitory effect of aphidicolin on DNA synthesis (Fig. 14b, red line). This decrease was not due to major differences in the amount of "background" or non-ori DNA, since the estimated amounts for the Prom3 segment in both preparations were only marginally

Figure 14. RIP mapping of the DBF4 ori in synchronized HeLa cells

(a) RIP primer extension reactions were carried out using primers pA or pC and 1–2 kb nascent DNA from HeLa cells released for 1 h from a double-thymidine arrest. Emetine was added to the growth media for one hour prior to sampling in the presence (+) or absence (–) of 5 μg/ml aphidicolin. L, digoxigenin-labeled molecular weight ladder (size in bases shown to the left); N, no template control reactions. Numbers in brackets to the right of each panel indicate the approximate position of RIPs.

(b) Q-PCR analysis of nascent strand abundance using the same DNA preparations as in panel (a). Equal volumes of each DNA sample (+/ – aphidicolin) were used as template for Q-PCR amplification with the indicated primers. Results are shown as the ratio between the amount of each segment and the amount of the control fragment amplified by Prom3 primer set. For reference, the raw amount of Prom3 DNA in each sample, as determined by Q-PCR, is also shown (in ng).



different (Fig. 14b, 0.71 ng in the sample without aphidicolin versus 0.54 ng in the aphidicolin-treated preparation).

Importantly, when **equal amounts of nascent DNA** from aphidicolin-treated and untreated cells were used as template in RIP mapping reactions, diminished signals were obtained with both pA and pC in the aphidicolin-treated sample, clearly showing that the RIP products are sensitive to inhibitors of DNA replication, as would be expected if they are the result of amplification of replication intermediates (compare lanes "–" and "+" of pA and pC in Fig. 14a).

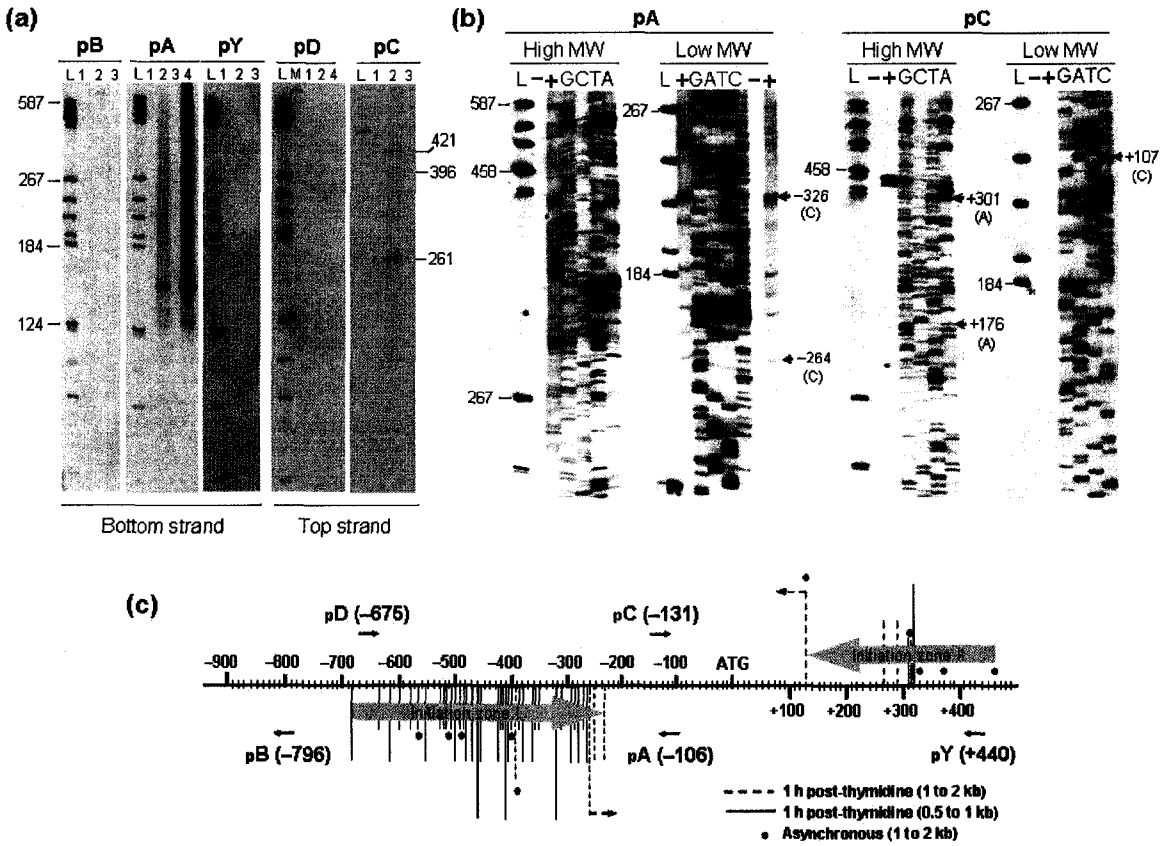
Although emetine treatment provides critical information in ori mapping, it is formally possible that it could generate artifactual results. To rule out this possibility, RIP mapping assays were performed using cells at 1 h post-G1/S in the presence or absence of emetine. When pA was used for primer extension with 1–2 kb nascent DNA from synchronized cells **without emetine treatment**, the shortest observed product (usually considered to be a TP) was detected at -235 (Fig. 15a, lane 4 of panel pA). In addition to this band, other products were also observed with this primer. This was expected because, in the absence of emetine, Okazaki fragments will ligate to the leading strands as replication progresses, generating longer DNA templates. However, multiple primer extension products were also obtained with primer pA in the **emetine-treated sample** (Fig. 15a, lane 2 of panel pA). This is consistent with the idea that DNA synthesis on the sense strand starts from multiple potential sites. Based on the sizes of the strongest bands, the most frequently used initiation sites were located at positions -235, -255, -260 and -390 (Fig. 15c, broken lines). As was observed previously, no RIP was detected in any of the samples when using primer pB (Fig. 15a, panel pB). Similarly, no amplification was observed with primer Y (pY,

Figure 15. The DBF4 ori contains two initiation zones in the presence or absence of emetine

(a) Leading strand RIPs were determined by primer extension using 1–2 kb nascent DNA templates. Primers used (pB, pA, pY, pD and pC; Appendix IV) were 5'-end labeled with digoxigenin. Nascent DNA was isolated from HeLa cells arrested by double-thymidine and released for 1 h into medium containing emetine (lane 2). Control PCR reactions were performed in the absence of template DNA (lane 1) or using DNA from emetine-treated, **mitotic** cells (lane 3). Two samples were not treated with emetine (lanes 4 in panels pA and pD). L, digoxigenin-labeled molecular weight ladder. M, digoxigenin-labeled 738-base PCR product used as an additional size marker.

(b) RIP assays were performed with 0.5–1.0 kb nascent DNA from emetine-treated cells 1 h after thymidine arrest (+). Control reactions without template DNA were also included (–). The positions of some of the RIPs identified are marked by black arrows. Top and bottom portions of sequencing gels were separated in order to better resolve high molecular weight products. Sequencing ladders are shown (G, A, T, C).

(c) Summary of replication initiation sites. Vertical dashed lines show leading strand initiation sites for both the sense (bottom) and antisense (top) strands as determined in (a), using 1–2 kb nascent DNA template from1 h post-thymidine cells treated with emetine. Solid lines show the position of RIPs determined in (b) using 0.5–1.0 kb nascent DNA from synchronized cells. Longer lines indicate RIPs with stronger signals. Replication direction from initiation zone I (sense strand) and zone II (antisense strand) is shown by gray arrows. The RIPs of asynchronous HeLa cells (from Fig. 13) are marked by black dots.



positions +421 to +440), suggesting that initiation does not occur upstream of position -815 or downstream of -235 on the sense strand (Fig. 15a, panels pB and PY).

Amplification with primer pC confirmed that replication initiation on the antisense strand starts around positions +130, +265 and +290 (Fig. 15a, lane 2 of panel pC, and Fig. 15c, broken lines), similar to what was observed in asynchronous cells. In the case of primer pD, however, no extension products were obtained in synchronized cells (Fig. 15a, panel pD), even though a faint ~800-bases band was detected with the same primer using asynchronous cells (Fig. 13b). This is probably because 800 bases may be the upper limit that can be efficiently amplified by the one-way PCR-based primer extension. Nonetheless, no low molecular weight products were observed with pD in any of the samples, including the one that was not treated with emetine. This data suggests that, even in the absence of emetine, replication from the DBF4 ori antisense strand occurs at least 350 bases downstream of the initiation sites on the opposite strand, consistent with the idea that the two initiation zones identified at the *DBF4* locus are not an artifact of emetine treatment.

It is worth mentioning that no amplification was observed with any of the primers when emetine-treated **mitotic cells** were used as the source of nascent DNA (Fig. 15a, lane 3 in all panels). This corroborates that the RIP products generated by one-way PCR did not result from non-specific amplification of randomly cleaved genomic DNA.

Lastly, DBF4 ori RIPs were determined using shorter DNA templates (0.5–1.0 kb) isolated from emetine-treated cells synchronized at 1 h post-G1/S. In order to map the precise nucleotide position of each initiation site, one-way PCR products were separated by electrophoresis along with DNA sequencing reactions (Fig. 15b, lanes G, A, T and C in all panels). Forty nine RIPs were detected on the sense strand with pA, all of which were

clustered within a 420-bp segment that extends from -264 to -683 (Fig. 15b, panel pA, and Fig. 15c, solid lines). Among these RIPs, initiation at a C residue was the most common (20/49, 41%), followed by T (11/49, 22%), G (10/49, 20%) and A (8/49, 16%). For the antisense strand, replication start-sites were detected at +311, +312 and +320 (Fig. 15b, panel pC, and Fig. 15c, solid lines). A summary of all the initiation sites identified in this set of experiments is shown in Appendix V.

Collectively, the data presented here suggests the existence of two replication initiation zones at the DBF4 ori. Initiation zone I, on the sense strand, spans ~449 nucleotides (from -235 to -683) and includes at least 49 potential RIPs, some of which are used more frequently than others. Initiation zone II, on the other hand, occupies ~331 bases on the antisense strand (from +130 to +460) and contains only a few RIPs. The two zones are separated by a 365-bp spacer (from -235 to +130) in which replication initiation either does not occur or occurs very rarely.

3.3 Asymmetric initiation of bidirectional replication at the *DBF4* ori locus

It is generally assumed that DNA synthesis from eukaryotic oris begins simultaneously from both parental strands and then proceeds in the form of two replication forks moving in opposite directions. However, in some viral oris (Hay and DePamphilis, 1982; Decker et al., 1986; Hendrickson et al., 1987) as well as in oriC of *E. coli* (Hirose et al., 1983; Kohara et al., 1985) and in the D-loop ori of mouse mitochondrial DNA (Brown et al., 2005), leading strand synthesis from the two parental strands appears to initiate sequentially, with one strand firing first as a prerequisite for the activation of the oppositely moving fork.

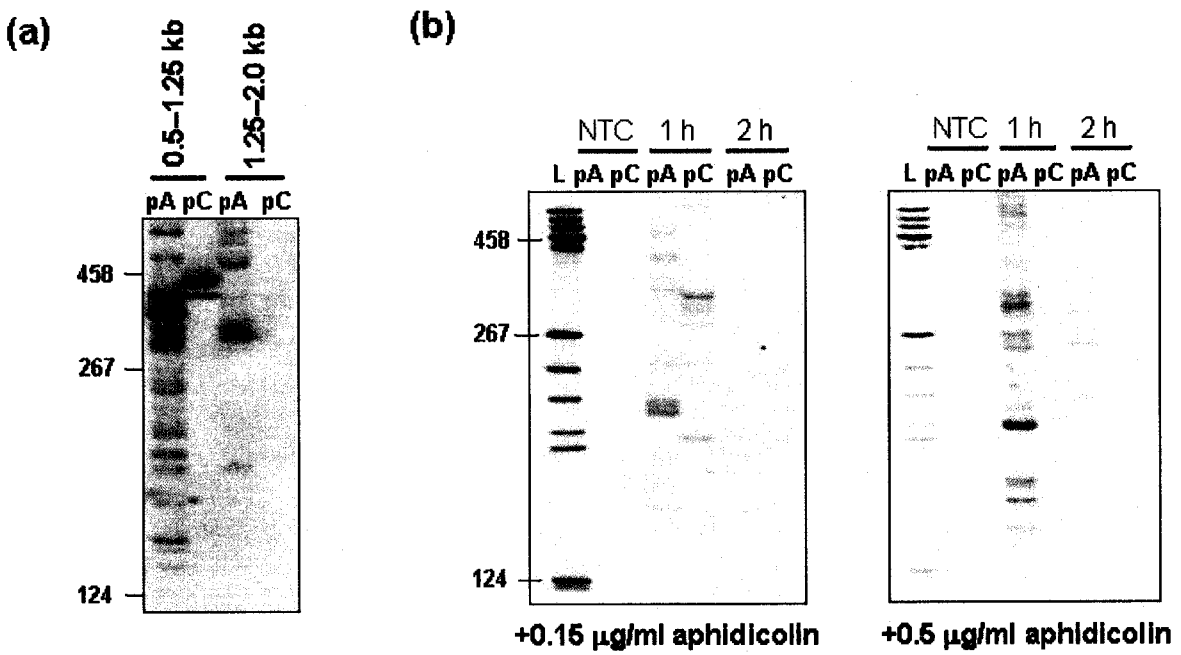
The distribution of RIPs at the DBF4 ori greatly differed from what is observed in other eukaryotic oris, including the lamin B2 ori (Abdurashidova et al., 2000). Instead, the two replication initiation zones of the DBF4 ori were reminiscent of prokaryotic oris, particularly the *E. coli* oriC locus (Kohara et al., 1985). This observation, along with the fact that replication forks emanating from the two DBF4 initiation zones must initially move towards each other as they progress through the intervening "non-initiation" region, led to the hypothesis that replication initiation from the sense and antisense strands of the DBF4 ori does not occur simultaneously. In fact, some of the initial data seemed to suggest that this could be the case. For example, as shown in Figure 14a, although DNA replication was effectively blocked by addition of aphidicolin, the drug appeared to have a more pronounced effect on the antisense strands than on the strands synthesized from the bottom parental helix (Fig. 14a, compare lanes "–" and "+" in panels pA and pC). Since, at low concentrations, aphidicolin is a more potent inhibitor of chain elongation than replication initiation (Decker et al., 1986; Dimitrova and Gilbert, 2000), one possible interpretation of this result could be that the antisense strands can only be initiated after chains moving in the sense direction have been elongated.

To further explore the possibility of asymmetric initiation, RIP assays were performed using short (0.5–1.25 kb) and long (1.25–2.0 kb) nascent strands isolated from the same population of cells synchronized at 1 h post-G1/S arrest. When the smaller nascent DNA was used as template, several RIPs were detected with both pA and pC (Fig. 16a). In contrast, only the primer pA generated any detectable products with the larger strands. Since both DNA samples were isolated from the same cell lysate, the lack of amplification with

Figure 16. Replication initiation occurs from replication zone I prior to zone II

(a) RIP analysis was carried out using short (0.5–1.25 kb) or long (1.25–2.0 kb) nascent DNA isolated from emetine-treated HeLa cells at 1 h post-G1/S arrest. Numbers to the left of the panel indicate the position of size markers; pA and pC are the digoxigenin-labeled primers used for primer extension.

(b) RIP assays were carried out with primers pA and pC using **0.5–1 kb** nascent DNA isolated from synchronized HeLa cells as described in (a), except that aphidicolin was added to the cell culture media at the indicated concentrations. "1 h" and "2 h" are the times after release from the double-thymidine block. NTC, no template controls. L, digoxigenin-labeled molecular weight ladder (sizes are indicated to the left).



pC in the longer DNA fraction seemed consistent with the hypothesis that DNA synthesis from the sense strand initiates earlier than that from the antisense strand.

Additional evidence supporting the asymmetric initiation of DNA synthesis at the *DBF4* ori locus came from RIP experiments performed with short (0.5–1.0 kb) single-stranded DNA templates isolated from cells released from thymidine arrest for 1- or 2 h in the presence of low levels of aphidicolin (and emetine). This approach has been previously used to determine the initial direction of DNA synthesis at the SV40 ori (Decker et al., 1986). Two different concentrations of aphidicolin were used. At 0.15 μg ml^{-1} aphidicolin, several RIPs were detected with both pA and pC in the 1 h sample (Fig. 16b, 1 h in left panel). No products were detected with either primer 2 h after release from the G1/S arrest, suggesting that most of the nascent DNA initiated from both strands has already progressed beyond detection (i.e., >1.0 kb) by this time. This is consistent with the data shown in Fig. 11d. When replication forks were further slowed down by increasing the concentration of aphidicolin to 0.5 μg ml^{-1}, RIPs on the sense strand were detected at 1- and 2 h post-thymidine with pA (Fig. 16b, right panel). In contrast, no appreciable replication activity was detected with pC at either timepoint (Fig. 16b, right panel). Thus, these data suggest that initiation from zone I on the sense strand occurs prior to initiation from zone II on the antisense strand.

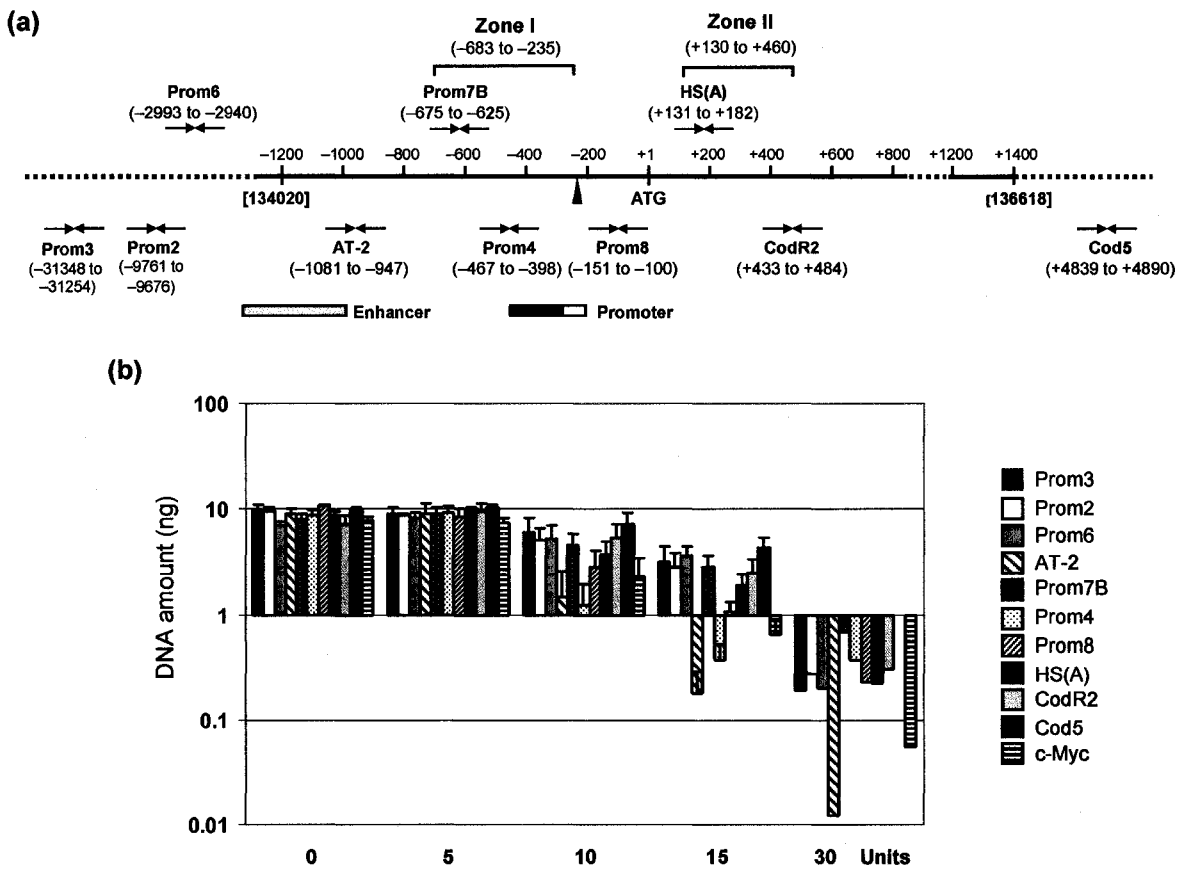
3.4 Chromatin structure at the *DBF4* ori locus

To gain insight into chromatin organization around the DBF4 ori, which could potentially be involved in the establishment of asymmetric initiation, Q-PCR-based DNase I-sensitivity assays were carried out (Fig. 17). The results of these experiments showed that

Figure 17. The *DBF4* locus contains two DNase I-sensitive regions

(a) Map of the DBF4 ori region showing the location of the core and auxiliary promoter elements (white and green boxes, respectively), along with the position of a putative enhancer (gray box) and the primers used for Q-PCR analysis of DNase I-sensitivity. The two replication initiation zones identified by RIP mapping are also indicated.

(b) To map DNase I-sensitive sites, HeLa nuclei were treated with increasing amounts of DNase I (x-axis). Eight nanograms of DNA from the digested nuclei were quantified by Q-PCR with the indicated primer pairs using standard curves generated with total, undigested genomic DNA. A nuclease-sensitive region of the human *c-myc* promoter (c-Myc) was used as positive control (Kumar and Leffak, 1991). Error bars, standard error of the mean of two or more experiments.



the DNA segment corresponding to primer AT-2 (-1081 to -947) was much more sensitive to DNase I-digestion than two flanking regions (Prom6 and Prom7B), suggesting that this area has a more "open" chromatin structure. Similar to AT-2, the segment amplified by primer pair Prom4, which maps within replication initiation zone I, was also more sensitive to DNase I-treatment. In this case, DNase I-sensitivity (i.e., chromatin openness) gradually decreased towards the region downstream of Prom4, with the segments corresponding to Prom8, HS(A) and CodR2 being progressively more resistant to nuclease treatment (Fig. 17b). Thus, the segment from -467 to -398 is generally in a more relaxed chromatin arrangement, with nucleosome packing or chromatin compaction gradually increasing downstream of this area.

3.5 Protein-DNA interactions at the *DBF4* locus

As mentioned in Chapter 1, DNA replication and gene transcription are thought to be closely interrelated cellular processes, although the exact nature of this relationship in higher eukaryotes remains unclear (Françon et al., 1999). Since the region around the *DBF4* promoter also serves as an ori, the *DBF4* locus constitutes an ideal model to study how origin activity and gene transcription are coordinated in human cells.

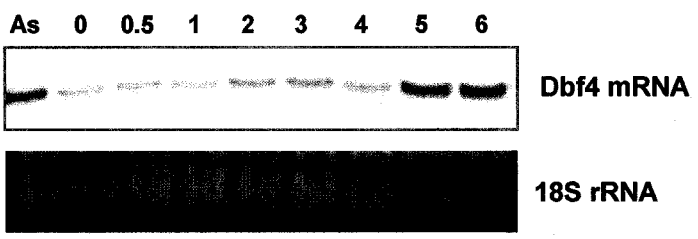
Toward this end, *DBF4* mRNA levels were measured at different timepoints post-G1/S arrest. Northern blot analysis was performed using a probe against the Chinese hamster DBF4 coding sequence (Guo and Lee, 2001). As shown in Figure 18, *DBF4* mRNA levels were high in asynchronous cells (As, set as 100% in Fig. 18b), decreased to about 25% at the G1/S border (0 h), and then remained at this level for up to 4 h following release from the G1/S block (Fig. 18a and b). A substantial increase in transcript level was

Figure 18. Changes in *DBF4* mRNA levels in different cell cycle compartments

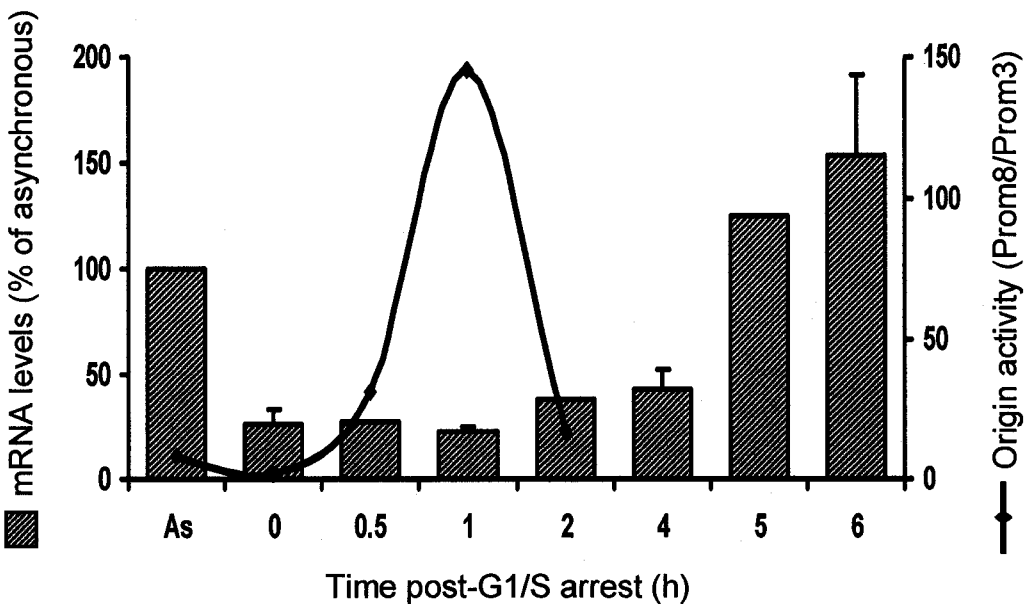
(a) Northern blot analysis of DBF4 mRNA expression along the cell cycle. HeLa cells were synchronized at the G1/S border by double-thymidine arrest and released into regular medium for the indicated period. As, asynchronous cells. Total RNA was isolated, separated by electrophoresis, transferred to a nylon membrane and hybridized with a digoxigenin-labeled probe against Chinese hamster *DBF4* mRNA. Ethidium bromide staining of 18S rRNA is shown as loading control.

(b) *DBF4* mRNA levels (hatched bars) were estimated from the blot in (a) and from additional experiments, by comparing the image optical intensity at each timepoint with the value obtained for asynchronous cells (As), which was taken as 100%. DBF4 origin activity at the different timepoints is shown by the gray line (data from Fig. 11d).

(a)



(b)



observed at later timepoints (5 h and 6 h), when, as estimated by flow cytometry, most cells were around G2/M (Fig. 11c). In contrast to this profile, DBF4 ori activity peaked at the 1 h timepoint and decreased by 2 h after release from the G1/S block (black line in Fig. 18b; data is the same as in Fig. 11d). Thus, high activities of transcription and replication initiation at the *DBF4* ori locus appeared to be mutually exclusive.

The core and auxiliary promoter elements of the human *DBF4* gene overlap with replication initiation zone I of the DBF4 ori (Fig. 17a). To study in more detail how the DNA replication and transcriptional machineries may be organized within such a small genomic area, anti-ORC4 and anti-Sp1 ChIP assays were performed using cells synchronized at different cell cycle positions.

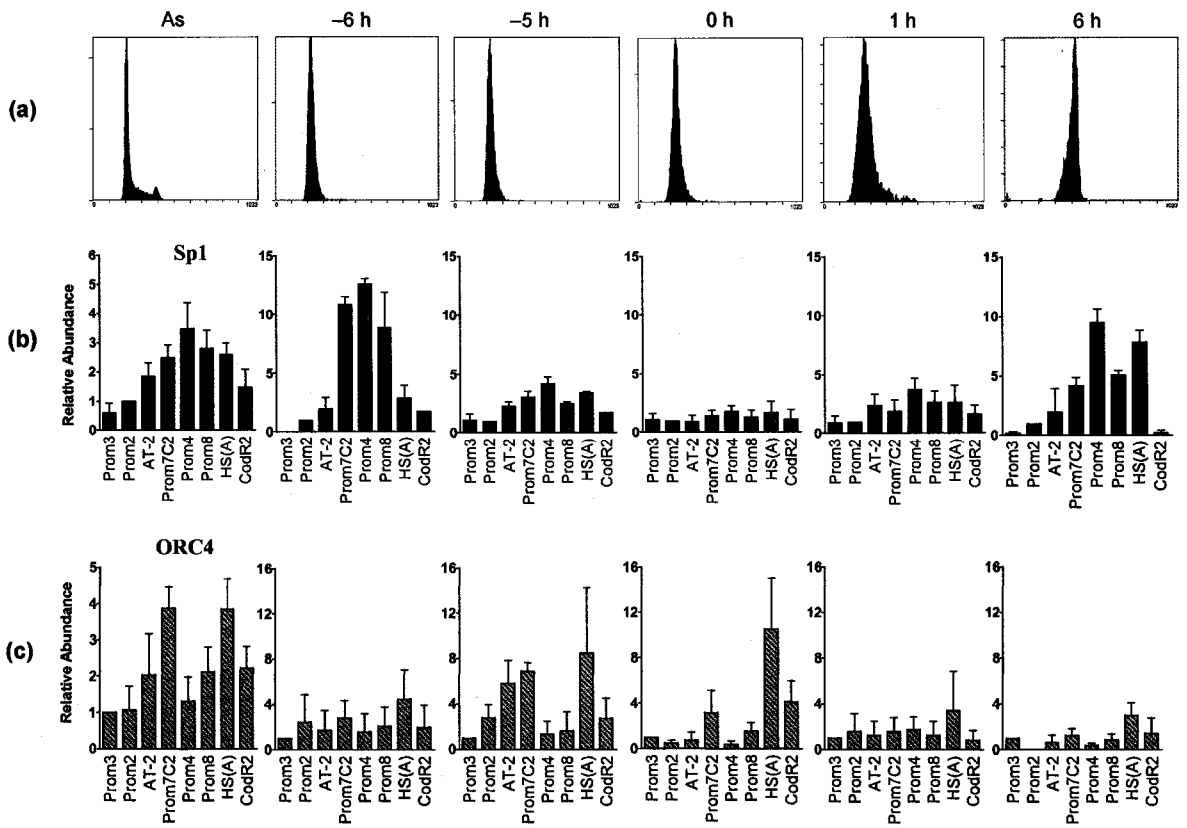
No preferential binding of Sp1 to the *DBF4* promoter was observed in cells arrested at G1/S with double-thymidine (Fig. 19b, 0 h timepoint). However, a progressive increase in Sp1 binding to the promoter-proximal segments over the distal regions was evident after cells were released from the G1/S block. Notably, very high enrichments for the Prom4 and HS(A) segments were obtained at the 6 h timepoint, coinciding with the time when DBF4 mRNA levels peaked (Fig. 18). Although it cannot be ruled out that the high transcript levels observed at this timepoint result from cell cycle-regulated changes in RNA stability, these results suggest that the Sp1 transcription factor is required for efficient transcription from the *DBF4* promoter, as described previously (Wu and Lee, 2002). High levels of Sp1 binding to the promoter were also observed at an earlier timepoint, 6 h prior to the completion of the second thymidine block (Fig. 19b, -6 h). As incubation in thymidine continued and cells progressed further towards S phase (i.e., -5 h and 0 h timepoints), these values progressively decreased, reaching the almost negligible levels described above by the

Figure 19. ***In vivo*** **protein-DNA interactions at the** ***DBF4*** **locus during cell cycle progression**

(a) Flow cytometric analysis of HeLa cells synchronized at different cell cycle positions. As, asynchronous cells; 0 h, cells arrested at the G1/S border; 1 h and 6 h, cells released from double-thymidine arrest for the indicated times (in hours); -5 h and -6 h, cells collected at the indicated times (in hours) **prior to the completion** of the second thymidine incubation.

(b) ChIP assays were performed using anti-Sp1 antibodies and formaldehyde-fixed cells synchronized at the indicated times. Sequence abundance in the immunoprecipitated DNA was determined by Q-PCR with the specified primer pairs. Relative abundance was calculated as described in the legend to Figure 9b, except that the Prom2 segment was used for normalization. Error bars, standard error of the mean of at least two different experiments.

(c) ChIP assays were performed as described in (b) but with antibodies against the human ORC4 protein. Sequence abundance was normalized to that of the Prom3 segment. Error bars, standard error of the mean of at least three different experiments.



end of the block (Fig. 19b).

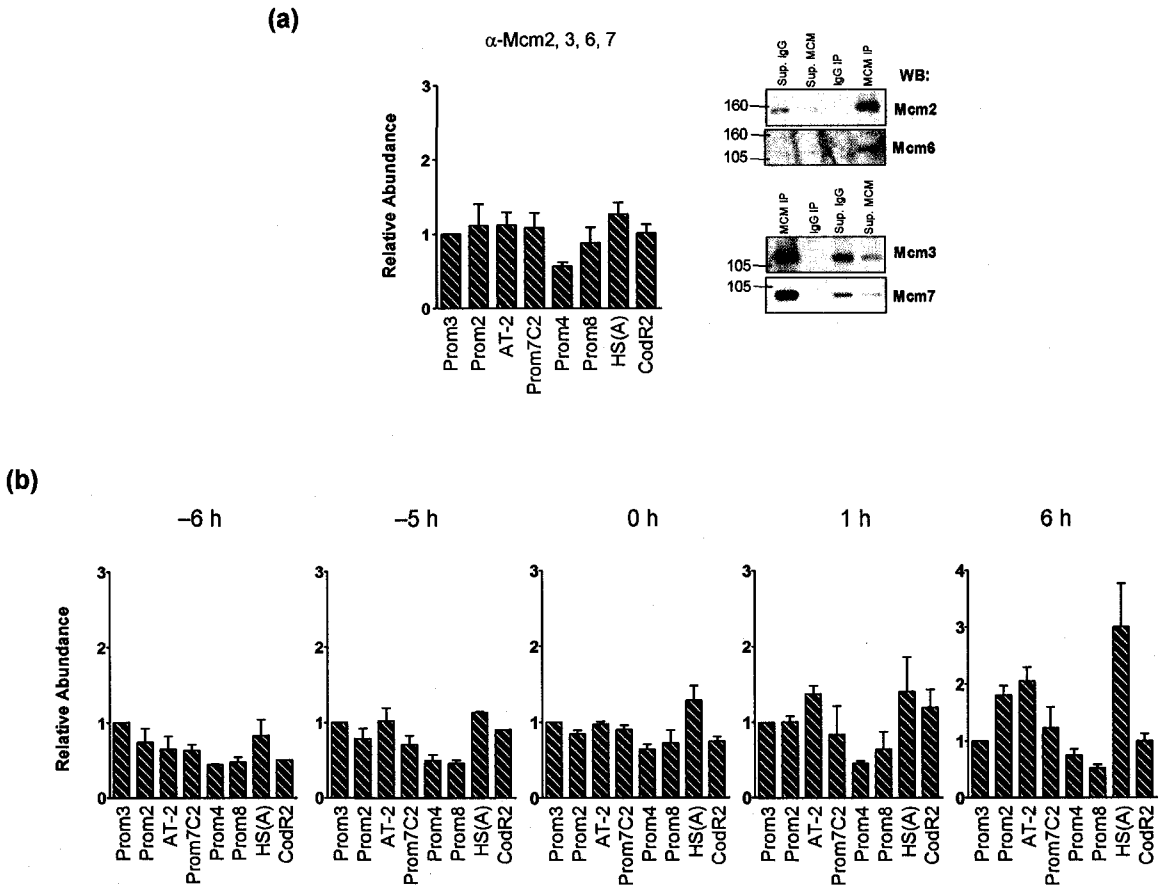
In the case of ORC4, the ChIP experiments with asynchronous cell populations showed that this protein is preferentially bound to two distinct regions at the *DBF4* locus: one is represented by primer pair HS(A), and the other is around the segment amplified by primer set Prom7C2 (Fig. 9). As anticipated, both of these DNA segments were enriched in ORC4-immunoprecipitates from synchronized HeLa cells, with the highest recovery observed at the -5 h timepoint for both regions (Fig. 19c). Somewhat unexpectedly, however, these high enrichment values for both Prom7C2 and HS(A) were not maintained throughout the cell cycle. Instead, the enrichment was gradually diminished as cells progressed from late G1 to G2/M (Fig. 19c, compare -5 h and 6 h timepoints).

Finally, ChIP assays were also performed using a cocktail of antibodies against the MCM complex, another component of the pre-RCs assembled at eukaryotic oris. Despite effective protein precipitation and DNA recovery, no enrichment was observed for any of the analyzed segments when a combination of anti-Mcm2, -Mcm3, -Mcm6 and -Mcm7 antibodies was used to immunoprecipitate chromatin from asynchronous HeLa cells (Fig. 20a). Instead, slightly lower levels of MCM-binding were observed for the DNA segment represented by Prom4 primer set. Similar results were obtained using synchronized cell populations (Fig. 20b). Reduced binding of MCM proteins to the ori region was observed at the -6 h, -5 h and 0 h timepoints although the differences did not appear to be significant (Fig. 20b). In addition, a small increase in MCM-binding to the HS(A) and AT-2 regions occurred at the 6 h timepoint. However, the significance of this observation is not certain because comparable values were also obtained for the Prom2 sequence, which is normally considered to be background. Thus, despite some small variations, the overall results of the

Figure 20. *In vivo* binding of MCM complexes to the *DBF4* locus

(a) ChIP assays were performed using asynchronous HeLa cells and a cocktail of antibodies against the human Mcm2, Mcm3, Mcm6, and Mcm7 proteins. Abundance of DNA sequences in the immunoprecipitated DNA was determined by Q-PCR with indicated primer pairs, and relative abundance was calculated as described in the legend to Figure 9b. Error bars, standard error of the mean of three different experiments. All four proteins were efficiently precipitated, as shown in the Western blots to the right. Sup., cell extracts after immunoprecipitation; IP, precipitated protein fraction. IgG, control immunoprecipitation reactions using pre-immune antibodies.

(b) ChIP assays were performed as in (a) but using synchronized HeLa cells. Synchronization conditions and timepoints were the same as those employed for the Sp1 and ORC4 ChIP experiments of Figure 19.



MCM-ChIP assays point to an even distribution of MCM complexes (or some of its subunits) throughout the entire *DBF4* locus, including the ori, the enhancer and the coding region.

Chapter 4

Discussion

4.1 Summary

The existence of a strong ori in the *DBF4* promoter region was initially determined by nascent strand abundance analysis, and was then confirmed by ChIP assays and RIP mapping (Figs. 6, 9 and 13-15). Certainly, one of the most striking features of this novel ori is that it contains multiple potential start-sites for leading strand DNA synthesis, all of which are distributed within two "replication initiation zones" separated by approximately 400 bp (Fig. 15c). This makes the DBF4 ori unique because all of the other eukaryotic oris that have been characterized in such detail contain single replication start-sites located in close proximity to each other on both parental DNA strands (Bielinsky and Gerbi, 1999; Gómez and Antequera, 1999, Abdurashidova et al., 2000; Kong and DePamphilis, 2002).

Data presented in this thesis also demonstrate that the asymmetric initiation of DNA synthesis described for some prokaryotic oris can be extended to the DBF4 ori and possibly to additional oris in higher eukaryotes. Since this is the first time that this mechanism has been reported for a mammalian chromosomal ori, we have termed it asymmetric bidirectional replication or ABR.

The model shown in Figure 21 compiles a large portion of the data described here and explains how replication initiation at the DBF4 ori may occur through the ABR mechanism. According to this model, leading strand DNA synthesis first starts within replication initiation zone I (which overlaps with the DBF4 core and auxiliary promoter elements) and progresses in the direction of *DBF4* transcription (Fig. 21c). At this stage, initiation from the complementary (antisense) strand could potentially be suppressed by transcription factors bound to the promoter, by transcription activity of the *DBF4* gene, or

Figure 21. Model for asymmetric bidirectional replication (ABR) at the DBF4 ori

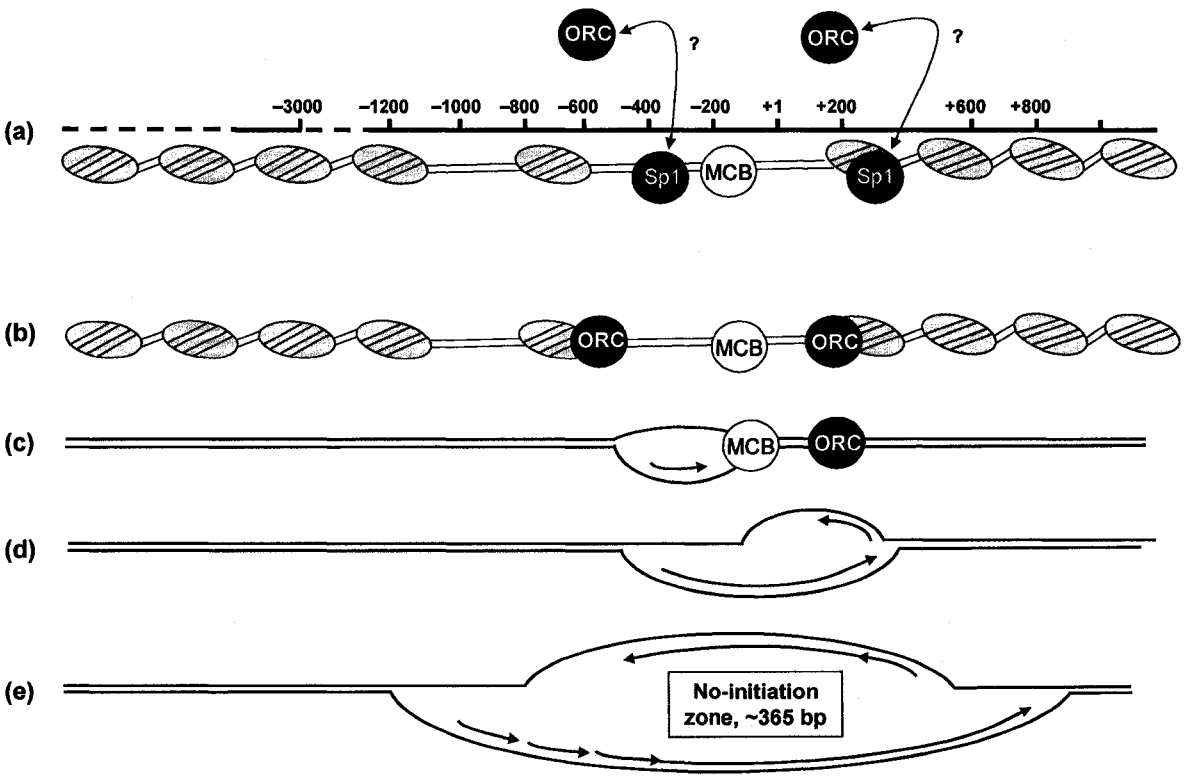
(a) The Sp1 transcription factor may be involved in positioning ORC at the human DBF4 ori. Gray ovals show the proposed position of nucleosomes around the *DBF4* locus. MCB, putative binding site for the MCB transcription factor.

(b) ORC binds to two regions of the *DBF4* locus, both of them located close to each of the two replication initiation zones. Sp1 is released from chromatin following ORC binding.

(c) Leading strand synthesis starts from initiation zone I more frequently and earlier than from zone II. Chromatin structure around zone I is more open than in zone II which could facilitate the assembly and activation of replication protein complexes in this region. Short leading strands emanating from zone I appear to have an unusually long half life, which could result from the transient pause of the replication fork due to the binding of transcription factors such as MCB to the downstream promoter area. ORC4 is released from zone I during the process of activation of DNA synthesis.

(d) Elongation of leading strands emanating from zone I can temporarily disrupt the transcription machinery bound at the promoter and provide a "window of opportunity" for antisense leading strand synthesis from initiation zone II. In addition, sense leading strand movement could "open up" chromatin around zone II to facilitate the initiation of DNA synthesis. Activation of DNA synthesis from zone II is also accompanied by ORC4 release.

(e) Bidirectional replication from two oppositely moving replication forks is finally established at the human DBF4 ori. Leading strand synthesis does not start or starts very rarely within the ~365 bp region that separates initiation zones I and II.



by the fact that a more compact chromatin structure exists around initiation zone II (Figure 21a-c). Eventually, replication forks emanated from zone I can temporarily disrupt this block, providing a short window of opportunity during which initiation of the antisense strand can occur (Fig. 21d). This finally leads to the establishment of two oppositely moving replication forks (Fig. 21e). The most important aspects of this model and its relevance to the DNA replication field will be discussed in detail in the sections that follow.

4.2 Oris and transcriptional promoters

Although no consensus sequence analogous to the yeast ACS has been found in mammalian oris, a subset of them does share certain similarities that could be important for origin selection and activation (Table 1). Indeed, the fact that an ori was successfully identified in the human *DBF4* locus based exclusively on the fact that a transcriptional promoter, a CpG island and two AT-rich stretches were co-localized in this genomic area, highlights how relevant some of these features might be for the specification and function of mammalian oris.

How are each of these elements related to mammalian DNA replication? It is well-known that loosening of chromatin structure is a pre-requisite for the assembly of the big multiprotein complexes that are involved in the initiation of both DNA replication and gene transcription (reviewed by Melendy and Li, 2001). Thus, it is generally accepted that it would be beneficial and economical for a cell to initiate DNA synthesis from sites in which chromatin organization has already been disrupted by transcription. In agreement with this, the results presented here show that the DBF4 ori, particularly replication initiation zone I, maps within a region of relaxed or "open" chromatin structure, which coincides with the

basal and auxiliary promoter elements of the *DBF4* gene (Fig. 17). However, besides what could be perceived as a coincidental relationship between oris and transcriptional promoters due to their common requirement for an open chromatin configuration, it has been suggested that promoters and transcription factors could play a more direct role in the initiation of DNA synthesis. For instance, transcription factors (both cellular and viral) have been shown to positively affect DNA replication in eukaryotic viruses by tethering replication proteins to oris, sometimes via direct protein-protein interactions (Sanders and Stenlund, 1998; Nguyen-Huynh et al., 1998; Baumann et al., 1999; Schepers et al., 2001). The idea of polypeptides such as transcription factors being involved in the recruitment of replication initiation proteins to oris is particularly attractive in mammalian cells because, as reported by others (Vashee et al., 2003; Remus et al., 2004) and demonstrated by the *in vitro* DNA-binding assays of Figure 10, metazoan ORC does not seem to have any intrinsic preferential affinity for ori DNA compared to non-ori sequences (Kohzaki and Murakami, 2005).

In the case of CpG islands, recent estimates indicate that these G +C-rich DNA regions are associated with the promoters of all human housekeeping genes and with approximately 50% of tissue-specific genes (Antequera, 2003). Thus, their proximity to mammalian oris is probably not surprising and might be secondary to their co-localization with gene promoters. Finally, the presence of AT-rich stretches near oris is believed to reflect the requirement for DNA segments that can be easily denatured to facilitate entry of the initiation machinery, although this might not be their only function (Aladjem et al., 2006).

4.3 RIP mapping at the *DBF4* ori locus

The RIP mapping technique has been successfully used to determine precise initiation sites at several yeast oris and at the *S. coprophila* II/9A amplification ori (Gerbi and Bielinsky, 1997; Bielinsky et al., 2001). However, it wasn't really expected that this method could be directly applied to mammalian oris, which is why LM-PCR was utilized to map RIPs at the human lamin B2 ori (Abdurashidova et al., 2000).

The results presented in this thesis clearly show that the one-way PCR-based RIP mapping methodology can indeed be sensitive and powerful enough to study mammalian oris, provided that it is combined with specific sample preparation methods which include: (i) a relatively simple alkaline gel electrophoresis procedure for nascent strand isolation, (ii) direct cell lysis in the wells of alkaline gels to avoid potential DNA fragmentation, and (iii) emetine treatment of the cells prior to nascent DNA isolation in order to enrich for leading strands and prevent the ambiguous results that can be generated by the ligation of Okazaki fragments.

Although some nucleotide positions within each replication initiation zone appeared to be used more frequently than others as replication start-sites (for example -390 within zone I and +310 within zone II in Fig. 15c), the precise location of initiation sites varied between synchronized and asynchronous cells as well as amongst different experiments. However, what remained constant in all the experiments was the fact that DNA synthesis always initiated within the boundaries of the two replication initiation zones. This suggests that, at least in some mammalian oris such as the DBF4 ori, chromatin context and organization might be more important than specific nucleotide residues in determining replication initiation sites.

Interestingly, the presence of two separate replication initiation zones in the DBF4 ori was not observed in the initial nascent strand abundance assays (Fig. 6b). In these experiments, the regions amplified by primers Prom7B and Prom8 (close to initiation zone I) were clearly highly enriched compared to the area around primer set CodR2, which maps near initiation zone II. The existence of replication initiation zone II was only evident when RIP mapping analysis was performed. A possible explanation for this apparent discrepancy came from the observation that, most of the time, many more sense strand RIPs were detected when shorter nascent DNA were used as template than when longer strands were employed. For instance, while numerous RIPs were detected with pA in early S phase using 0.5–1.25 kb templates, only a small number of products were observed with the same primer with 1.25–2 kb-long nascent strands (Fig. 16a). Similarly, up to 49 RIP sites were obtained with pA using 0.5–1 kb DNA from cells at 1 h post-G1/S arrest (Fig. 15b, panel A), whereas only a few bands were observed with 1–2 kb strands (Fig. 15a). These data appear to suggest that, once DNA synthesis from zone I has occurred, replication forks become transiently arrested in the downstream area (maybe due to the presence of transcription-related proteins such as the MCB factor), generating a population of short nascent strands with an unusually long half-life (Fig. 21c). The fact that these strands are long-lived increases the probability of detecting their initiation sites by RIP mapping, and explains why replication initiation zone I is preferentially enriched in the experiments shown in Figure 6b.

4.4 ORC binding to the DBF4 ori

Two different areas of the *DBF4* ori locus [(Prom7C2 and HS(A)] were found to bind ORC4 *in vivo* (Figs. 9c and 21b). This contrasts with the single ORC-binding site

identified in the mammalian lamin B2 and TOP1 oris (Keller et al., 2002; Abdurashidova et al., 2003). Interestingly, both ORC-binding regions at the *DBF4* locus are located close to the two replication initiation zones and near putative binding sites for the Sp1 transcription factor (Fig. 9a and Fig. 21). Thus, it is tempting to hypothesize that Sp1 might be involved in positioning ORC within the DBF4 ori, similar to the mechanism employed by some eukaryotic viruses. As mentioned above, a role for transcription factors in the specification of ORC binding sites in metazoans has been previously suggested (Kohzaki and Murakami, 2005). Furthermore, a potential relationship between Sp1 and ORC and the possibility of direct contact between these proteins at mammalian oris has been reported in the past in the human TOP1 ori (Keller et al., 2002). In this case, based on the results of ChIP assays using exponentially growing cells, it was concluded that ORC2 and Sp1 most likely reside at closely adjacent sites in the ori region, in agreement with the observations for Sp1 and ORC4 at the DBF4 ori.

Notably, however, the ChIP analyses presented here for synchronized cell populations (Fig. 19c) appear to add another level of complexity to the potential interaction between Sp1 and ORC4 at mammalian oris. Indeed, the results of these experiments suggest that the two proteins do not bind simultaneously to DNA. Instead, whenever high levels of Sp1-binding are detected in the DBF4 promoter (for instance at the -6 h and +6 h timepoints), low ORC4-binding is observed in the area, and vice versa. How can this observation be reconciled with the idea that Sp1 is involved in recruitment of ORC to the DBF4 ori? One plausible explanation could be that Sp1 molecules bound to their specific recognition sites in the DBF4 promoter initially direct the binding of ORC to the adjacent DNA sequences. Once this has occurred, the Sp1 proteins may be released from chromatin,

clearing the promoter/ori area to allow further assembly of the DNA replication machinery (Fig. 21a-b). This implies that Sp1 and ORC are only bound simultaneously to DNA for a very brief period of time, making it difficult to detect by ChIP in synchronized cells. Additionally, since the Sp1 transcription factor seems to be necessary for efficient activation of the *DBF4* gene promoter (Wu and Lee, 2002), its dissociation from the DBF4 ori following ORC recruitment might result in reduced rates of *DBF4* gene transcription which could also facilitate DNA replication by minimizing the probability of collisions between replication forks and the transcription machinery. Since it has been suggested that the *DBF4* promoter is bidirectional (Yamada e al., 2002), this effect could also extend to the oppositely transcribed *MCFP* gene. Certainly, further investigation is required to study the relationship, if any, between Sp1 and the ORC at the DBF4 ori. Future experiments should also focus on studying transcription of the *MCFP* gene along the cell cycle and its correlation with DBF4 ori activity.

The model described above does not exclude the possibility that other factors besides Sp1 might also participate in the process of ORC recruitment to the DBF4 ori. For example, DNA topology has recently been identified as a contributing factor in high-affinity binding of *Drosophila* ORC to DNA (Remus et al., 2004) and could certainly play a role in origin specification in mammalian cells. Similarly, other proteins such as the high-mobility-group protein HMGA1a and the transcriptional repressor A1F-C have been implicated in the binding of ORC to specific mammalian DNA sequences, which makes it highly likely that additional factors might work to direct ORC to the DBF4 ori (Minami et al., 2006; Thomae et al., 2008).

Interestingly, recovery of DBF4 ori sequences in ORC4-immunoprecipitates progressively decreased in synchronized cells after the -5 h timepoint (Fig. 19c). This suggests that the protein is either released from the DBF4 ori before or during the process of origin firing, or is modified at this point in such a way that the antibodies used are not able to efficiently recognize it anymore (although unmodified ORC4 molecules associated to late-firing oris could still be pulled down).

Although ORC complexes remain constantly bound to oris in *S. cerevisiae* (Aparicio et al., 1997), cell-cycle dependent dissociation of ORC subunits has been reported in some mammalian oris. The ORC1 protein, for instance, was shown to be released from the lamin B2 ori during S phase (Abdurashidova et al., 2003; Ladenburger et al., 2002). At the global level, ORC proteins are also clearly removed from chromatin as cells pass through S phase (Kreitz et al., 2001; Li and DePamphilis, 2002). For ORC2, contradictory results have been obtained by different ChIP studies. In some cases, the protein has been shown to remain bound to oris during S phase (Abdurashidova et al., 2003), whereas in others it appears to be released (Gerhardt et al., 2006). The reasons for this discrepancy are still unclear, although they might be related to the use of different immunoprecipitation antibodies, different cell fixation techniques or different synchronization conditions. Lastly, a recent study has shown that the stable association of ORC4 with the ORC complex depends on the presence of the ORC1 subunit (Siddiqui and Stillman, 2007). This raises the possibility that ORC4 may be released from oris during S phase due to the dissociation of ORC1 (Siddiqui and Stillman, 2007). The data presented here for the DBF4 ori appear to support this model, although, as mentioned above, it is also possible that the ChIP results reflect conformational changes in ORC4 and not its chromatin dissociation. Further studies should be performed to rule out

this possibility, perhaps using antibodies to different regions of the ORC4 protein as well as to additional subunits of the ORC complex.

Strikingly, the observed decrease in the interaction of ORC4 with the DBF4 ori did not seem to occur simultaneously at the two ORC4-binding sites. Instead, although **both** Prom7C2 and HS(A) were bound by ORC4 at the -5 h timepoint, enrichment values for the Prom7C2 sequence (located within the replication initiation zone I) significantly decreased at the 0 h timepoint, while those for HS(A) still remained very high (Fig. 19c). Recovery of the HS(A) segment (i.e., initiation zone II) was only diminished at a later timepoint (i.e., 1 h post-G1/S arrest), coinciding with maximal DBF4 ori firing. These findings are in remarkable agreement with the ABR model and with the idea that DNA synthesis from the DBF4 ori starts sequentially from the two replication initiation zones. Thus, although pre-replication complexes seem to assemble simultaneously at both zones (Fig. 21b), some molecular event(s) appear to take place during late G1 or G1/S only within zone I, which leads to the release (or conformational change) of ORC4 and the synthesis of DNA leading strands that pause in the downstream area (Fig. 21c). The preferential activation of initiation zone I over zone II may be dictated by the more accessible chromatin arrangement that exists around this area (Fig. 17 and discussion below). Once cells enter S phase, however, fork movement is resumed and the sense leading strands progress through initiation zone II. This opens up chromatin around this region and facilitates the activation of DNA synthesis from the antisense strand, with the concomitant release/modification of the ORC4 molecules bound in this area (Fig. 21d).

4.5 MCM binding to the DBF4 ori

Contrary to what was observed for ORC4, MCM ChIP experiments using asynchronous HeLa cells did not reveal preferential binding of this complex to any particular sequence within the *DBF4* locus (Fig. 20a). Similar observations have been made by others (Alexandrow et al., 2002; Schaarschmidt et al., 2002) although there are reports in the literature of ori sequences being enriched over non-ori DNA in MCM-immunoprecipitates from exponentially growing cells (Ghosh et al., 2006). Likewise, ChIP assays failed to detect any enrichment for DBF4 ori DNA in synchronized cells at any of the cell cycle timepoints analyzed (Fig. 20b), even though previous studies have shown that MCM proteins preferentially bind to oris during G1 and G1/S (Schaarschmidt et al., 2002; Abdurashidova et al., 2003).

The inconsistencies described above are not new to the field of DNA replication. In fact, although MCM complexes have been postulated to travel along with replication forks as the eukaryotic replicative helicase (You et al., 1999), this view is not universally held. According to some studies, MCM proteins are evenly distributed throughout chromosomal DNA, particularly in unreplicated genomic regions, and are gradually released from chromatin as a result of replication fork passage (Krude et al., 1996). Clearly, additional work is required to clarify the precise MCM binding sites at the DBF4 ori. It is possible that the widespread chromosomal distribution observed here resulted from the use of a cocktail of antibodies for immunoprecipitation. It would be interesting to perform similar experiments using antibodies against individual MCM proteins, in order to rule out the possibility that only certain subcomplexes or specific MCM subunits might be preferentially localized to oris.

4.6 Chromatin structure around the DBF4 ori

Two regions of increased nuclease sensitivity were observed in the *DBF4* locus. One of them, spanning from position -1081 to -947, is located close to a putative *DBF4* transcriptional enhancer (Fig. 17; Wu and Lee, 2002). Consequently, the observed sensitivity around this area might result from the binding of transcription factors. The other DNase I-sensitive region overlaps with the *DBF4* gene promoter. Interestingly, the upstream, more "open" boundary of this second hypersensitive area (corresponding to Prom4, from -467 to -398) maps within the DBF4 ori initiation zone I, whereas the downstream, more "resistant" end roughly coincides with replication initiation zone II (Fig. 17). According to these results, it would make sense that DNA replication at the DBF4 ori starts first from initiation zone I, since this area is located within a more relaxed chromatin region where it would be easier to fully assemble and activate the DNA replication machinery. Once initiation has occurred, replication fork progression through the non-initiation area towards zone II could serve to temporarily disrupt chromatin organization in this region, allowing DNA synthesis to begin from the complementary strand.

Detailed analysis of chromatin structure around budding yeast ARS1 has shown that this ori lies within a nucleosome-free DNA area which is flanked at both ends by precisely positioned nucleosomes. These nucleosome boundaries were shown to be imposed on one side by binding of the transcription factor Abf1 and, on the other, by the interaction of ORC with the A and B1 sequence elements (Lipford and Bell, 2001). This organization bears some resemblance to what is observed at the *DBF4* locus, where the ori region seems to be roughly nested within an area of increased DNase-I sensitivity, flanked on both sides by what could potentially be positioned nucleosomes (Fig. 17 and Fig. 21a). In this case, the

boundaries could be imposed by ORC on both sides or by a combination of ORC and transcription factors, which could prevent nucleosome assembly within the replication initiation area, particularly around initiation zone I. Interestingly, in ARS1, the two positioned nucleosomes were shown to have a positive influence on pre-RC assembly at the ori (Lipford and Bell, 2001), suggesting that the chromatin arrangement around the DBF4 ori could potentially have a more active role in origin activity.

4.7 ABR: a new model for mammalian chromosomal DNA replication

The ABR pattern reported here for the DBF4 ori is very similar to the initiation mode described *in vivo* for the *E. coli* oriC locus. In this bacterial ori, multiple, asymmetrically distributed replication start-sites have been mapped on both parental DNA strands, which fire in a temporal manner (Kohara et al., 1985). For the strand replicated in the counterclockwise direction, start-sites are clustered near the left half of oriC, whereas for the complementary strand, initiation sites seem to be located in a more distributed fashion outside of the left boundary of oriC (Kohara et al., 1985; Fang et al., 1999). Although some discrepancies exist regarding its exact size, the two initiation regions of oriC are separated by an area of 70 to 250 bp in which initiation events do not occur very frequently (Kohara et al., 1985; Fang et al., 1999). Interestingly, initiation from oriC in plasmids was more reminiscent of the classical OBR pattern (Seufert and Messer, 1987), suggesting that chromosomal context might be important in determining replication initiation sites and replication patterns even in bacterial cells.

It is possible that the ABR initiation mechanism described for the DBF4 ori could be widespread amongst mammalian oris. In fact, there are two other cases of metazoan oris that

seem to potentially operate through ABR: the mitochondrial D-loop ori, and the II9/A amplification ori of *S. coprophila.*

Although the exact mechanism of mitochondrial DNA replication is rather controversial, the most accepted model is still the strand-displacement model, which proposes that initiation occurs asymmetrically from separate oris located several kilobases apart on the two parental strands (Clayton, 1991; Brown et al., 2005). DNA synthesis is believed to start first from the "origin of heavy (H)-strand synthesis" (O_H) and proceed unidirectionally for about two-thirds of the circular mitochondrial genome. Fork movement eventually exposes potential "origins of light (L)-strand synthesis" (O_L) on the complementary strand which are then used to initiate DNA synthesis in the opposite direction (Brown et al., 2005; Falkenberg et al., 2007). Due to the large size of the region that separates the two replication initiation areas mitochondrial DNA replication constitutes a somewhat "extreme" example of the ABR initiation mechanism.

As mentioned in Chapter 1 (see section 1.4.4), a clear transitional point was observed in the II9/A ori on one of the strands, with several different initiation sites detected on the other (Bielinsky et al., 2001). Since emetine was not included in these experiments, it is difficult to establish at this time whether these sites correspond to leading strand RIPs or to growing chains ligated to Okazaki fragments. The latter scenario would imply that a precise TP might exist on both strands although one of them may lie outside of the analyzed area. In either case, as discussed by the authors of the study, ori II/9A seems to be more similar to *E. coli* oriC than to other eukaryotic oris. This, they propose, could be related to the fact that II/9A is an amplification ori and, for this reason, might not be regulated in the same way as oris involved in regular genome duplication (Bielinsky et al., 2001). However, based on the

data presented here, it is also possible that ori II9/A, just like the DBF4 ori, belongs to a potentially large group of metazoan oris that initiate DNA synthesis by ABR.

Notably, Dijkwel and colleagues have recently suggested that, at the hamster *DHFR* locus, two of the most frequently used oris within the 55 kb intergenic spacer (ori-β and ori-β') resemble *E. coli*'s oriC, and do not correspond to oris with precise initiation sites such as yeast ARS1 and the lamin B2 ori (Dijkwel et al., 2002). Data from the DBF4 ori could potentially support this hypothesis and provide an explanation for the highly delocalized initiation observed at DHFR ori, which could result from the presence of several evenly-spaced ABR-operating oris.

4.8 Perspectives

Out of an estimated number of at least 30,000 oris (Aladjem et al., 2006), the human lamin B2 ori has been for many years the only mammalian ori characterized in detail and mapped with nucleotide resolution. Somewhat surprisingly, this ori was shown to have a very similar mode of initiation to that of the oris of lower eukaryotes (i.e., OBR), even though notable differences exist between mammalian and yeast oris in terms of structure and organization.

The in-depth characterization of the DBF4 ori presented here points to the existence of a second replication initiation mechanism in mammalian cells. This suggests that mammalian oris can belong to one of at least two different groups: those, like the lamin B2 ori, in which initiation occurs according to the traditional OBR model, and others, represented by the DBF4 ori, in which initiation takes place asymmetrically from multiple potential sites following the ABR model. Based on the available data, ORC-binding might

play a crucial role in determining where DNA replication initiates, and whether OBR or ABR take place at a particular ori. Thus, while a single ORC-binding region has been identified in close proximity to the position of the RIPs at the lamin B2 ori (Abdurashidova et al., 2003), ORC binds to two areas at the DBF4 ori, both of them located in the vicinity of the two replication initiation zones.

Considering that the existence of two initiation zones at the DBF4 ori was only evident after performing RIP mapping, it is possible that ABR might be quite common amongst mammalian oris. This, however, has not been revealed so far because, with the exception of the lamin B2 ori, mammalian oris have only been mapped using techniques such as nascent strand abundance assays. Hopefully, the technical modifications described here for the RIP mapping method will facilitate the application of high-resolution mapping to additional oris, and will provide a better idea of how widespread ABR initiation is.

Although the study of mammalian oris has been somewhat dormant during the past few years, the results and conclusions presented in this thesis open up new and exciting possibilities in this area, and can supply the incentive needed to reignite interest in the field. In the future, RIP mapping combined with ChIP analysis of protein-DNA interactions at other oris will help elucidate what factors are involved in mammalian ori recognition and function. This information is essential to significantly advance our knowledge about mammalian DNA replication and to eventually understand how this process is deregulated in disease and how it might contribute to the development and progression of pathologies such as cancer.

REFERENCES

Abdurashidova, G., Danailov, M. B., Ochem, A., Triolo, G., Djeliova, V., Radulescu, S., Vindigni, A., Riva, S., and Falaschi, A. (2003). Localization of proteins bound to a replication origin of human DNA along the cell cycle. EMBO J. *22*, 4294–4303.

Abdurashidova, G., Deganuto, M., Klima, R., Riva, S., Biamonti, G., Giacca, M., and Falaschi, A. (2000). Start sites of bidirectional DNA synthesis at the human Lamin B2 origin. Science *287*, 2023–2026.

Aladjem, M. I., Falaschi, A. and Kowalski, D. (2006). Eukaryotic DNA replication origins. In DNA replication and human disease, M. L. DePamphilis, ed. (Cold Spring Harbor, USA: Cold Spring Harbor Laboratory Press), pp. 31–61.

Alexandrow, M. G., Ritzi, M., Pemov, A., and Hamlin, J. L. (2002). A potential role for mini-chromosome maintenance (MCM) proteins in initiation at the dihydrofolate reductase replication origin. J. Biol. Chem. *277*, 2702–2708.

Antequera, F. (2003). Structure, function and evolution of CpG island promoters. Cell. Mol. Life Sci. *60*, 647–1658.

Aparicio, O. M., Weinstein, D. M., and Bell, S. P. (1997). Components and dynamics of DNA replication complexes in S. cerevisiae: redistribution of MCM proteins and Cdc45p during S phase. Cell *91*, 59–69.

Austin, R. J., Orr-Weaver, T. L., and Bell, S. P. (1999). Drosophila ORC specifically binds to ACE3, an origin of DNA replication control element. Genes Dev. *13*, 2639–2649.

Barry, E. R., and Bell, S. D. (2006). DNA replication in the Archaea. Microbiol. Mol. Biol. Rev. *70*, 876–887.

Baumann, M., Feederle, R., Kremmer, E., and Hammerschmidt, W. (1999). Cellular transcription factors recruit viral replication proteins to activate the Epstein-Barr virus origin of lytic DNA replication, *oriLyt*. EMBO J. *18*, 6095–6105.

Bazar, L., Meighen, D., Harris, V., Duncan, R., Levens, D., and Avigan, M. (1995). Targeted melting and binding of a DNA regulatory element by a transactivator of c-myc. J. Biol. Chem. *270*, 8241–8248.

Bell, S.P., and Stillman, B. (1992). ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. Nature *357*, 128–134.

Berquist, B., and DasSarma, S. (2003). An archaeal chromosomal autonomously replicating sequence element from and extreme halophile, Halobacterium sp. strain NRC-1. J. Bacteriol. *185*, 5959–5966.

Biamonti, G., Giacca, M., Perini, G., Contreas, G., Zentilin, L., Weighardt, F., Guerra, M., Della Valle, G., Saccone, S., Riva, S., and Falaschi, A. (1992). The gene for a novel human lamin maps at a highly transcribed locus of chromosome 19 which replicates at the onset of S-phase. Mol. Cell. Biol. *12*, 3499–3506.

Bielinsky, A. –K. 2003. Replication origins: why do we need so many? Cell Cycle *2*, 307–309.

Bielinsky, A.-K., Blitzblau, H., Beall, E. L., Ezrokhi, M., Smith, H. S., Botchan, M. R., and Gerbi, S. A. (2001). Origin recognition complex binding to a metazoan replication origin. Curr. Biol. *11*, 1427–1431.

Bielinsky, A-K., and Gerbi, S. (1999). Chromosomal ARS1 has a single leading strand start site. Mol. Cell. *3*, 447–486.

Blow, J. J. 2001. Control of chromosomal DNA replication in the early *Xenopus* embryo. EMBO J. *20*, 3293–3297.

Bolon, Y.-T, and Bielinsky, A.-K. (2006). The spatial arrangement of ORC binding modules determines the functionality of replication origins in budding yeast. Nucleic Acids Res. *34*, 5069–5080.

Bousset, K., and Diffley, J. F. X. (1998). The Cdc7 protein kinase is required for origin firing during S phase. Genes Dev. *12*, 480–490.

Bozzoni, I., Baldari, C. T., Amaldi, F., and Buongiorno-Nardelli, M. (1981). Replication of ribosomal DNA in *Xenopus laevis*. Eur. J. Biochem. *118*, 585–590.

Bramhill, D., and Kornberg, A. (1988). A model for initiation at origins of DNA replication. Cell *54*, 915–918.

Brewer, B. J., and Fangman, W. L. (1987). The localization of replication origins on ARS plasmids in *S. cerevisiae*. Cell *51*, 463-471.

Brown, T. A., Cecconi, C., Tkachuk, A. N., Bustamante, C., and Clayton, D. A. (2005). Replication of mitochondrial DNA occurs by strand displacement with alternative light-strand origins, not via a strand-coupled mechanism. Genes Dev. *19*, 2466–2476.

Burhans, W. C., Vassilev, L. T., Wu, J., Sogo, J. M., Nallaseth, F. S., and DePamphilis, M. L. (1991). Emetine allows the identification of origins of mammalian DNA replication by imbalanced DNA synthesis, not through conservative nucleosome segregation. EMBO J. *10*, 4351–4360.

Burhans, W. C., Vassilev, L. T., Caddle, M. S., Heintz, N. H., and DePamphilis, M. L. (1990). Identification of an origin of bidirectional DNA replication in mammalian chromosomes. Cell *62*, 955–965.

Chong, J. P., Hayashi, M. K., Simon, M. N., Xu, R.-M., and Stillman, B. (2000). A double-hexamer archaeal minichromosome maintenance protein is an ATP-dependent DNA helicase. Proc. Natl. Acad. Sci. USA. *97*, 1530–1535.

Chuang, R.-Y., and Kelly, T. J. (1999). The fission yeast homologue of Orc4p binds to replication origin DNA via multiple AT-hooks. Proc. Natl. Acad. Sci. USA. *96*, 2656–2661.

Clayton, D. A. (1991). Replication and transcription of vertebrate mitochondrial DNA. Annu. Rev. Cell. Biol. *7*, 453–478.

Clyne, R. K., and Kelly, T. J. (1995). Genetic analysis of an ARS element from the fission yeast *Schizosaccharomyces pombe*. EMBO J. *14*, 6348–6357.

Dai, J., Chuang, R. –Y., and Kelly, T. J. (2005). DNA replication origins in the *Schizosaccharomyces pombe* genome. Proc. Natl. Acad. Sci. USA. *102*, 337–342.

Danis, E., Brodolin, K., Menut, S., Maiorano, D., and Girard-Reydet, C. (2004). Specification of a DNA replication origin by a transcription complex. Nature Cell Biol. *6*, 721–730.

Davey, M. J., Fang, L., McInerney, P., Georgescu, R. E., and O'Donnell, M. (2002). The DnaC helicase is a dual ATP/ADP switch protein. EMBO J. *21*, 3148–3159.

Decker, R. S., Yamaguchi, M., Possenti, R., and DePamphilis, M. L. (1986). Initiation of Simian Virus 40 DNA replication in vitro: aphidicolin causes accumulation of early-replicating intermediates and allows determination of the initial direction of DNA synthesis. Mol. Cell. Biol. *6*, 3815–3825.

Delgado, S., Gómez, M., Bird, A., and Antequera, F. (1998). Initiation of DNA replication at CpG islands in mammalian chromosomes. EMBO J. *17*, 2426–2435.

Delidakis, C., and Kafatos, F. C. (1989). Amplification enhancers and replication origins in the autosomal chorion gene cluster of *Drosophila*. EMBO J. *8*, 891–901.

DePamphilis, M. L. (1993). Origins of DNA replication on metazoan chromosomes. J. Biol. Chem. *286*, 1–4.

Diffley, J. F. X. (2004). Regulation of early events in chromosome replication. Curr. Biol. *14*, R778–R786.

Dijkwel, P. A., Vaughn, J. P., and Hamlin, J. L. (1994). Replication initiation sites are distributed widely in the amplified CHO dihydrofolate reductase domain. Nucleic Acids Res. *22*, 4989–4996.

Dijkwel, P. A., Mesner, L. D., Levenson, V. V., d'Anna, J., and Hamlin, J. L. (2000). Dispersive initiation of replication in the Chinese hamster rhodopsin locus. Exp. Cell. Res. *256*, 150–157.

Dijkwel, P. A., Wang, S., and Hamlin, J. L. (2002). Initiation sites are distributed at frequent intervals in the Chinese hamster dihydrofolate reductase origin of replication but are used with very different efficiencies. Mol. Cell. Biol. *22*, 3053–3065.

Dimitrova, D. S., and Gilbert, D. M. (2000). Temporally coordinated assembly and disassembly of replication factories in the absence of DNA synthesis. Nat. Cell Biol. *2*, 686–694.

Dimitrova, D.S., Giacca, M., Demarchi, F., Biamonti, G., Riva, S., and Falaschi, A. (1996). In vivo protein-DNA interactions at human DNA replication origin. Proc. Natl. Acad. Sci. USA. *93*, 1498–1503.

Donaldson, A. D., Fangman, W. L., and Brewer, B. J. (1998). Cdc7 is required throughout the yeast S phase to activate replication origins. Genes Dev. *12*, 491–501.

Edwards, M. C., Tutter, A. V., Cvetic, C., Gilbert, C. H., Prokhorova, T. A., and Walter, J. C. (2002). MCM2-7 complexes bind chromatin in a distributed pattern surrounding the origin recognition complex in *Xenopus* egg extracts. J. Biol. Chem. *277*, 33049–33057.

Falkenberg, M., Larsson, N.-G., and Gustafsson, C. M. (2007). DNA replication and transcription in mammalian mitochondria. Annu. Rev. Biochem. *76*, 679–699.

Fang, L., Davey, M. J., and O'Donnell, M. (1999). Replisome assembly at *oriC*, the replication origin of *E. coli*, reveals an explanation for initiation sites outside an origin. Mol. Cell. *4*, 541–553.

Feng, W., Collingwood, D., Boeck, M. E., Fox, L. A., Alvino, G. M., Fangman, W. L., Raghuraman, M. K., and Brewer, B. J. (2006). Genomic mapping of single-stranded DNA in hydroxyurea-challenged yeasts identified origins of replication. Nat. Cell Biol. *8*, 148–155.

Françon, P., Maiorano, D., and Méchali, M. (1999). Initiation of DNA replication in eukaryotes: questioning the origin. FEBS Letters *452*, 87–91.

Gerbi, S. A., and Bielinsky, A.-K. (1997). Replication initiation point mapping. Methods *13*, 271–280.

Gerbi, S. A., Strezoska, Z., and Waggener, J. M. (2003). Initiation of DNA replication in multicellular eukaryotes. J. Struc. Biol. *140*, 17–30.

Gerhardt, J., Jafar, S., Spindler, M.-P., Ott, E., and Schepers, A. (2006). Identification of new human origins of DNA replication by an origin-trapping assay. Mol. Cell. Biol. *26*, 7731–7746.
Ghosh, M., Kemp, M., Liu, G., Ritzi, M., Schepers, A., and Leffak, M. (2006). Differential binding of replication proteins across the human c-myc replicator. Mol. Cell. Biol. *26*, 5270 –5283.

Giacca, M., Pelizon, C., and Falaschi, A. (1997). Mapping replication origins by quantifying relative abundance of nascent DNA strands using competitive polymerase chain reaction. Methods *13*, 301–312.

Giacca, M., Zentilin, L., Norio, P., Diviacco, S., Dimitrova, D., Contreas, G., Biamonti, G., Perini, G., Weighardt, F., Riva, S., and Falaschi, A. (1994). Fine mapping of a replication origin of human DNA. Proc. Natl. Acad. Sci. *91*, 7119–7123.

Gilbert, D. M. (2001). Making sense of eukaryotic DNA replication origins. Science *294*, 96 –100.

Girard-Reydet, C., Grégoire, D., Vassetzky, Y., and Méchali, M. (2004). DNA replication initiates at domains overlapping with nuclear matrix attachment regions in the *Xenopus* and mouse *c-myc* promoter. Gene *332*, 129–138.

Gómez, M., and Antequera, F. (1999). Organization of DNA replication origins in the fission yeast genome. EMBO J. *18*, 5683–5690.

Goren, A., and Cedar, H. (2003). Replicating by the clock. Nat. Rev. Mol. Cell. Biol. *4*, 25–32.

Guo, B., and Lee, H. (2001). Cloning and characterization of Chinese hamster homologue of yeast DBF4 (ChDBF4). Gene *264*, 249–256.

Haitina, T., Lindblom, J., Renstrom, T, and Fredriksson, R. (2006). Fourteen novel human members of mitochondrial solute carrier family 25 (SLC25) widely expressed in the central nervous system. Genomics *88*, 779–790.

Hamlin, J. L. (1992). Initiation of replication in mammalian chromosomes. Crit. Rev. Eukaryot. Gene Expr. *2*, 359–381.

Handeli, S., Klar, A., Meuth, M., and Cedar, H. (1989). Mapping replication units in animal cells. Cell *57*, 909–920.

Hay, N., Bishop, J. M., and Levens, D. (1987). Regulatory elements that modulate expression of human c-myc. Genes Dev. *1*, 659–671.

Hay, R. T., and DePamphilis, M. L. (1982). Initiation of SV40 DNA replication in vivo: location and structure of 5' ends of DNA synthesized in the *ori* region. Cell *28*, 767–779.

Heck, M. M., and Spradling, A. C. (1990). Multiple replication origins are used during *Drosophila* chorion gene amplification. J. Cell. Biol. *110*, 903–914.
Hendrickson, E. A., Fritze, C. E., Folk, W. R., and DePamphilis, M. L. (1987). The origin of bidirectional DNA replication in polyoma virus. EMBO J. *6*, 2011–2018.

Heinemeyer, T., Wingender, E., Reuter, I., Hermjakob, H., Kel, A. E., Kel, O. V., Ignatieva, E. V., Ananko, E. A., Podkolodnaya O. A., Kolpakov, F. A.,., Podkolodny, N. L., and Kolchanov, N. A. (1998). Databases on transcriptional regulation: TRANSFAC, TRRD, and COMPEL. Nucleic Acids Res. *26*, 362-367.

Hendrickson, E. A., Fritze, C. E., Folk, W. R., and DePamphilis, M. L. (1987). The origin of bidirectional DNA replication in polyoma virus. EMBO J. *6*, 2011-2018.

Herrick, J., Stanislawski, P., Hyrien, O., and Bensimon, A. (2000). Replication fork density increases during DNA synthesis in X. laevis egg extracts. J. Mol. Biol. *300*, 1133–1142.

Hirose, S., Hiraga, S., and Okazaki, T. (1983). Initiation site of deoxyribonucleotide polymerization at the replication origin of the *Escherichia coli* chromosome. Mol. Gen. Genet. *189*, 422–431.

Huberman, J. A., Spotila, L. D., Nawotka, K. A., El-Assouli, S. M., and Davis, L. R. (1987). The in vivo replication origin of the yeast 2μm plasmid. Cell *51*, 473–481.

Hyrien, O., and Méchali, M. (1992). Plasmid replication in *Xenopus* eggs and egg extracts: a 2D gel electrophoretic analysis. Nuclei Acids Res. *20*, 1463–1469.

Hyrien, O., and Méchali, M. (1993). Chromosomal replication initiates and terminates at random sequences but at regular intervals in the ribosomal DNA of *Xenopus* early embryos. EMBO J. *12*, 4511–4520.

Hyrien, O., Maric, C., and Méchali, M. (1995). Transition in specification of embryonic metazoan DNA replication origins. Science *270*, 994–997.

Iguchi-Ariga, S. M. M., Okazaki, T., Itani, T., Ogata, M., Sato, Y., and Ariga, H. (1988). An initiation site of DNA replication with transcriptional enhancer activity present upstream of the *c-myc* gene. EMBO J. *7*, 3135–3142.

Jacob, F., Brenner, S., and Cuzin, F. (1963). On the regulation of DNA synthesis in bacteria: the hypothesis of the replicon. Cold Spring Harbor Symp. Quant. Biol. *28*, 329–348.

Jiang, W., McDonald, D., Hope, T. J., and Hunter, T. (1999). Mammalian Cdc7-Dbf4 protein kinase complex is essential for initiation of DNA replication. EMBO J. *18*, 5703–5713.

Johnston, L. H., and Barker, D. G. (1987). Characterization of an autonomously replicating sequence from the fission yeast *Schizosaccharomyces pombe*. Mol. Gen Genet. *207*, 161–164.

Kamath, S., and Leffak, M. (2001). Multiple sites of replication initiation in the human β-globin gene locus. Nucleic Acids Res. *29*, 809–817.

Keller, C., Ladenburger, E.-M., Kremer, M., and Knippers, R. (2002). The origin recognition complex marks a replication origin in the human *TOP1* gene promoter. J. Biol. Chem. *277*, 31430–31440

Kelman, L. M., and Kelman, Z. (2003). Archaea: an archetype for replication initiation studies? Mol. Microbiol. *48*, 605–615.

Kim, S.-M., and Huberman, J. A. (1998). Multiple orientation-dependent, synergistically interacting, similar domains in the ribosomal DNA replication origin of the fission yeast *Schizosaccharomyces pombe*. Mol. Cell. Biol. *18*, 7294–7303.

Kobayashi, T., Rein, T., and DePamphilis, M. L. (1998). Identification of primary initiation sites for DNA replication in the Hamster dihydrofolate reductase gene initiation zone. Mol. Cell. Biol. *18*, 3266–3277.

Kohara, Y., Tohdoh, N., Jiang, X.-W., and Okazaki, T. (1985). The distribution of RNA primed initiation sites of DNA synthesis at the replication origin of Escherichia coli chromosome. Nucleic Acids Res. *13*, 6847–6866.

Kohzaki, H., and Murakami, Y. (2005). Transcription factors and DNA replication origin selection. BioEssays *27*, 1107–1116.

Kong, D., and DePamphilis, M. L. (2002). Site-specific ORC binding, pre-replication complex assembly and DNA synthesis at *Schizosaccharomyces pombe* replication origins. EMBO J. *21*, 5567–5576.

Kong, D., Coleman, T., and DePamphilis, M. L. (2003). *Xenopus* origin recognition complex (ORC) initiates DNA replication preferentially at sequences targeted by *Schizosaccharomyces pombe* ORC. EMBO J. *22*, 3441–3450.

Kornberg, A., and Baker, T. A. (1992). DNA replication, 2nd edition. (New York, USA: WH Freeman and Company).

Kreitz, S., Ritzi, M., Baack, M., and Knippers, R. (2001). The human origin recognition complex protein 1 dissociates from chromatin during S phase in HeLa cells. J. Biol. Chem. *276*, 6337–6342.

Krude, T., Musahl, C., Laskey, R. A., and Knippers, R. (1996). Human replication proteins hCdc21, hCdc46 and P1Mcm3 bind chromatin uniformly before S-phase and are displaced locally during DNA replication. J. Cell Sci. *109*, 309–318.

Kumagai, H., Sato, N., Yamada, M., Mahony, D., Seghezzi, W., Lees, E., Arai, K., and Masai, H. (1999). A novel growth- and cell cycle-regulated protein, ASK, activates human Cdc7-related kinase and is essential for G1/S transition in mammalian cells. Mol. Cell. Biol. *19*, 5083–5095.

Kusic, J., Kojic, S., Divac, A., and Stefanovic, D. (2005). Noncanonical DNA elements in the lamin B2 origin of DNA replication. J. Biol. Chem. *280*, 9848–9854.

Labib, K., and Gambus, A. (2007). A key role for the GINS complex at DNA replication forks. TRENDS in Cell Biol. *17*, 271–278.

Ladenburger, E. –M., Keller, C., and Knippers, R. 2002. Identification of a binding region for human origin recognition complex proteins 1 and 2 that coincides with an origin of DNA replication. Mol. Cell. Biol. *22*, 1036-1048.

Lee, J. -K., Moon, K. –Y., Jiang, Y., and Hurwitz, J. (2001). The *Schizosaccharomyces pombe* origin recognition complex interacts with multiple AT-rich regions of the replication origin DNA by means of the AT-hook domains of the spOrc4 protein. Proc. Natl. Acad. Sci. USA. *98*, 13589–13594.

Lemaitre, J. M., Geraud, G., and Méchali, M. (1998). Dynamics of the genome during early *Xenopus laevis* development: karyomeres as independent units of replication. J. Cell. Biol. *142*, 1159–1166.

Li, C. J., and DePamphilis, M. L. (2002). Mammalian Orc1 protein is selectively released from chromatin and ubiquitinated during the S-to-M transition on the cell division cycle. Mol. Cell. Biol. *22*, 105–116.

Liang, C., Spitzer, J. D., Smith, H. S., and Gerbi, S. A. (1993). Replication initiates at a confined region during DNA amplification in *Sciara* DNA puff II/9A. Genes Dev. *7*, 1072–1084.

Lipford, J. R., and Bell S. P. (2001). Nucleosomes positioned by ORC facilitate the initiation of DNA replication. Mol. Cell *7*, 21–30.

Little, R. D., Platt, T. H., and Schildkraut, C. L. (1993). Initiation and termination of DNA replication in human rRNA genes. Mol. Cell. Biol. *13*, 6600–6613.

Liu, G., Malott, M., and Leffak, M. (2003). Multiple functional elements comprise a mammalian chromosomal replicator. Mol. Cell. Biol. *23*, 1832–1842.

Lopez, P., Philippe, H., Myllykallio, H., and Forterre, P. (1999). Identification of putative chromosomal origins of replication in Archaea. Mol. Microbiol. *32*, 883–886.

Lu, L., Zhang, H., and Tower, J. (2001). Functionally distinct, sequence-specific replicator and origin elements are required for Drosophila chorion gene amplification. Genes Dev. *15*, 134–146.

Lucas, I., Palakodeti, A., Jiang, Y., Young, D. J., Jiang, N., Fernald, A. A., and Le Beau, M. M. (2007). High-throughput mapping of origins of DNA replication in human cells. EMBO Rep. *8*, 770–777.

Lundgren, M., Andersson, A., Chen, L., Nilsson, P., and Bernander, R. (2004). Three replication origins in Sulfolobus species: synchronous initiation of chromosomal replication and asynchronous termination. Proc. Natl. Acad. Sci. USA. *101*, 7046–7051.

MacAlpine, D. M., Rodríguez, H. K., and Bell, S. P. (2004). Coordination of replication and transcription along a Drosophila chromosome. Genes Dev. *18*, 3094–3105.

Mahbubani, H. M., Paull, T., Elder, J. K., and Blow, J. J. (1992). DNA replication initiates at multiple sites on plasmid DNA in *Xenopus* egg extracts. Nucleic Acids Res. *20*, 1457–1462.

Majerník, A. I., Jenkinson, E. R., and Chong, J. P. (2004). DNA replication in thermophiles. Biochem. Soc. Trans. *32*, 236–239.

Malott, M., and Leffak, M. (1999). Activity of the *c-myc* replicator at an ectopic chromosomal location. Mol. Cell. Biol. *19*, 5685–5695.

Marahrens, Y., and Stillman, B. (1992). A yeast chromosomal origin of DNA replication defined by multiple functional elements. Science *255*, 817–823.

Matsunaga, F., Norais, C., Forterre, P., and Myllykallio, H. (2003). Identification of short "eukaryotic" Okazaki fragments synthesized from a prokaryotic replication origin. EMBO Rep. *4*, 154–158.

Matsunaga, F., Forterre, P., Ishino, Y., and Myllykallio, H. (2001). *In vivo* interactions of archaeal Cdc6/Orc1 and minichromosome maintenance proteins with the replication origin. Proc. Natl. Acad. Sci. USA. *98*, 11152–11157.

McWhinney, C., and Leffak, M. (1990). Autonomous replication of a DNA fragment containing the chromosomal replication origin of the human *c-myc* gene. Nucleic Acids Res. *18*, 1233–1242.

Méchali, M., and Kearsey, S. (1984). Lack of specific sequence requirement for DNA replication in *Xenopus* eggs compared with high sequence specificity in yeast. Cell *38*, 55–64.

Melendy, T., and Li, R. (2001). Chromatin remodeling and initiation of DNA replication. Front. Biosci. *6*, D1048–D1053.

Messer, W. (2002). The bacterial replication initiator DnaA. DnaA and oriC, the bacterial mode to initiate DNA replication. FEMS Microbiol. Rev. *26*, 355–374.

Minami, H., Takahashi, J., Suto, A., Sayito, Y., and Tsutsumi, K. (2006). Binding of AIF-C, an Orc1-binding transcriptional regulator, enhances replicator activity of the rat aldolase B origin. Mol. Cell. Biol. *26*, 8770–8780.

Myllykallio, H., Lopez, P., Lopez-Garcia, P., Heilig, R., Saurin, W., Zivanovic, Y., Philippe, H., and Forterre, P. (2000). Bacterial mode of replication with eukaryotic-like machinery in a hyperthermophilic archaeon. Science *288*, 2212–2215.

Newlon, C. S., and Theis, J. F. (1993). The structure and function of yeast ARS elements. Curr. Opin. Gen. Dev. *3*, 752–758.

Nguyen-Huynh, A. T., and Schaffer, P. A. (1998). Cellular transcription factors enhance herpes simplex virus type 1 oriS-dependent DNA replication. J. Virol. *72*, 3635–3645.

Oka, A., Sugimoto, K., Takanami, M., and Hirota, Y. (1980). Replication origin of the Escherichia coli K-12 chromosome: the size and structure of the minimum DNA segment carrying the information for autonomous replication. Mol. Gen. Genet. *178*, 9-20.

Okuno, Y., Satoh, H., Sekiguchi, M., and Masukata, H. (1999). Clustered adenine/thymine stretches are essential for function of a fission yeast replication origin. Mol. Cell. Biol. *19*, 6699–6709.

Olsen, G. J., and Woese, C. R. (1997). Archaeal genomics: an overview. Cell *89*, 991–994.

Patel, P. K., Arcangioli, B., Baker, S. P., Bensimon, A., and Rhind, N. (2006). DNA replication origins fire stochastically in fission yeast. Mol. Biol. Cell. *17*, 308–316.

Pearson, C. E., Zorbas, H., Price, G. B., and Zannis-Hadjopoulos, M. (1996). Inverted repeats, stem-loops, and cruciforms: significance for initiation of DNA replication. J. Cell. Biochem. *63*, 1–22.

Phi-van, L., Sellke, C., vo Bodenhausen, A., and Strätling, W. H. (1998). An initiation zone of chromosomal DNA replication at the chicken lysozyme gene locus. J. Biol. Chem. *273*, 18300–18307.

Raghuraman, M. K., Winzeler, E. A., Collingwood, D., Hunt, S., Wodicka, L., Conway, A., Lockhart, D. J., Davis, R. W., Brewer, B. J., and Fangman, W. L. (2001). Replication dynamics of the yeast genome. Science *294*, 115–121

Rao, H., and Stillman, B. (1995). The origin recognition complex interacts with a bipartite DNA binding site within yeast replicators. Proc. Natl. Acad. Sci. USA. *92*, 2224–2228.

Remus, D., Beall, E. L., and Botchan, M. R. (2004). DNA topology, not DNA sequence, is a critical determinant for Drosophila ORC-DNA binding. EMBO J. *23*, 897–907.

Robinson, N. P., and Bell, S. D. (2005). Origins of DNA replication in the three domains of life. FEBS J. *272*, 3757–3766.

Robinson, N. P., Dionne, I., Lundgren, M., Marsh, V. L., Bernander, R., and Bell, S. D. (2004). Identification of two origins of replication in the single chromosome of the archaeon *Sulfolobus solfataricus*. Cell *116*, 25–38.

Salzberg, S. L., Salzberg, A. J., Kerlavage, A. R., and Tomb, J.-F. (1998). Skewed oligomers and origins of replication. Gene *217*, 57–67.

Sanders, C. M., and Stenlund, A. (1998). Recruitment and loading of the E1 initiator protein: an ATP-dependent process catalysed by a transcription factor. EMBO J. *17*, 7044–7055.

Sasaki, T., Sawado, T., Yamaguchi, M., and Shinomiya, T. (1999). Specification of regions of DNA replication initiation during embryogenesis in the 65-Kilobase DNApolα-dE2F locus of *Drosophila melanogaster*. Mol. Cell. Biol. *19*, 547–555.

Schaarschmidt, D., Ladenburger, E. –M., Keller, C., and Knippers, R. (2002). Human Mcm proteins at a replication origin during the G1 to S phase transition. Nucleic Acids Res. *30*, 4176–4185.

Schepers, A., Ritzi, M., Bousset, K., Kremmer, E., Yates, J. L., Harwood, J., Diffley, J. F. X., and Hammerschmidt, W. (2001). Human origin recognition complex binds to the region of the latent origin of DNA replication of Epstein-Barr virus. EMBO J. *20*, 4588–4602.

Schreiber, E., Matthias, P., Müller, M. M., and Schaffner, W. (1989). Rapid detection of octamer binding proteins with "mini-extracts" prepared from a small number of cells. Nucleic Acids Res. *17*, 6419.

Seufert, W., and Messer, W. (1987). Start sites for bidirectional *in vitro* DNA replication inside the replication origin, oriC, of *Escherichia coli*. EMBO J. *6*, 2469–2472.

Shinomiya, T., and Ina, S. (1993). DNA replication of histone gene repeats in *Drosophila melanogaster* tissue culture cells: multiple initiation sites and replication pause sites. Mol. Cell. Biol. *13*, 4098–4106.

Shinomiya, T., and Ina, S. (1994). Mapping an initiation region of DNA replication at a single-copy chromosomal locus in *Drosophila melanogaster* cells by two-dimensional gel methods and PCR-mediated nascent-strand analysis: multiple replication origins in a broad zone. Mol. Cell. Biol. *14*, 7394–7403.

Siddiqui, K., and Stillman, B. (2007). ATP-dependent assembly of the human origin recognition complex. J. Biol. Chem. *282*, 32370–32383.

Simpson, R. T. (1990). Nucleosome positioning can affect the function of a cis-acting DNA element *in vivo*. Nature *343*, 387–389.

Sivaprasad, U., Machida, Y. J., and Dutta, A. (2007). APC/C – the master controller of origin licensing? Cell Div. *2*, 8.

Stefanovic, D., Stanojcic, S., Vindigni, A., Ochem, A., and Falaschi, A. (2003). *In vitro* protein-DNA interactions at the human Lamin B2 replication origin. J. Biol. Chem. *278*, 42737–42743.

Stinchcomb, D. T., Struhl, K., and Davis, R. W. (1979). Isolation and characterization of a yeast chromosomal replicator. Nature *282*, 39–43.

Takahashi, T. S., Wigley, D. B., and Walter, J. C. (2005). Pumps, paradoxes and ploughshares: mechanism of the MCM2-7 DNA helicase. Trends. Bioch. Sci. *30*, 437–444.

Takahashi, T., Ohara, E., Nishitani, H., and Masukata, H. (2003). Multiple ORC-binding sites are required for efficient MCM loading and origin firing in fission yeast. EMBO J. *22*, 964–974.

Tao, L., Dong, Z., Leffak, M., Zannis-Hadjopoulos, M., and Price, G. (2000). Major DNA replication initiation sites in the *c-myc* locus in human cells. J. Cell. Biochem. *78*, 442–457.

Theis, J. F., and Newlon, C. S. (1994). Domain B of ARS307 contains two functional elements and contributes to chromosomal replication origin function. Mol. Cell. Biol. *14*, 7652–7659.

Thomae, A. W., Pich, D., Brocher, J., Spindler, M. –P., Berens, C., Hock, R., Hammerschmidt, W., and Schepers, A. (2008). Interaction between HMGA1a and the origin recognition complex creates site-specific replication origins. Proc. Natl. Acad. Sci. USA. *105*, 1692–1697.

Todorovic, V., Giadrossi, S., Pelizon, C., Mendoza-Maldonado, R., Masai, H., and Giacca, M. (2005). Human origins of DNA replication selected from a library of nascent DNA. Mol. Cell *19*, 567–575.

Touchon, M., Nicolay, S., Audit, B., Brodie of Brodie, E. B., d'Aubenton-Carafa, Y., Arneodo, A., and Thermes, C. (2005). Replication-associated strand asymmetries in mammalian genomes: toward detection of replication origins. Proc. Natl. Acad. Sci. USA. *102*, 9836–9841.

Tribioli, C., Biamonti, G., Giacca, M., Colonna, M., Riva, S., and Falaschi, A. (1987). Characterization of human DNA sequences synthesized at the onset of S-phase. Nucleic Acids Res. *15*, 10211–10232.

Trivedi, A., Waltz, S. E., Kamath, S., and Leffak, M. (1998). Multiple initiations in the *c-myc* replication origin independent of chromosomal location. DNA Cell Biol. *17*, 885–896.

Vashee, S., Cvetic, C., Lu, W., Simancek, P., Kelly, T. J., and Walter, J. C. (2003). Sequence-independent DNA binding and replication initiation by the human origin recognition complex. Genes Dev. *17*, 1894–1908.

Vassilev, L., and Johnson, E. M. (1990). An initiation zone of chromosomal DNA replication located upstream of the *c-myc* gene in proliferating HeLa cells. Mol. Cell. Biol. *10*, 4899–4904.

Waltz, S. E., Trivedi, A. A., and Leffak, M. (1996). DNA replication initiates non-randomly at multiple sites near the *c-myc* gene in HeLa cells. Nucleic Acids Res. *24*, 1887–1894.

Wohlschlegel, J. A., Dwyer, B. T., Dhar, S. K., Cvetic, C., Walter, J. C., and Dutta, A. (2000). Inhibition of eukaryotic DNA replication by geminin binding to Cdt1. Science *290*, 2309–2312.

Wu, X., and Lee, H. (2002). Human Dbf4/ASK promoter is activated through the Sp1 and *Mlu*I-cell-cycle box (MCB) transcription elements. Oncogene *21*, 7786–7796.

Wyrick, J.J., Aparicio, J.G., Chen, T., Barnett, J.D., Jennings, E.G., Young, R. S., Bell, S. P., and Aparicio, O. M. (2001). ORC and MCM proteins in *S. cerevisiae*: high-resolution mapping of replication origins. Science *294*, 2357–2360.

Yamada, M., Sato, N., Taniyama, C., Ohtani, K., Arai, K., and Masai, H. (2002). A 63-base pair DNA segment containing an Sp1 site but not a canonical E2F site can confer growth-dependent and E2F-mediated transcriptional stimulation of the human ASK gene encoding the regulatory subunit for human Cdc7-related kinase. J. Biol. Chem. *277*, 27668–27681.

Yan, J., Xu, L., Crawford, G., Wang, Z., and Burgess, S. M. (2006). The forkhead transcription factor FoxI1 remains bound to mitotic chromosomes and stably remodels chromatin structure. Mol. Cell. Biol. *26*, 155–168.

Yompakdee, C., and Huberman, J. A. (2004). Enforcement of late replication origin firing by clusters of short G-rich DNA sequences. J. Biol.Chem. *279*, 42337–42344.

Yoon, Y., Sanchez, J. A., Brun, C., and Huberman, J. A. (1995). Mapping of replication initiation sites in human ribosomal DNA by nascent-strand abundance analysis. Mol. Cell. Biol. *15*, 2482–2489.

You, Z., Komamura, Y., and Ishimi, Y. (1999). Biochemical analysis of the intrinsic Mcm4-Mcm6-Mcm7 DNA helicase activity. Mol. Cell. Biol. *19*, 8003–8015.

Zhang, H., and Tower, J. (2004). Sequence requirements for function of *Drosophila* chorion gene locus ACE3 replicator and ori-β origin elements. Development *131*, 2089–2099.

Zhu, J., Carlson, D. L., Dubey, D. D., Sharma, K., and Huberman, J. A. (1994). Comparison of the two major ARS elements of the ura4 replication origin region with other ARS elements in the fission yeast, *Schizosaccharomyces pombe*. Chromosoma *103*, 414–422.

Zou, L., and Stillman, B. (2000). Assembly of a complex containing Cdc45p, replication protein A, and Mcm2p at replication origins controlled by S-phase cyclin-dependent kinases and Cdc7-Dbf4p kinase. Mol. Cell. Biol. *20*, 3086–3096.

Appendices

Appendix I. Primers used for Q-PCR analysis.

All primer pairs were designed using Primer Express (Applied Biosystems). Dissociation curve analysis was performed after each PCR reaction to confirm the presence of a single amplification product. Amplicon size was also verified by gel electrophoresis.

(*) Unless otherwise indicated, position is based on the sequence of BAC clone CTB-60N22 (Accession #AC003083). Numbers with (§), (#), (Ψ) and (~) refer to BAC clone CTB-135C18 (Accession #AC005164), GenBank file M94363, GenBank file X00364, and GenBank file NT_008583.16, respectively. "–" and "+" denote up- and downstream of the DBF4 ATG start-codon, respectively. N/A, not applicable. The *c-myc* locus primer pair utilized in the DNase I-sensitivity experiments is marked "DNase I".

Primer	Sequence (5'-3')	Position (*)	Position (to ATG)	Tm (°C)	Amplicon size (bp)
Prom3 F	ATAAGGACCACGCAACCCAG	103965-103946	-31254 to -31273	58.3	95
Prom3 R	GCTCCACCTCCTGTCAGATCA	103871-103891	-31348 to -31328	59.5	
Prom2 F	CACACCACCCAGTGGGTACC	125543-125524	-9676 to -9695	60.8	86
Prom2 R	CCCTGGATGCGCCTTTTAA	125458-125476	-9761 to -9743	56.5	
Prom6 F	TTTTATTCATTGCTGGCCCTG	132279-132259	-2940 to -2960	53.8	54
Prom6 R	ACATTTGCCCATGGTCACAA	132226-132245	-2993 to -2974	55.3	
AT-2F	GAGATGTTACTTTCCACTCC	134138-134157	-1081 to -1062	50.1	135
AT-2R	GCCTCCAGGTCACTTTGT	134253-134272	-966 to -947	54.9	
Prom7B F	CCTTCTGCCCCAGGTATGC	134594-134576	-625 to -643	58.3	51
Prom7B R	AGCTTCCCAGACGTCCGAC	134544-134562	-675 to -657	59.5	
Prom7C2 F	AGTCCTGGCAAACCTAGAAG	134611-134630	-608 to -589	59.9	80
Prom7C2 R	ACCTAGTGGATGCAGTTAGC	134671-134690	-548 to -529	58.4	
Prom4 F	CTGCAGGTTGTGTTTCCGC	134821-134803	-398 to -416	57.3	70
Prom4 R	GAGAGCCCGGGTCTGAGAC	134752-134770	-467 to -449	59.7	
Prom8 F	TCTCTCCTTCCGCCAGCTAC	135119-135100	-100 to -119	58.6	52
Prom8 R	TGCAGACGCGGTACCTCTACT	135068-135088	-151 to -131	60.0	
HS(A) F	TCCTTCCTTCAGACCCCAGAG	135381-135401	+162 to +182	58.1	52
HS(A) R	CGGGTCCTCAGCTTCTTTTCT	135350-135370	+131 to +151	57.0	
CodR2 F	CCAAAAATCCACCCACAATCA	135683-135703	+464 to +484	53.7	52
CodR2 R	TCGTGGCTTTTCGCTTCAG	135652-135670	+433 to +451	55.9	
Cod5 F	CACACAACTGGAACACGGCTT	3433-3453 (§)	+4839 to +4859	58.3	52
Cod5 R	TCCCACATGCCCAAATCTTC	3484-3465 (§)	+4890 to +4871	55.4	
LaminB2 F	CTCTTCCAAGCCCTGCGTC	3770-3788 (#)	N/A	58.5	51
LaminB2 R	CGTTTTTGCAGGTTGTGCTGT	3800-3820 (#)	N/A	57.1	
c-Myc F	TTCCGCCTGCGATGATTTATAC	1982-2003 (Ψ)	N/A	55.2	52
c-Myc R	ACTGTTTGACAAACCGCATCC	2015-2035 (Ψ)	N/A	56.2	
c-Myc F (DNase I)	CGATGCATTTTTTGTGCATGA	1073-1093 (Ψ)	N/A	52.7	54
c-Myc R (DNase I)	CGGACAAACCGGACGTTTAA	1157-1138 (Ψ)	N/A	55.5	
Ex2 F	TGATAACAGGCCAGAAAAATCCA	136594-136616	+1376 to +1398	54.6	51
Ex2 R	AATACTTTTCCCCAAAGTGGCTT	136622-136644	+1405 to +1427	55.1	
Chr10 UPS F	CCCCATTGGCAACTTTTGAA	14482226-14482245 (~)	-479 to -460	54.1	54
Chr10 UPS R	AACGCTCACTTGCCAGGAA	14482261-14482279 (~)	-513 to -495	57.3	
Chr10 Cont F	CAAAGCCCCAACCTTTGATC	14474466-14474485 (~)	+7281 to +7300	54.4	51
Chr10 Cont R	AAGGGCTGATTGGAAATTGCT	14474496-14474516 (~)	+7250 to +7270	55.3	

Appendix II. Primers used for detection of the *DBF4* pseudogene.

Primer	Sequence (5'-3')	Tm (°C)
I	TGCAGACGCGGTACCTCTACT	60.0
II	TGGTGACAGAAGGTAAGTCAAGGT	57.4
III	AACTTCCATCATGGCTGGGATGAG	60.2

Appendix III. Primers used to synthesize DNA fragments for *in vitro* binding assays. All forward primers (marked with an "F") were labeled on their 5'-end with biotin (IDLabs Biotechnology). (*) Unless otherwise indicated, position is based on the nucleotide sequence of BAC clone CTB-60N22 (Accession #AC003083). (§) Numbering is based on the sequence of human chromosome 19 contig NT_011255. "-" and "+" denote up- and downstream of the DBF4 ATG start-codon, respectively. N/A, not applicable. The size of the expected products is shown in the last column.

Primer	Sequence (5'-3')	Position (*)	Position (to ATG)	Amplicon (bp)
ORI F	TGGCAAACCTAGAAGGCGGGAATA	134616-134639	-603 to -580	827
ORI R	TCAGGATCCGTTCCTTCTTAGGCA	135442-135419	+224 to +201	
Control F	AGTGAGCTGAGATCATGCCACTGT	124510-124533	-10709 to -10686	813
Control R	ACTCGTGAGTAGCAGTGGTGGTTT	125299-125322	-9920 to -9897	
I F	TCTTGTCACAACCCACCCTTGTCT	133852-133875	-1367 to -1344	788
I R	TATTCCCGCCTTCTAGGTTTGCCA	134639-134616	-580 to -603	
II F	TAGCAAACAATGTGCAGATCCGGG	135518-135541	+300 to +323	795
II R	TCGTGTGTTTATTCCTGAGGCACG	136312-136289	+1094 to +1071	
LB2 F	CAAAAACGGAGCTGGGCTCAGCTG	2368334-2368310 (§)	N/A	240
LB2 R	GACATCCGCTTCATTAGGGCAGAGGCC	2368094-2368120 (§)	N/A	
B13 F	CCTCAGAACCCAGCTGTGGA	2363655-2363636 (§)	N/A	166
B13 R	GCCAGCTGGGTGGTGATAGA	2363490-2363509 (§)	N/A	

Appendix IV. Oligonucleotides used for RIP mapping. All primers were labeled with digoxigenin at their 5'-end (ID Labs Biotechnology). (*) Position is based on the nucleotide sequence of BAC clone CTB-60N22 (Accession #AC003083). "–" and "+" denote up- and downstream of the DBF4 ATG start-codon, respectively. N/A, not applicable. C1 corresponds to the primer used for RIP mapping of the human lamin B2 ori. The name and sequence of this primer are the same as reported by Dimitrova et al. (1996).

Primer	Sequence (5'-3')	Position (*)	Position (to ATG)	Tm (°C)	GC-content	Annealing temp (°C)
pA	CTTCCGCCAGCTACGGCCTC	135113–135094	–106 to –125	63.0	70.0%	68
pB	CAGCAAGCGGGGTTTTCCCG	134423–134404	–796 to –815	62.0	65.0%	68
pC	GCGTAGAGGCCGTAGCTGGC	135089–135108	–130 to –111	62.9	70.0%	68
pD	AGCTTCCCAGACGTCCGAC	134544–134562	–675 to –657	59.5	63.2%	64
pY	GCCACGAGCGCGTCTGACTT	135658–135639	+440 to +421	62.8	65.0%	68
C1	TCGCATCACGTGACGAAGAGTCAGC	N/A	N/A	62.7	56.0%	64

Appendix V. Summary of RIPs shown in Figure 15b. The letters in red are precise replication initiation points. The complementary DNA sequences (bottom) are the template DNA sequence of the newly replicated leading strand DNA. "H," "M" and "L" denote RIPs used with high, medium and low frequencies, respectively.

(a) RIPs determined using pA

Nucleotide sequence	Position (to ATG)	Position (relative to BAC clone #AC003083)
gcggCgccg cgccGcggc	-264 (M)	134955
ccgaTacgc ggctAtgcg	-270 (L)	134949
gcgcCccgc cgcgGggcg	-279 (M)	134940
ccgcCccgt ggcgGggca	-284 (L)	134935
cgccCcgtc gcggGgcag	-285 (L)	134934
tctcAcgcc agagTgcgg	-292 (M)	134927
ccccGcctc ggggCggag	-311 (L)	134908
ccgcCtcgc ggcgGagcg	-313 (L)	134906
ctcgCccct gagcGggga	-323 (H)	134896
cgccCctcg gcggGgtgc	-325 (H)	134894
gcccCtcgt cgggGagca	-326 (H)	134893
cgccCcgcc gcggGgcgg	-355 (L)	134864
cccgCcccc gggcGgggg	-358 (L)	134861
ccccCtcac ggggGagtg	-362 (M)	134857
ccttCtcgc ggaaGagcg	-379 (M)	134840
ctcgCgcct gagcGcgga	-383 (L)	134836
gcccTgcag cgggAcgtc	-399 (L)	134820
tgcaGgttg acgtCcaac	-403 (M)	134816
aggtTgtgt tccaAcaca	-406 (M)	134813
ggttGtgtt ccaaCacaa	-407 (M)	134812

gttgTgttt caacAcaaa	-408 (M)	134811
gtgtTtccg cacaAaggc	-411 (H)	134808
tccgCctct aggcGgaga	-416 (L)	134803
tctaCcccg agatGgggc	-422 (L)	134797
ccccGcctc ggggCggag	-426 (M)	134793
tctcAgacc agagTctgg	-454 (M)	134765
acccGggct tgggCccga	-460 (H)	134759
cccgGgctc gggcCcgag	-461 (H)	134758
ccggGctct ggccCgaga	-462 (H)	134757
tctcGgtcc agagCcagg	-468 (M)	134751
gccgCttca cggcGaagt	-477 (M)	134742
ccgcTtcag ggcgAagtc	-478 (M)	134741
aggtCttgg tccaGaacc	-485 (L)	134734
tggcGcagc accgCgtcg	-491 (L)	134728
ggcgCagcc ccgcGtcgg	-492 (L)	134727
cctcAgaga ggagTctct	-499 (M)	134720
agagTtggc tctcAaccg	-505 (L)	134714
gagtTggcg ctcaAccgc	-506 (L)	134713
tctgTgttg agacAcaac	-519 (L)	134700
gtgtTgacc cacaActgg	-522 (L)	134697
gttgAccaa caacTggtt	-524 (L)	134695
accaAacct tggtTtgga	-528 (L)	134691
ggagCccgg cctcGggcc	-556 (M)	134663
gcccCgccc cgggGcggg	-564 (L)	134655
caggGttat gtccCaata	-578 (L)	134641
ggttTgcca ccaaAcggt	-599 (L)	134620
ctgcAgctg gacgTcgac	-620 (M)	134599
aggtAtgct	-640 (L)	134579

tccaTacga		
ggcgAgggg ccgcTcccc	-683 (M)	134536

pA: 49 sites in total ----- C, 20/49 (41%); T, 11/49 (22%); G, 10/49 (20%); A, 8/49 (16%)

(b) RIPs determined using pC

Nucleotide sequence	Position (relative to ATG)	Position (relative to BAC clone #AC003083)
caatGtgca gttaCacgt	+311	135529
caatgTgca gttacAcgt	+312	135530
gatcCggga ctagGccct	+320	135538

Appendix VI

Published journal articles

nature
structural &
molecular biology

Asymmetric bidirectional replication at the human *DBF4* origin

Julia Romero[1,2] & Hoyun Lee[1–3]

Faithful replication of the entire genome once per cell cycle is essential for maintaining genetic integrity, and the origin of DNA replication is key in this regulation. Unlike that in unicellular organisms, the replication initiation mechanism in mammalian cells is not well understood. We have identified a strong origin of replication at the *DBF4* promoter locus, which contains two initiation zones, two origin recognition complex (ORC) binding sites and two DNase I–hypersensitive regions within ~1.5 kb. Notably, similar to the *Escherichia coli oriC*, replication at the *DBF4* locus starts from initiation zone I, which contains an ORC-binding site, and progresses in the direction of transcription toward initiation zone II, located ~0.4 kb downstream. Replication on the opposite strand from zone II, which contains another ORC-binding site, may be activated or facilitated by replication from zone I. We term this new mammalian replication mode 'asymmetric bidirectional replication'.

DNA replication in prokaryotes and lower eukaryotes usually initiates from a discrete *cis*-acting DNA element termed the origin of DNA replication (ori)[1,2]. There are lines of evidence that replication in the mammalian cells also starts from a defined location called an origin of bidirectional replication (OBR)[2–5]. However, there is equally compelling evidence that replication initiates from a broad initiation zone[6–9]. It is now generally accepted that several potential ori sites exist within a small or large initiation zone, although certain sites within the zone may be preferentially used[10–12]. However, there is no experimental data that can reconcile the OBR and diffused/delocalized ori models. Even in the case of known mammalian OBRs, the exact mode of replication is still to be demonstrated, since the resolution of all known OBRs is several hundreds to thousands of nucleotides range, except the one in the human lamin-B2 (*LMNB2*) locus[5].

Although ori sites of mammalian cells apparently do not contain a common motif or consensus sequence, several pronounced features are often found at replication initiation loci[13]. These include asymmetric AT stretches[14], CpG islands[15] and transcriptional control elements[16–19]. As the first two features are also frequently found at transcription promoter regions, there is a close interplay between regulation of transcription and initiation of DNA replication.

We have been studying DNA replication using the *DBF4* locus as a model. Dbf4 is the regulatory partner of the Cdc7 serine/threonine protein kinase, which is essential for the activation of individual ori sites throughout S phase in eukaryotic cells[20–22]. After the assembly of pre-replication complexes onto an ori, Cdc7-Dbf4 kinase activates initiation of DNA replication in conjunction with S phase–specific cyclin-dependent kinases[21]. Cdc7-Dbf4 activity is tightly regulated as part of the many mechanisms that cells have in place to ensure that DNA replication occurs at the right time and only once per cell cycle[23]. As the abundance of the Cdc7 catalytic subunit remains largely constant throughout the cell cycle, the regulation of Cdc7-Dbf4 activity is thought to occur primarily through modulation of the Dbf4 regulatory subunit[21,23].

To understand how Dbf4 is regulated in mammalian cells, we and others previously identified and characterized the *DBF4* promoter[24,25]. We found that the 237-bp DNA segment from –211 to –447 (that is, upstream of the translation start-codon numbered as +1) has nearly full promoter activity. The *DBF4* promoter contains putative binding sites for several transcription factors, including a MluI cell-cycle box (MCB) and binding sites for Sp1 (four sites) and TFIIB. Among these elements, the MCB (–250 to –245) is essential for the basal promoter activity, and the Sp1 site at –361 to –353 (Sp1 box 3) is required for efficient promoter activity[24]. In addition, a putative enhancer is found within the 865-bp DNA segment from –1,628 to –994, in which there is a 184-bp DNA segment (–1,186 to –1,003) containing several A-T tracts[24] (**Supplementary Fig. 1** online). The *DBF4* locus also contains a CpG island (**Fig. 1a**). Of note, the putative mitochondrial carrier protein gene (*MCFP*), which is located ~20 kb upstream of the *DBF4* gene, may also use the *DBF4* promoter for its transcription[25]. The activity of the *DBF4* promoter increases substantially at the G1/S transition[26]. These features of the *DBF4* promoter led us to postulate that it may contain an ori with high activity in early S phase.

Here, we report the mapping of an ori and replication initiation points (RIPs) in the *DBF4* promoter region. Replication at this locus

[1]Department of Biochemistry, Microbiology and Immunology, Faculty of Medicine, University of Ottawa, 451 Smyth Road, Ottawa, Ontario K1H 8M5, Canada. [2]Tumour Biology Group, Northeastern Ontario Regional Cancer Centre, Sudbury Regional Hospital, 41 Ramsey Lake Road, Sudbury, Ontario P3E 5J1, Canada. [3]Department of Medical Sciences, the Northern Ontario School of Medicine, 935 Ramsey Lake Road, Sudbury, Ontario P3E 2C6, Canada. Correspondence should be addressed to H.L. (hlee@hrsrh.on.ca).

Received 12 March; accepted 6 May; published online 8 June 2008; doi:10.1038/nsmb.1439

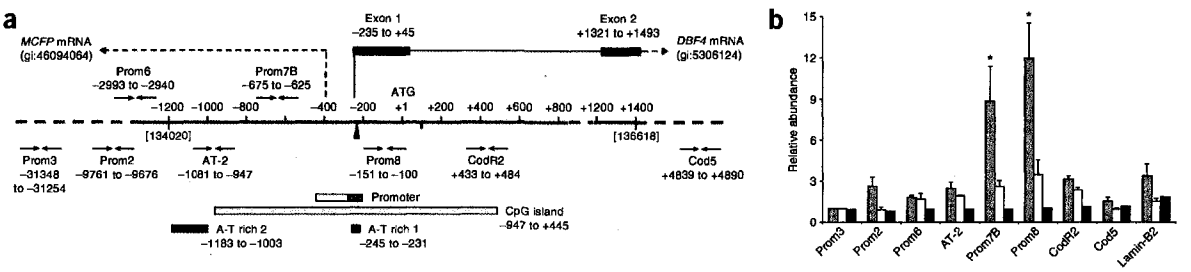


Figure 1 The *DBF4* promoter locus contains a strong ori. **(a)** Numbers are relative positions to the A (+1) of the translational start-codon. The numbering in brackets is according to the CTB-60N22 BAC clone. The *DBF4* major transcription initiation site[24] is shown as a vertically elongated triangle. The putative transcription initiation site for MCFP is shown as a broken line. Primers used to quantify nascent strands are shown as converging arrows (**Supplementary Table 1**). Gray and white in the *DBF4* promoter are the core and auxiliary elements, respectively. **(b)** *DBF4* ori is localized within the DNA segments corresponding to Prom7B and Prom8. The relative abundance of nascent DNA was quantified by Q-PCR and was normalized to the amount of nascent DNA amplified with Prom3. Lamin-B2 ori was included as positive control. Gray, white and black are nascent DNAs from asynchronous cells, nascent DNAs from mitotic cells and sonicated total DNA, respectively. Error bars, s.e.m. of at least three independent experiments. $*P < 0.01$.

initially starts from initiation zone I and then proceeds in the sense direction (that is, the direction of *DBF4* transcription) toward initiation zone II. The replication of antisense strand from initiation zone II may occur only when the replication on the sense strand has reached or passed through this initiation zone. This mode of replication in human cells is notably similar to that at the bacterial *oriC* locus, raising the possibility that asymmetric bidirectional replication may be common at ori sites of many organisms.

RESULTS

The *DBF4* promoter region contains a strong ori

To determine whether the *DBF4* promoter contains an ori, we carried out a nascent strand abundance assay[27]. To avoid DNA damage during nascent strand preparation, we directly loaded HeLa cells into a gel, where they were lysed. DNA was then separated by gel electrophoresis under denaturing conditions, followed by isolation of short nascent strands (1–2 kb) from the gel as described previously[28]. The analysis of nascent-strand abundance was carried out by quantitative PCR (Q-PCR) using primer pairs corresponding to a large part of the *DBF4* locus, including the promoter, putative enhancer and part of the coding region (**Fig. 1a** and **Supplementary Table 1** online). DNA segments corresponding to primers Prom7B (–675 to –625) and Prom8 (–151 to –100) were 8–10 times more abundant than adjacent regions (**Fig. 1b**). In contrast, no substantial enrichment was found for any sequence using mitotic cells or sonicated total genomic DNA (**Fig. 1b**). These data suggest that an ori exists within the 574-bp DNA segment from –675 to –100, which includes the principal transcription initiation site (–235) (ref. 24). We obtained similar results with HEK293 and HEK293T cells (data not shown). We found that initiation activities at the *DBF4* locus in HeLa cells were at least 2–3 times greater than those of the *MYC* (*c-myc*) (not shown) or lamin-B2 ori (**Fig. 1b**). In addition, data from synchronized HeLa cells showed that the *DBF4* ori fired extensively for a relatively short period in early S phase (7 h after mitosis or 1 h after G1/S arrest) (**Fig. 2**), providing an excellent model for studying mammalian DNA replication.

The *DBF4* ori contains two separate initiation zones

Precise replication initiation points (RIPs) at the human lamin-B2 locus have been determined using the technically challenging ligation-mediated (LM) PCR method[5]. To our knowledge, however, no other mammalian RIP has thus far been mapped by LM-PCR since that study was published in the year 2000. This is probably due, at least in part, to the difficulties of employing LM-PCR for RIP mapping.

A different research group successfully used a common one-way PCR-based primer extension method to map RIPs at the II/9A amplified ori of *Sciara coprophila*[29,30]. We therefore examined one-way PCR-based primer extension on the well characterized lamin-B2 ori in an attempt to develop technically a less challenging mammalian RIP mapping method. Using emetine-treated synchronized HeLa cells, we were able to detect a leading strand RIP of the 'bottom' strand at the expected position[5] (**Supplementary Fig. 2** online). Emetine was used to detect synthesis of the leading strand only, as it inhibits Okazaki fragment synthesis without affecting leading strand progression[31].

As the *DBF4* locus showed strong initiation activities (**Fig. 1b**), we determined leading strand RIPs in asynchronous HeLa cells using one-way PCR-based primer extension. Data from 1- to 2-kb nascent strand templates with primer A (pA, corresponding to –106 to –125; **Supplementary Table 2** online) suggested that approximate replication start-sites on the sense strand were –390, –400, –485, –510 and –565 (**Fig. 3a**). As we detected no primer extension product with pB (corresponding to –796 to –815), replication initiation may not occur upstream of –815 on the sense strand (**Fig. 3a**). Data obtained with pC (corresponding to –130 to –111) suggested that RIPs on the antisense strand were located at or around +130, +310, +330, +370 and +460 (**Fig. 3a**). Primer extension with pD (corresponding to –675 to –657) amplified a single ~800-nt strand, suggesting that replication initiated around +130 on the antisense strand (**Fig. 3b,c**). Other RIPs detected using pC were not detected with pD, presumably because the primer extension by one-way PCR was not optimal to amplify a long template. However, the main result of the primer extension with pD was that no RIP smaller than 800 nt was detected, even when smaller products were actively sought (**Fig. 3b**, right panel). Together, our data suggest that the initiation at the *DBF4* locus does not follow the conventional OBR model, since no initiation was observed on the opposite strand from RIPs detected with pA (**Fig. 3c**).

Next, we determined RIPs of synchronized HeLa cells in the presence (or absence, in some cases) of emetine (**Fig. 4a**). When pA was used for primer extension with a sample taken at 1 h after G1/S in the absence of emetine, a replication transitional point was detected at –235. (This position is calculated by adding –106 (pA position) and –129 (the approximate size of the band determined by the gel

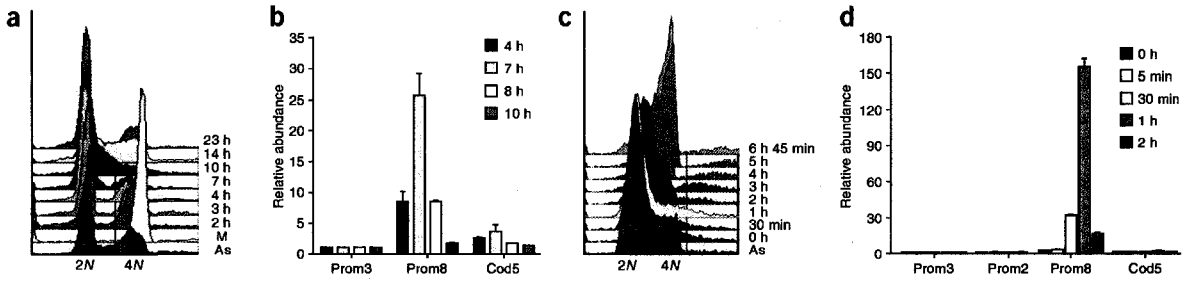


Figure 2 Replication initiation at the *DBF4* locus is largely confined to around 7 h after mitosis or 1 h after G1/S. **(a)** Flow cytometric profile of DNA content in postmitotic HeLa cells. M, cells arrested at mitosis by nocodazole. As, asynchronous cells; *N*, haploid DNA content. **(b)** Short nascent DNA corresponding to Prom8 is the most abundant at 7 h after mitosis. Nascent strands (1–2 kb) were isolated at the indicated time points after HeLa cells were released from nocodazole-based mitotic arrest. The relative abundance of DNA corresponding to each primer set was measured by Q-PCR. **(c)** Flow cytometric profile of post-G1/S HeLa cells. 0 h, cells arrested at the G1/S boundary by double-thymidine treatment. **(d)** Replication at the *DBF4* ori (corresponding to Prom8) is most active 1 h after G1/S.

electrophoresis)). In addition to the band corresponding to –235, many primer extension products were observed in this non-emetine-treated sample (**Fig. 4a**). This was expected, as Okazaki fragments would ligate to the leading strand DNA as replication continued. However, it seemed that some of the bands might not have been generated by ligation of Okazaki fragments because certain high-molecular-mass bands were unusually strong. Consistent with this interpretation, the sample treated with emetine also showed several primer-extension products. Together, our data suggested that leading strand synthesis on the sense strand started from several sites, including –235, –255, –260 and –390 (**Fig. 4a**, panel pA). Combining the RIP data from synchronous and asynchronous cells, we conclude that there are multiple initiation sites within the DNA segment amplified with pA. It should be noted that the products generated by one-way PCR-based primer extension were not the results of nonspecific amplification of randomly cleaved genomic DNA, as no product was generated from emetine-treated mitotic cells (**Fig. 4a**). Before the RIP assays, we also tested primers using cloned template DNA to ensure that polymerization was not prematurely terminated under our experimental conditions (data not shown).

Consistent with data from asynchronous cells, no RIP was detected with pB or pY (corresponding to +421 to +440) in synchronized cells, confirming that initiation does not occur upstream of –815 or downstream of –235 on the sense strand (**Fig. 4a**). Similarly, replication initiation on the antisense strand starts around +130, +265 and +290, but not upstream of +130 (**Fig. 4a**, panel pC).

To pinpoint the leading strand start-points, we determined precise RIPs using sequencing-gel electrophoresis. For this experiment, HeLa cells synchronized at the G1/S interface were released for 1 h into complete medium containing emetine. One-way PCR-based primer extension with 0.5–1.0-kb long leading-strand templates showed 49 RIPs on the sense strand, all of which clustered within a 420-bp segment from –264 to –683 (**Fig. 4b,c**; **Supplementary Table 3** online). Among these RIPs, initiation of DNA replication with a C residue was dominant (20/49, 41%), followed by T (11/49, 22%), G (10/49, 20%) and A (8/49, 16%). In a similar experiment, we found that replication initiation on the antisense strand started at positions +311, +312 and +320 (**Fig. 4b,c**).

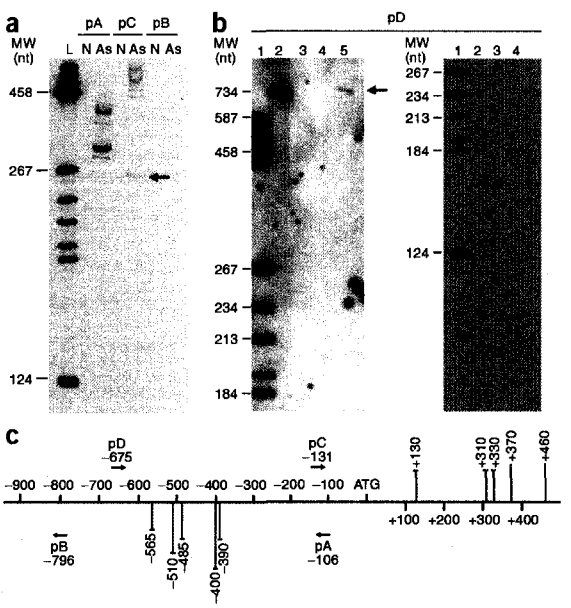


Figure 3 Determining RIPs at the *DBF4* locus in asynchronous HeLa cells. **(a)** Primer extension with the indicated digoxigenin-labeled primers (pA, pB and pC) using nascent DNA (1–2 kb) isolated from asynchronous HeLa cells that had been incubated for 1 h with emetine (As). L, digoxigenin-labeled molecular weight ladder; N, no template; arrow, the same RIP identified by both pC and pD (+130 position); MW, molecular weight. **(b)** Replication initiation on the antisense strand in asynchronous HeLa cells occurs far downstream from initiation zone I. Left panel, nascent strand templates (1–2 kb) were subjected to PCR-based primer extension with pD. The resultant products were separated by gel electrophoresis to resolve high-MW fragments. Lanes: 1, MW marker; 2, 5′-digoxigenin-labeled PCR product (734 nt); 3, no sample; 4, sham reaction control including primer pD but no template; 5, primer extension on nascent DNA templates (annealing temperature, 64 °C). Arrow, DNA amplified by primer extension with pD. The gel at right in **b** was designed to resolve smaller DNA fragments. Lanes: 1, MW ladder; 2, no sample; 3, sham reaction control; and 4, the same as in lane 5 of the left panel. **(c)** Summary of data in panels **a** and **b**. Note that pA and pB primers are on the sense strand and pC and pD are on the antisense strand.

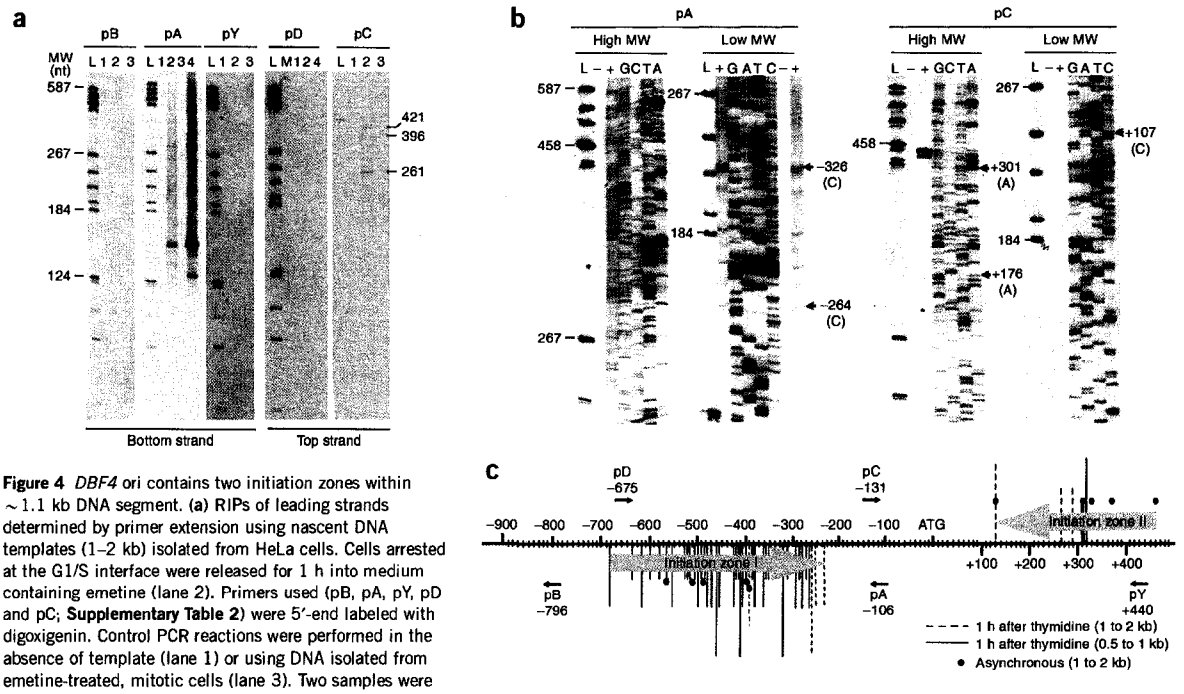


Figure 4 *DBF4* ori contains two initiation zones within ~1.1 kb DNA segment. **(a)** RIPs of leading strands determined by primer extension using nascent DNA templates (1–2 kb) isolated from HeLa cells. Cells arrested at the G1/S interface were released for 1 h into medium containing emetine (lane 2). Primers used (pB, pA, pY, pD and pC; **Supplementary Table 2**) were 5′-end labeled with digoxigenin. Control PCR reactions were performed in the absence of template (lane 1) or using DNA isolated from emetine-treated, mitotic cells (lane 3). Two samples were not treated with emetine (lane 4 in panels pA and pD). MW, molecular weight; L, digoxigenin-labeled MW ladder; M, digoxigenin-labeled 738-nt PCR product as an extra MW marker. **(b)** RIP assays performed with 0.5–1.0 kb nascent DNA (+) from cells at 1 h in S phase (see above) or in the absence of template (–). Arrows, some of RIPs and their first nucleotides. **(c)** Summary of initiation sites. Vertical dashed and solid lines, leading-strand DNA initiation sites as determined from data in **a** and **b**, respectively. The taller vertical lines are RIPs with stronger signals. Gray arrows, replication directions from initiation zones I (sense strand) and II (antisense strand). Black dots, the RIPs of asynchronous HeLa cells (from **Fig. 3a**).

Together, our data suggest that there are two initiation zones at the *DBF4* ori locus. The 449-nt-long initiation zone I spans from –235 to –683, within which at least 49 RIPs were detected. The 331-nt-long initiation zone II, on the antisense strand, stretches from +130 to +460 and contains only a few RIPs. Replication initiation was not detected within the 365-bp space between these two initiation zones (that is, from –235 to +130), suggesting that replication initiation seldom occurs within this DNA segment.

Asymmetric initiation of DNA replication at the *DBF4* locus

Because replication from initiation zones I and II occurs in opposite directions and progresses through the overlapping 'no-initiation' zone (**Fig. 4c**), replication initiation on the sense and antisense strands may not occur simultaneously. To gain insight into potential asymmetric replication, we carried out RIP assays with small (0.5–1.25 kb) and large (1.25–2.0 kb) nascent-strand templates isolated from cells synchronized at early S phase. RIPs were detected from the smaller templates amplified with either pA or pC (**Fig. 5a**). Primer extension with pA also showed several RIPs from the larger templates. In contrast, no RIPs were detected with pC from the larger templates, raising the possibility that DNA synthesis on the sense strand may initiate earlier than that on the antisense strand (**Fig. 5a**).

To further examine the possibility of asymmetric replication, we conducted primer extension using short single-stranded DNA templates (0.5–1.0 kb) isolated from HeLa cells that had been released from G1/S arrest for 1 or 2 h in the presence of low concentrations of aphidicolin (and emetine). Because aphidicolin can effectively slow down DNA chain elongation (but not initiation) at a low concentration, it has been used to determine the initial direction of DNA synthesis at the SV40 ori locus[32]. At 0.15 μg ml^{-1} aphidicolin, we

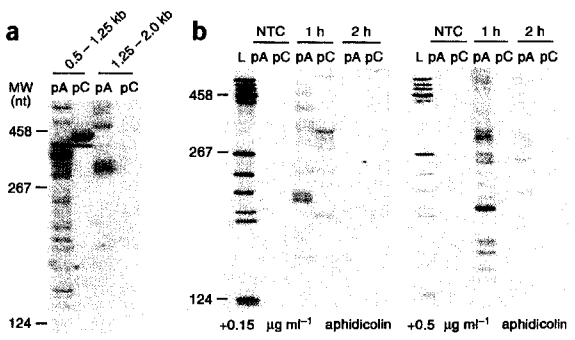


Figure 5 Replication initiation occurs from replication zone I before zone II. **(a)** At 1 h after G1/S, much more abundant RIPs were detected from replication zone I than zone II. RIP analysis was carried out using short (0.5–1.25 kb) or long (1.25–2.0 kb) nascent DNA isolated from emetine-treated HeLa cells 1 h after G1/S. MW, molecular weight; L, MW ladder; pA and pC, the digoxigenin-labeled primers used for primer extension. **(b)** RIP assays carried out with 0.5–1.0-kb nascent DNA isolated from HeLa cells treated as in **a**, except that aphidicolin was added at the indicated concentrations. NTC, no template control.

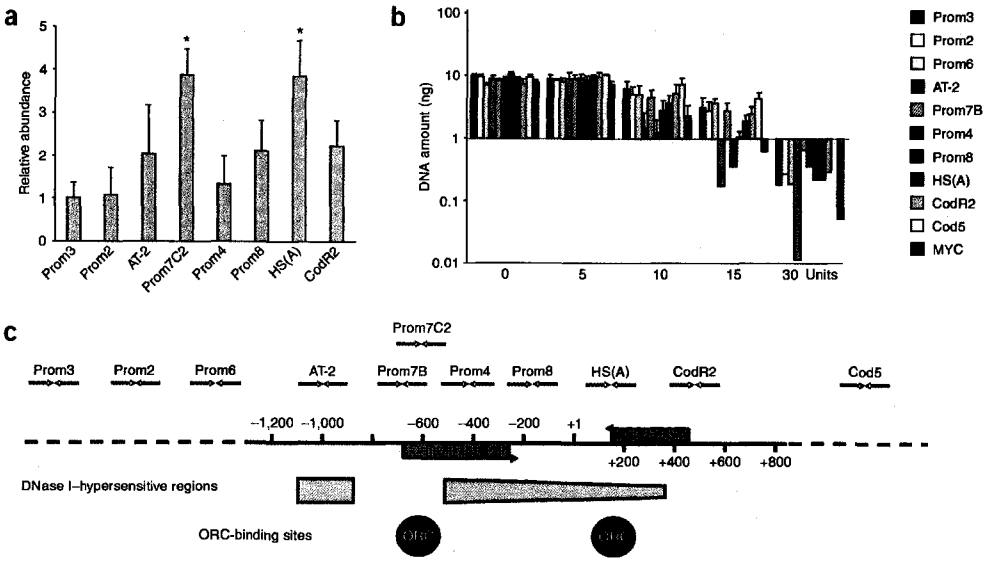


Figure 6 Mapping of the ORC-binding and DNase I–hypersensitive sites. (**a**) ORC binds preferentially to the segments corresponding to Prom7C2 and HS(A). The binding of ORC to DNA in asynchronous HeLa cells was determined by ChIP with antibodies to ORC4. The abundance of DNA segments in ChIP was determined by Q-PCR. The resultant values were standardized by subtracting the amount recovered with preimmune IgG. Relative abundance was calculated as percentage of input DNA and then normalized to the values obtained for Prom3 primer. $*P < 0.05$. (**b**) To map DNase I–hypersensitive sites, HeLa nuclei were treated with different concentrations of DNase I. Eight nanograms of DNA from the nuclei were quantified by Q-PCR with primers pairs indicated. Standard curves were generated for each primer pair using total, undigested genomic DNA to determine the amount of DNase I–resistant DNA. A nuclease-sensitive region of the *MYC* (*c-Myc*) promoter was a positive control[41]. Error bars, s.e.m. of two or more experiments. (**c**) Summary of the ORC-binding and DNase I–hypersensitive regions.

detected several RIPs using both pA and pC at 1 h after G1/S (**Fig. 5b**, left panel). However, we detected no RIPs with either primer by 2 h after G1/S, suggesting that all of the nascent DNAs initiated on both strands had progressed beyond detection (that is, >1.0 kb) by 2 h after G1/S. This conclusion is consistent with the data shown in **Figure 2d**. When replication forks were further slowed down by increasing the concentration of aphidicolin to 0.5 μg ml^{-1}, RIPs on the sense strands were detected at 1 and 2 h after G1/S with pA (**Fig. 5b**, right panel). In contrast, no appreciable replication activity was detected with pC at either 1 or 2 h after G1/S under the same conditions (**Fig. 5b**, right panel). These data suggest that replication from zone I on the sense strand occurs before initiation from zone II on the antisense strand.

ORC binds replication initiation zones I and II

Binding of the origin recognition complex (ORC) to an ori is a prerequisite for 'licensing' the initiation of DNA replication[2,33] by which the positions of replication initiation may be determined[29]. To begin examining whether *DBF4* ori is bound by ORC, we carried out chromatin immunoprecipitation (ChIP) with antibodies to ORC4 (**Supplementary Methods** online). DNA segments corresponding to Prom7C2 and HS(A) primer sets (**Supplementary Table 1**) were most abundantly bound to ORC (**Fig. 6a**). The abundance of flanking regions gradually decreased in both directions from the two peaks. These data therefore suggest that there are two ORC-binding sites: one at –608 to –529 (corresponding to Prom7C2) and the other at +131 to +182 (corresponding to HS(A)). The ORC-binding sites and other features of the *DBF4* ori/promoter are summarized in **Supplementary Figure 1**.

Two pronounced DNase I–hypersensitive regions in *DBF4* ori

To gain insight into the chromatin structure at the *DBF4* locus, we carried out DNase I sensitivity assays (**Fig. 6b**). The DNA segment corresponding to primer AT-2, but not two flanking regions (Prom6 and Prom7B), were DNase I hypersensitive, suggesting that the DNA segment spanning –1,081 to –947 (AT-2) has an 'open' chromatin structure. The segment corresponding to Prom4 (–467 to –398) within initiation zone I was also DNase I hypersensitive. The open chromatin structure at this locus extended to Prom8 (and to HS(A) to some extent); however, the openness gradually decreased as the distance from the Prom4 site increased further downstream (**Fig. 6b,c**). The downstream boundary of this DNase I–hypersensitive region approximately coincides with that of initiation zone II (**Fig. 6c**).

DISCUSSION

Identification of an ori at the *DBF4* promoter region

The existence of a strong ori at the *DBF4* promoter region was initially determined by a nascent-strand abundance assay and then confirmed by RIP mapping. Although the same locus was identified by these two independent methods, the boundary of the ori identified by the former is largely confined within initiation zone I (see below for discussion).

We found that one-way PCR-based primer extension[30] could be very sensitive and powerful in determining a RIP when combined with the following sample preparation methods: (i) lysis of mammalian cells directly in the gel to avoid DNA damage; (ii) isolation of nascent strands directly from gel after DNA separation by gel electrophoresis under denaturing conditions, which gives a high yield of short single-stranded DNAs; and (iii) treatment of cells with emetine so

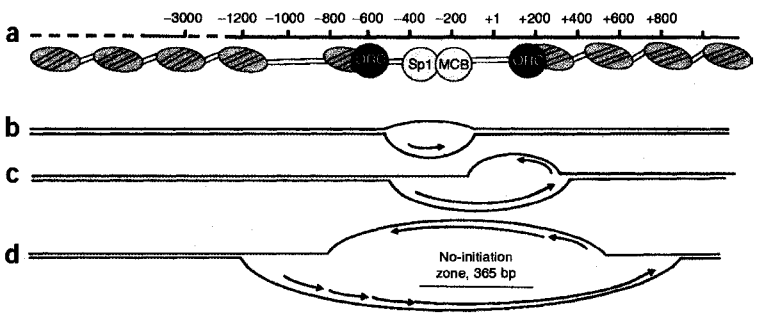


Figure 7 An ABR initiation model. **(a)** The numbers and map shown are as described in the **Figure 1** legend. The elongated circles and the associated double line are putative nucleosome locations and duplex DNA, respectively. The Sp1 binding site (–353/–361) and the MCB (–250/–245) were described previously[24]. ORC denotes ORC-binding sites. **(b–d)** The detailed ABR initiation mechanism is described in the text. Long and short arrows represent nascent leading and lagging strands, respectively.

that only leading strands are generated and isolated. Because PCR is routinely used in many laboratories, this PCR-based RIP mapping approach may greatly facilitate the study of mammalian DNA replication.

Identification of replication initiation zones I and II

One of the most notable features of the *DBF4* ori was that it contained two initiation zones, separated by ~400 bp no-initiation zone (**Fig. 4**). Many replication initiation sites were clustered within the 449-nt initiation zone I (–683 to –235), whereas only few initiation sites occurred within the 331-nt initiation zone II (+130 to +460). Although we cannot completely rule out the possibility that some of the leading-strand RIPs detected in the emetine-treated cells were generated by ligation of Okazaki fragments to the leading strand DNA, most of the RIPs, if not all, are likely to be genuine replication initiation sites. This interpretation is supported by the fact that certain bands in the non-emetine-treated sample showed very strong intensities. If these bands were generated by ligation of Okazaki fragments to the leading strand DNA, the signal intensities of the bands should have gradually decreased as the lengths of leading strand increased. However, we did not see this pattern.

The upstream boundary of initiation zone I approximately coincided with the ORC binding site (–608 to –529), and the downstream boundary overlapped the putative TFIIB- and MCB-binding sites (–267 to –261 and –250 to –245, respectively)[24]. The second ORC-binding site (+132 to +182) also roughly coincided with the upstream boundary of initiation zone II. Thus, ORC binding may determine the boundary of an initiation zone, rather than determining the precise positions of individual RIPs. We found that although specific nucleotides were generally used, slightly different initiation sites were observed between synchronized and asynchronous cell populations (summarized in **Fig. 4c**) as well as among different experiments (data not shown). However, the overall boundaries of the initiation zones never varied, suggesting that that chromatin context but not nucleotide sequence is essential for the activation of initiation in mammalian cells.

In early S phase, numerous RIPS were detected in the shorter (0.5–1.25 kb) nascent strands within initiation zone I, but only a few RIPs were found in the larger (1.25–2.0 kb) nascent strands (**Fig. 5a**). A similar trend was also found in the cells treated with low concentrations of aphidicolin to slow down replication fork movement (**Fig. 5b**). Together, these data raise the possibility that the half-life of the short leading strands in initiation zone I is unusually long, perhaps because of a transient pause of the replication fork (see below for discussion). This interpretation is supported by the fact that the *DBF4* ori identified by the nascent strand abundance assay was largely confined within initiation zone I.

A new model: asymmetric bidirectional replication at the *DBF4* locus

Initiation of DNA replication starts from initiation zone I more frequently and earlier than from zone II. This makes sense because it would be easier for the DNA to replicate, at least initially, in the same direction as that of vigorously transcribing genes such as *DBF4*. In addition, the chromatin structure at or around initiation zone I was more open than that around zone II, which could facilitate the assembly of the replication protein complexes at this region. Consistent with this expectation, initiation zone I contained a strong ORC-binding site. It seems possible that the open chromatin structure facilitates the binding of ORC to the upstream boundary of initiation zone I, allowing the chromatin to become competent for replication initiation (**Fig. 7a**). The chromatin in this region may become further competent for DNA replication by the binding of transcription factors at the downstream boundary of zone I (that is, binding of ORC at one end of initiation zone I and binding of transcription factors at the other end).

The Sp1 transcription element has been shown to positively affect DNA replication initiation in virus and yeast[34–36]. Our data **Figure 4c** is in line with these previous reports, potentially expanding the positive role of Sp1 on DNA replication to mammalian cells. In contrast to Sp1, MCB (and possibly TFIIB) may negatively affect replication initiation (and, thus, define the downstream boundary of initiation zone I). It may be possible that the MCB transcription factor forces the replication fork movement to stall transiently while MCB 'negotiates' with the replication machinery before letting it pass through the downstream boundary. The initiation of DNA replication on the antisense strand could be suppressed by transcription factors bound to the chromatin and/or high transcriptional activities in the opposite direction in early S phase. It is possible that DNA replication from initiation zone I can temporarily disrupt the transcriptional machinery bound at the promoter and, thus, provide a short 'window of opportunity' for antisense replication. It is therefore likely that the initiation on the antisense strand can occur only when the replication fork on the sense strand has reached or passed through initiation zone II (**Fig. 7b–d**). We term this newly identified mode of mammalian DNA replication asymmetric bidirectional replication (ABR).

OBR, diffused/delocalized initiation, and ori of ABR

An important question raised by our work is whether ABR is a general phenomenon or an exception. We cannot definitely answer this question until RIPs of many more mammalian ori sites have been characterized in details. It is, however, possible that many mammalian ori sites known at present (including some OBRs) may actually replicate asymmetrically. The maximum resolution of replicon-mapping techniques available at present, except RIP mapping, is 0.5 to 1.0 kb. At this low resolution, one cannot definitely conclude how many initiation sites or zones an ori contains, or whether replication

initiation occurs asymmetrically. When the ori at the *DBF4* locus was identified by a nascent strand abundance assay alone, we did not suspect asymmetric replication to occur at this locus. We realized the *DBF4* ori operated through ABR only when we systematically mapped the RIPs.

Only a single ORC-binding site is found at the lamin-B2 locus, where a single OBR operates. In contrast, the *DBF4* locus contains two initiation zones, each of which has ORC-binding site. Because ORC may determine the position of replication initiation sites[29] or initiation zones (this report), the number and position(s) of ORCs at the ori locus could be the determinant of the OBR or ABR replication mode. Overall, ABR replication at the *DBF4* does not extend to the lamin-B2 ori, instead showing a clearly different paradigm.

Although we have unequivocally shown in this paper that the *DBF4* locus contains two initiation zones, it could have been mistaken as a single (or two) OBR(s) under different assay conditions (for example, using only nascent-strand abundance assay, or using different synchronization time points and/or sizes of nascent strand DNA templates, and so forth). It is also possible that low-resolution replicon-mapping methods could identify the *DBF4* locus being a 'diffused or delocalized' initiation zone because replication bubbles could be detected within the entire 1.5 kb DNA segment from −1,000 to +500. Thus, the ABR model, at least in part, can explain and reconcile the OBR and the diffused or delocalized replication initiation models.

ABR in other organisms

Notably, the ABR replication pattern found at the *DBF4* locus is very similar to that at the *E. coli oriC* locus[37]. *In vivo* studies showed that leading strand replication in *E. coli* initiates at multiple sites within and around *oriC*. In addition, approximately 250 bp between the two initiation zones at the *oriC* locus is largely devoid of initiation[37]. Also similarly to the *DBF4* ori, the two initiation zones at the *oriC* locus seem to initiate in a temporally separable manner[37]. Although replication at *oriC* shows a more classical OBR pattern in the plasmid context[38], the mode of *oriC* replication in the chromosomal context is highly similar to that of *DBF4*. (This may suggest that the chromosomal context is important in determining initiation sites even in bacteria.) Furthermore, a similar mode of multiple initiations and asymmetric replication has also been found at ori sites of mouse mitochondria and the amplified II/9A of *S. coprophila*[29,39]. With these data taken together, ABR may be quite common in organisms, including mammals. In this context, we note that it has recently been hypothesized[7] that oriβ and β′ at the *Dhfr* locus in the Chinese hamster ovary cell lines may regulate replication similarly to *oriC*. In light of our findings, their hypothesis makes sense, especially if several sets of origins of ABR exist in the *Dhfr* locus.

METHODS

Nascent DNA isolation. We carried out separation and isolation of short single stranded nascent DNA by alkaline gel electrophoresis as described previously[28] with some modifications. Briefly, approximately 10^7 cells were collected by trypsinization, washed once with 1× PBS and then resuspended in 260 μl of 1× PBS containing 10% (v/v) glycerol. To prevent DNA damage during sample preparation, the cell suspension (85 μl) was directly loaded into a well of the 1.25% agarose gel containing 50 mM NaOH and 1 mM EDTA. The cells were then immediately lysed by adding 25 μl of 5× SDS lysis buffer (50 mM Tris HCl buffer, 5 mM EDTA, 0.5% (w/v) SDS, pH 8.0). After incubation for 10 min at room temperature, EDTA was added to a concentration of 20 mM and then NaOH to 50 mM. Electrophoresis was carried out for >8 h at 15 V at room temperature in freshly made alkaline running buffer (50 mM NaOH, 1 mM EDTA). After neutralization with 1× TAE (40 mM Tris, pH 8.0, 20 mM acetic acid, 1 mM EDTA) for 45 min, the lane containing DNA size markers was separated from the rest of the gel and stained with 1× TAE buffer containing 0.6 μg ml^{-1} ethidium bromide for 15 min. The size of the nascent DNA in the unstained gel was determined by comparing it with the ethidium bromide–stained DNA size marker. Subsequently, nascent DNA fragments were extracted from the gel using a QIAGEN gel-extraction kit.

Nascent strand abundance assay by Q-PCR. Q-PCR was carried out using a 7900HT ABI Prism (Applied Biosystems) in 25 μl Q-PCR solution containing 12.5 μl of 2× SYBR Green Master Mix (Applied Biosystems) and 300 nM of each primer (Integrated DNA Technologies). Standard curves were generated for each primer pair using serial dilutions of total genomic DNA. An equal volume of each nascent strand preparation was used for amplification. After initial denaturation for 10 min at 95 °C, amplification was carried out for 40 cycles as follows: 95 °C for 30 s, 57 °C for 30 s and then 72 °C for 30 s. Amplification of a single product at a correct DNA size by a given primer pair was verified by agarose gel electrophoresis. Relative abundance of nascent strands was estimated by calculating the ratio between the amount of each DNA segment (as estimated by the standard curves) and the amount of the DNA amplified by the Prom3 set (which consistently showed the lowest amplification).

RIP mapping. To analyze RIPs of leading strand synthesis, 2 μM emetine (Sigma) was added to the cell culture media for 1 h before DNA isolation. Where indicated, aphidicolin (Sigma) was also added. An equal volume of each DNA preparation (or an equal amount measured by Q-PCR when comparing different DNA preparations) was used as template in the primer extension reactions with 5′-digoxigenin-labeled primers. One-way PCR reactions were carried out in a final volume of 30 μl containing 1× PCR buffer (ID Labs Biotechnology), 200 μM dNTPs, 3 mM $MgSO_4$, 400 nM digoxigenin-labeled oligonucleotides (Integrated DNA Technologies), 1 M betaine (Sigma) and 1 unit of IDPol DNA polymerase (ID Labs Biotechnology). For the RIP mapping of the lamin-B2 ori, reactions were carried out in the absence of betaine. PCR conditions were as follows: initial incubation for 5 min at 95 °C, 30 cycles of amplification (1 min at 94 °C, 1 min at the appropriate annealing temperature as shown in **Supplementary Table 2**, and 1.5 min at 72 °C) and final extension for 7 min at 72 °C. As a negative control, a sample without template DNA was included in every experiment. After amplification, 20 μl of 'sequencing' loading buffer (98% (v/v) formamide, 10 mM EDTA, 0.1% (w/v) xylene cyanol, 0.1% (w/v) bromophenol blue) were added to each sample. Samples were then heated for 4 min at 90 °C, chilled quickly on ice and separated by PAGE in a gel containing 5% acrylamide and 7.7 M urea in 44.5 mM Tris, 44.5 mM boric acid, 1 mM EDTA, pH 8.4. After transfer of DNA onto a Hybond N^+ membrane (Amersham) for 1 h at 400 mA using a semi-dry transfer apparatus, the DNA was fixed with UV for 6 min. Detection of digoxigenin-labeled PCR products with an antibody to digoxigenin (Roche) was performed according to the manufacturer's instructions.

DNase I hypersensitivity assay. The isolation of nuclei and subsequent limited digestion with DNase I were carried out as described previously[40]. Briefly, asynchronously growing HeLa cells (~10^7) were trypsinized, resuspended in 1 ml RSB buffer (10 mM Tris, pH 7.5, 10 mM NaCl, 3 mM $MgCl_2$) containing 0.2% (v/v) NP-40, and incubated for 5 min on ice. Nuclei were isolated by centrifugation for 5 min at 500*g*. The nuclear pellet was resuspended in 600 μl of RSB buffer, and then divided into six 100-μl aliquots. To each aliquot, an appropriate amount of DNase I (Invitrogen) was added, and the aliquot was incubated for 5 min at 37 °C. The digestion reaction was stopped by adding an equal volume of Stop solution (1% (w/v) SDS, 600 mM NaCl, 10 mM EDTA, 20 mM Tris HCl, pH 8.0). Proteins in each sample were digested overnight with Proteinase K (Sigma), followed by phenol-chloroform extraction and ethanol precipitation. Finally, purified DNA was resuspended in 10 mM Tris HCl, pH 8.0, 1 mM EDTA, and quantified by spectrophotometry (Beckman DU-530).

Q-PCR was performed using 8 ng of DNA template for each sample. Standard curves were generated using serial dilutions of independently isolated, undigested total genomic DNA. Reaction mixture (25 μl) contained 12.5 μl of 2× SYBR Green Master Mix and 300 nM of each primer. After amplification, the amount of target DNA segment remaining in the digested samples was

determined using the standard curve. The results were compared to the amount detected in the control sample that was not treated with DNase I.

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

ACKNOWLEDGMENTS

We are grateful to S.-Y. Kim for her initial development of the ChIP assay protocol. This work was supported by grants from the Canadian Institutes of Health Research (MOP79473) to H.L. J.R. was supported in part by a graduate scholarship of the University of Ottawa.

AUTHOR CONTRIBUTIONS

J.R. planned, developed protocols, carried out experiments, analyzed data and drafted the manuscript; H.L. conceived and guided the overall research project, analyzed and interpreted data, and wrote the final version of the manuscript.

Published online at http://www.nature.com/nsmb/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Aladjem, M.I. & Fanning, E. The replicon revisited: an old model learns new tricks in metazoan chromosomes. *EMBO Rep.* **5**, 686–691 (2004).
2. Bell, S.P. & Dutta, A. DNA replication in eukaryotic cells. *Annu. Rev. Biochem.* **71**, 333–374 (2002).
3. Mechali, M. DNA replication origins: from sequence specificity to epigenetics. *Nat. Rev. Genet.* **2**, 640–645 (2001).
4. Burhans, W.C. *et al.* Identification of an origin of bidirectional DNA replication in mammalian chromosomes. *Cell* **62**, 955–965 (1990).
5. Abdurashidova, G. *et al.* Start sites of bidirectional DNA synthesis at the human lamin B2 origin. *Science* **287**, 2023–2026 (2000).
6. Vaughn, J.P., Dijkwel, P.A. & Hamlin, J.L. Replication initiates in a broad zone in the amplified CHO dihydrofolate reductase domain. *Cell* **61**, 1075–1087 (1990).
7. Dijkwel, P.A., Wang, S. & Hamlin, J.L. Initiation sites are distributed at frequent intervals in the Chinese hamster dihydrofolate reductase origin of replication but are used with very different efficiencies. *Mol. Cell. Biol.* **22**, 3053–3065 (2002).
8. Gencheva, M., Anachkova, B. & Russev, G. Mapping the sites of initiation of DNA replication in rat and human rRNA genes. *J. Biol. Chem.* **271**, 2608–2614 (1996).
9. Little, R.D., Platt, T.H. & Schildkraut, C.L. Initiation and termination of DNA replication in human rRNA genes. *Mol. Cell. Biol.* **13**, 6600–6613 (1993).
10. Dijkwel, P.A. & Hamlin, J.L. Sequence and context effects on origin function in mammalian cells. *J. Cell. Biochem.* **62**, 210–222 (1996).
11. Burhans, W.C. & Huberman, J.A. DNA replication origins in animal cells: a question of context? *Science* **263**, 639–640 (1994).
12. DePamphilis, M.L. Replication origins in metazoan chromosomes: fact or fiction? *Bioessays* **21**, 5–16 (1999).
13. Aladjem, M.I., Falaschi, A. & Kowalski, D. Eukaryotic DNA replication origins. in *DNA Replication and Human Disease* (ed. DePamphilis, M.L.) 31–61 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA, 2006).
14. Stefanovic, D. *et al.* In vitro protein-DNA interactions at the human lamin B2 replication origin. *J. Biol. Chem.* **278**, 42737–42743 (2003).
15. Delgado, S., Gomez, M., Bird, A. & Antequera, F. Initiation of DNA replication at CpG islands in mammalian chromosomes. *EMBO J.* **17**, 2426–2435 (1998).
16. Ladenburger, E.M., Keller, C. & Knippers, R. Identification of a binding region for human origin recognition complex proteins 1 and 2 that coincides with an origin of DNA replication. *Mol. Cell. Biol.* **22**, 1036–1048 (2002).
17. Beall, E.L. *et al.* Role for a *Drosophila* Myb-containing protein complex in site-specific DNA replication. *Nature* **420**, 833–837 (2002).
18. Bosco, G., Du, W. & Orr-Weaver, T.L. DNA replication control through interaction of E2F-RB and the origin recognition complex. *Nat. Cell Biol.* **3**, 289–295 (2001).
19. Abdurashidova, G. *et al.* Localization of proteins bound to a replication origin of human DNA along the cell cycle. *EMBO J.* **22**, 4294–4303 (2003).
20. Bousset, K. & Diffley, J.F. The Cdc7 protein kinase is required for origin firing during S phase. *Genes Dev.* **12**, 480–490 (1998).
21. Jiang, W., McDonald, D., Hope, T.J. & Hunter, T. Mammalian Cdc7-Dbf4 protein kinase complex is essential for initiation of DNA replication. *EMBO J.* **18**, 5703–5713 (1999).
22. Donaldson, A.D., Fangman, W.L. & Brewer, B.J. Cdc7 is required throughout the yeast S phase to activate replication origins. *Genes Dev.* **12**, 491–501 (1998).
23. Kumagai, H. *et al.* A novel growth- and cell cycle-regulated protein, ASK, activates human Cdc7-related kinase and is essential for G1/S transition in mammalian cells. *Mol. Cell. Biol.* **19**, 5083–5095 (1999).
24. Wu, X. & Lee, H. Human Dbf4/ASK promoter is activated through the Sp1 and MluI cell-cycle box (MCB) transcription elements. *Oncogene* **21**, 7786–7796 (2002).
25. Yamada, M. *et al.* A 63-base pair DNA segment containing an Sp1 site but not a canonical E2F site can confer growth-dependent and E2F-mediated transcriptional stimulation of the human ASK gene encoding the regulatory subunit for human Cdc7-related kinase. *J. Biol. Chem.* **277**, 27668–27681 (2002).
26. Guo, B. & Lee, H. Cloning and characterization of Chinese hamster homologue of yeast DBF4 (ChDBF4). *Gene* **264**, 249–256 (2001).
27. Giacca, M., Pelizon, C. & Falaschi, A. Mapping replication origins by quantifying relative abundance of nascent DNA strands using competitive polymerase chain reaction. *Methods* **13**, 301–312 (1997).
28. Kamath, S. & Leffak, M. Multiple sites of replication initiation in the human beta-globin gene locus. *Nucleic Acids Res.* **29**, 809–817 (2001).
29. Bielinsky, A.K. *et al.* Origin recognition complex binding to a metazoan replication origin. *Curr. Biol.* **11**, 1427–1431 (2001).
30. Gerbi, S.A. & Bielinsky, A.K. Replication initiation point mapping. *Methods* **13**, 271–280 (1997).
31. Burhans, W.C. *et al.* Emetine allows identification of origins of mammalian DNA replication by imbalanced DNA synthesis, not through conservative nucleosome segregation. *EMBO J.* **10**, 4351–4360 (1991).
32. Decker, R.S., Yamaguchi, M., Possenti, R. & DePamphilis, M.L. Initiation of simian virus 40 DNA replication in vitro: aphidicolin causes accumulation of early-replicating intermediates and allows determination of the initial direction of DNA synthesis. *Mol. Cell. Biol.* **6**, 3815–3825 (1986).
33. Bell, S.P. The origin recognition complex: from simple origins to complex functions. *Genes Dev.* **16**, 659–672 (2002).
34. Nguyen-Huynh, A.T. & Schaffer, P.A. Cellular transcription factors enhance herpes simplex virus type 1 oriS-dependent DNA replication. *J. Virol.* **72**, 3635–3645 (1998).
35. Li, R. Stimulation of DNA replication in *Saccharomyces cerevisiae* by a glutamine- and proline-rich transcriptional activation domain. *J. Biol. Chem.* **274**, 30310–30314 (1999).
36. Yan, P. *et al.* HBV C promoter Sp1 binding sequence functionally substitutes for the yeast ARS1 ABF1 binding site. *DNA Cell Biol.* **21**, 737–742 (2002).
37. Kohara, Y., Tohdoh, N., Jiang, X.W. & Okazaki, T. The distribution and properties of RNA primed initiation sites of DNA synthesis at the replication origin of *Escherichia coli* chromosome. *Nucleic Acids Res.* **13**, 6847–6866 (1985).
38. Seufert, W. & Messer, W. Start sites for bidirectional in vitro DNA replication inside the replication origin, oriC, of *Escherichia coli*. *EMBO J.* **6**, 2469–2472 (1987).
39. Brown, T.A. *et al.* Replication of mitochondrial DNA occurs by strand displacement with alternative light-strand origins, not via a strand-coupled mechanism. *Genes Dev.* **19**, 2466–2476 (2005).
40. Yan, J. *et al.* The forkhead transcription factor Foxl1 remains bound to condensed mitotic chromosomes and stably remodels chromatin structure. *Mol. Cell. Biol.* **26**, 155–168 (2006).
41. Kumar, S. & Leffak, M. Conserved chromatin structure in c-myc 5'flanking DNA after viral transduction. *J. Mol. Biol.* **222**, 45–57 (1991).

ONLINE SUPPLEMENTARY DATA

Asymmetric bidirectional replication (ABR) at the human DBF4 origin

Julia Romero and Hoyun Lee*

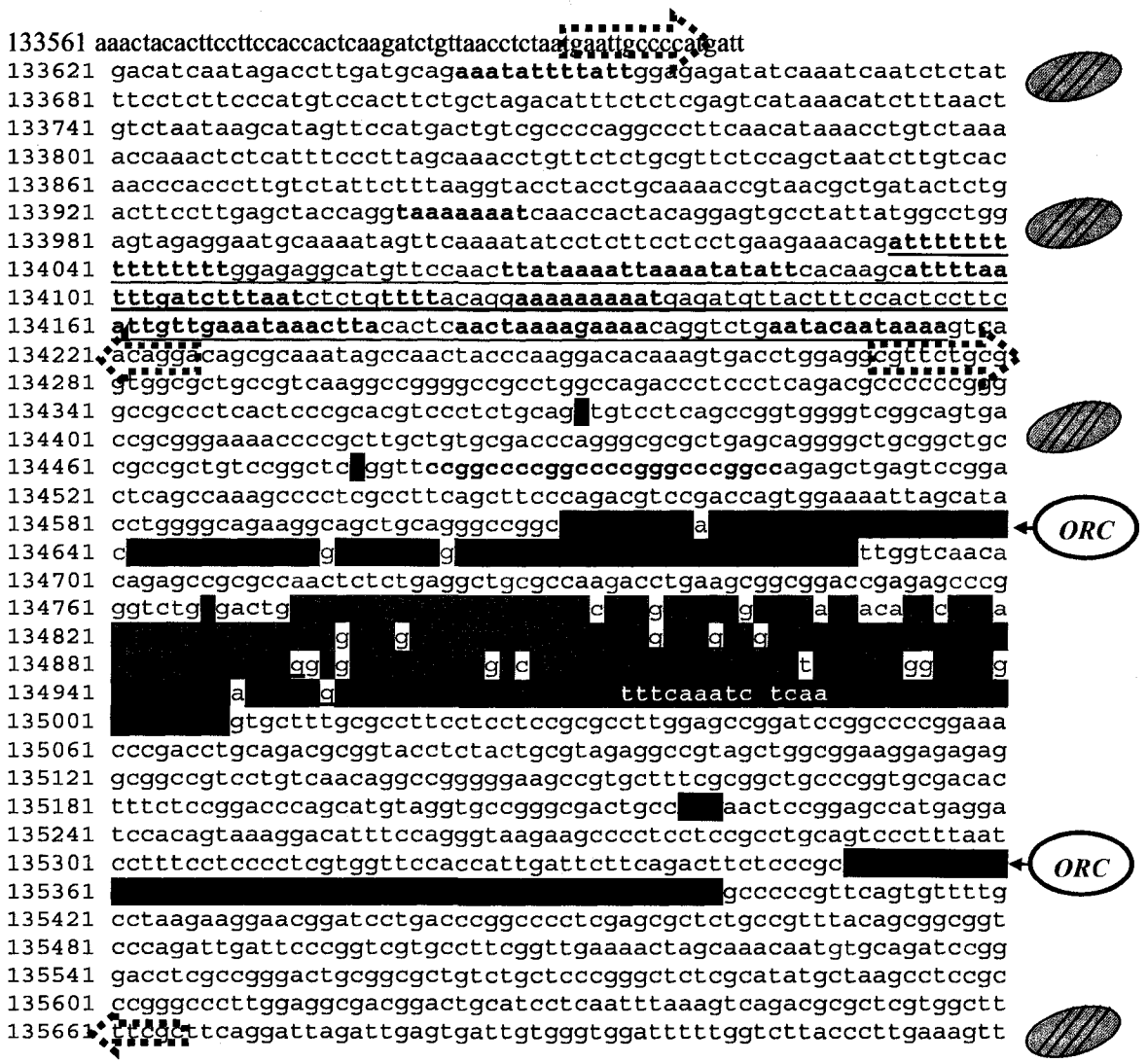
* **To whom correspondence should be addressed.**

Tumour Biology Group, Northeastern Ontario Regional Cancer Program at the Sudbury Regional Hospital, 41 Ramsey Lake Road, Sudbury, Ontario P3E 5J1, Canada

E-mail: hlee@hrsrh.on.ca

Supplementary Data include:

(1) Supplementary Figures 1 and 2.
(2) Supplementary Tables 1, 2, and 3
(3) Supplementary experimental procedures
(4) Supplementary References



Supplementary Figure 1. The nucleotide sequence and other features of the *DBF4* promoter region.

(1) Numbering scheme is according to that of the accession number AC003083; the +1 position of translational initiation (i.e., the A of the ATG) is corresponding to 135219 of the BAC clone; the numbers in the brackets are relative position from the ATG; "–" and "+" denote up- and down-stream of the ATG start-codon.

(2) The core promoter region (green highlight) is from 134934 (–285) to 135008 (–211). A 15-mer A+T-rich sequence (134974–134988) in the core promoter is in white (and red) letters (Wu and Lee, 2002).

(3) The auxiliary promoter (blue highlight) is from 134772 (–447) to 134934 (–286) (Wu and Lee, 2002).

(4) The initiations on the sense strand (transcriptional direction) are clustered around 134536 (–683) to 134955 (–264), and those on the antisense strand are clustered between 135529 (+311) and 135538 (+320). Nucleotides from which DNA replication initiates (determined by RIP assays) are written in red and highlighted in yellow. Data is from Figure 4B.

(5) Confirmed ORC binding sites are highlighted in purple (Prom7C2, 134611 [–608] to 134690 [–529]; HS(A), 135350 [+131] to 135401 [+185]). Data is from Figure 6A.

(6) Approximate nucleosome sites are shown on the right of the sequence. Data is from Figure 6B.

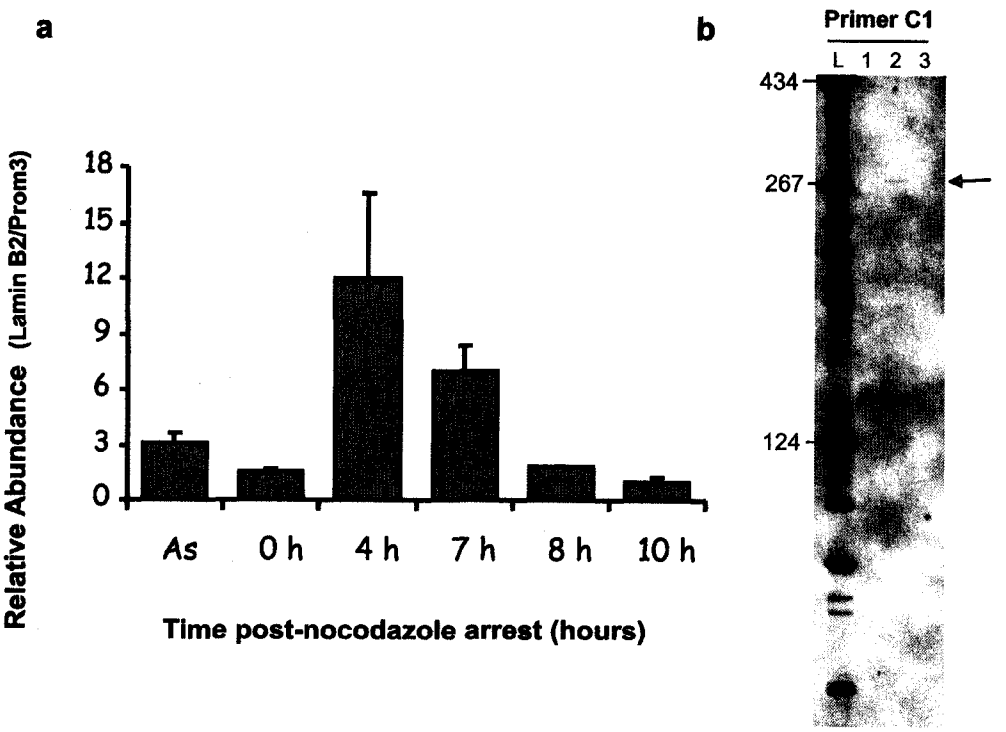
(7) The 865-bp DNA segment between 133591 (–1628) and 134225 (–994) (the start and end are marked with dotted green arrows) contains putative enhancer elements. The 184-bp DNA segment spanning from 134033 (–1186) to 134216 (–1003) (underlined in green) within this putative enhancer region is highly A+T-rich (75%), and contains several A/T tracks (bold letters) (Wu and Lee, 2002).

(8) According to NCBI gene bank analysis, the DNA segment spanning from 134272 (–947) to 135664 (+445) (% of G/C=66.5) is a putative CpG island (the start and end are marked as dotted blue arrows); an extremely G+C-rich region (letters in bolded blue) is present within this CpG island (134482 [–737] to 134505 [–714]).

(9) Transcription start sites (blue letters and purple highlight): (a) Bolded upper case "**T**" at134984 (–235), the major transcription-start site; (b) other transcription-start sites include "**t**"

at14999 (–220), "**t**" at 134974 (–245), "**a**" at 134914 (–305), "**c**" at 134897 (–322), "**c**" at 134888 (–331). Initiation also occur far upstream of these sites including, "**a**" at 134767 (–452), "**c**" at 134477 (–742), and "**c**" at 134372 (–847) (Wu and Lee, 2002).

(10) Nucleotide sequences underlined in red are putative/confirmed binding sites for the following proteins: ACGCGT, ***Mlu*1**; CGGCGCC, **TFIIB** (putative); GGGGCGGGG, **Sp1** (confirmed); CACGAG, **HES 1** (putative) (Wu and Lee, 2002).



Lee, **Supplementary Figure 2**

Supplementary Figure 2. Determination of a RIP at the human lamin B2 locus using one-way PCR-based primer extension. (**a**) Replication initiation at the lamin B2 locus was the maximum at 4 h post-mitosis. Short nascent DNA templates (1–2 kb) were isolated from HeLa at several different timepoints post-mitosis. Analysis of DNA replication by nascent strand abundance assay at the lamin B2 locus was carried out with the primers shown in **Supplementary Table 1** and the protocol described in the Methods. As, asynchronous cells; 0 h, cells arrested at mitosis by nocodazole; 4 h–10 h, 4–10 h post-mitosis. (**b**) One-way PCR-based primer extension was carried out using primer C1 (**Supplementary Table 2**) (Dimitrova et al., 1996) to determine the leading strand RIPs on the sense strand of the human lamin B2 origin. Nascent DNA (1–2 kb) was isolated from emetine-treated (for 1 h) HeLa cells released for 4 h from the nocodazole-based mitotic arrest. L, digoxigenin-labeled molecular weight ladder (sizes in nucleotides); lane 1, no template control; lane 2, nascent DNA 4 h post-mitosis; lane 3, nascent DNA from nocodazole-arrested cells. The expected size of the amplified product is 273 nucleotides (Abdurashidova et al., 2000).

Supplementary Table 1. Primers used for Q-PCR. All primer pairs were designed using the Primer Express software (Applied Biosystems). Dissociation curve analysis was performed after each PCR reaction to confirm the presence of a single amplification product. Amplicon size was verified by agarose gel electrophoresis. (*) Positions are based on the sequence of the CTB-60N22 BAC clone (Accession #AC003083). The numbers with (§), (#) and (Ψ) are based on the BAC clone CTB-135C18 (Accession #AC005164), the GenBank file M94363, and the GenBank file X00364, respectively. "–"and "+" denote, respectively, up- and down-stream of the ATG start-codon of the *DBF4*.

Primer	Sequence (5'-3')	Position (*) (BAC clone)	Position (to ATG)	Tm (°C)	Amplicon size (bp)
Prom3 F	ATAAGGACCACGCAACCCAG	103965–103946	–31254 to –31273	58.3	95
Prom3 R	GCTCCACCTCCTGTCAGATCA	103871–103891	–31348 to –31328	59.5	
Prom2 F	CACACCACCCAGTGGGTACC	125543–125524	–9676 to –9695	60.8	86
Prom2 R	CCCTGGATGCGCCTTTTAA	125458–125476	–9761 to –9743	56.5	
Prom6 F	TTTTATTCATTGCTGGCCCTG	132279–132259	–2940 to –2960	53.8	54
Prom6 R	ACATTTGCCCATGGTCACAA	132226–132245	–2993 to –2974	55.3	
AT-2F	GAGATGTTACTTTCCACTCC	134138–134157	–1081 to –1062	50.1	135
AT-2R	GCCTCCAGGTCACTTTGT	134253–134272	–966 to –947	54.9	
Prom7B F	CCTTCTGCCCCAGGTATGC	134594–134576	–625 to –643	58.3	51
Prom7B R	AGCTTCCCAGACGTCCGAC	134544–134562	–675 to –657	59.5	
Prom7C2 F	AGTCCTGGCAAACCTAGAAG	134611–134630	–608 to –589	59.9	80
Prom7C2 R	ACCTAGTGGATGCAGTTAGC	134671–134690	–548 to –529	58.4	
Prom4 F	CTGCAGGTTGTGTTTCCGC	134821–134803	–398 to –416	57.3	70
Prom4 R	GAGAGCCCGGGTCTGAGAC	134752–134770	–467 to –449	59.7	
Prom8 F	TCTCTCCTTCCGCCAGCTAC	135119–135100	–100 to –119	58.6	52
Prom8 R	TGCAGACGCGGTACCTCTACT	135068–135088	–151 to –131	60.0	
HS(A) F	TCCTTCCTTCAGACCCCAGAG	135381–135401	+162 to +182	58.1	52
HS(A) R	CGGGTCCTCAGCTTCTTTTCT	135350–135370	+131 to +151	57.0	
CodR2 F	CCAAAAATCCACCCACAATCA	135683–135703	+464 to +484	53.7	52
CodR2 R	TCGTGGCTTTTCGCTTCAG	135652–135670	+433 to +451	55.9	
Cod5 F	CACACAACTGGAACACGGCTT	3433–3453 (§)	+4839 to +4859	58.3	52
Cod5 R	TCCCACATGCCCAAATCTTC	3484–3465 (§)	+4890 to +4871	55.4	
LaminB2F	CTCTTCCAAGCCCTGCGTC	3770–3788 (#)	N/A	58.5	51
LaminB2R	CGTTTTTGCAGGTTGTGCTGT	3800–3820 (#)	N/A	57.1	
Myc F (5)	TTCCGCCTGCGATGATTTATAC	1982–2003 (Ψ)	N/A	55.2	52 Ori map
Myc R (5)	ACTGTTTGACAAACCGCATCC	2015–2035 (Ψ)	N/A	56.2	
Myc2F	CGATGCATTTTTTGTGCATGA	1073–1093 (Ψ)	N/A	52.7	54 DNase I
Myc2R	CGGACAAACCGGACGTTTAA	1157–1138 (Ψ)	N/A	55.5	

Supplementary Table 2. Oligonucleotide primers used for replication initiation point (RIP) mapping experiments. (*) Position is based on the nucleotide sequence of the BAC clone CTB-60N22 (Accession #AC003083). "–" and "+" denote, respectively, up- and down-stream of the *DBF4* ATG start-codon. N/A, not applicable. (§) Primer name and sequence as reported by Dimitrova et al. (Dimitrova et al., 1996)

Primer	Sequence (5'-3')	Position (*) (BAC clone)	Position (to ATG)	Tm (°C)	GC-content	Annealing temp (°C)
pA	CTTCCGCCAGCTACGGCCTC	135113–135094	–106 to –125	63.0	70.0%	68
pB	CAGCAAGCGGGGTTTTCCCG	134423–134404	–796 to –815	62.0	65.0%	68
pC	GCGTAGAGGCCGTAGCTGGC	135089–135108	–130 to –111	62.9	70.0%	68
pD	AGCTTCCCAGACGTCCGAC	134544–134562	–675 to –657	59.5	63.2%	64
pY	GCCACGAGCGCGTCTGACTT	135658–135639	+440 to +421	62.8	65.0%	68
C1 (§)	TCGCATCACGTGACGAAGAGTCAGC	N/A	N/A	62.7	56.0%	64

Supplementary Table 3. Summary of replication initiation points (RIPs) shown in **Fig. 4**. (**a**) RIPs determined using primer A. (**b**) RIPs determined using primer C. The letters in red are precise replication initiation points. The complementary DNA sequences (bottom) are the template DNA sequence of the newly replicated leading strand DNA. "H," "M" and "L" denote RIPs used with high, medium and low frequencies, respectively.

(**a**) RIPs determined using pA

Nucleotide sequence	Position (relative to ATG)	Position (relative to BAC clone #AC003083)
gcggCgccg cgccGcggc	264 (M)	134955
ccgaTacgc ggctAtgcg	-270 (L)	134949
gcgcCccgc cgcgGggcg	-279 (M)	134940
ccgcCccgt ggcgGggca	-284 (L)	134935
cgccCcgtc gcggGgcag	-285 (L)	134934
tctcAcgcc agagTgcgg	-292 (M)	134927
ccccGcctc ggggCggag	-311 (L)	134908
ccgcCtcgc ggcgGagcg	-313 (L)	134906
ctcgCccct gagcGggga	-323 (H)	134896
cgccCctcg gcggGgtgc	-325 (H)	134894
gcccCtcgt cgggGagca	-326 (H)	134893
cgccCcgcc gcggGgcgg	-355 (L)	134864
cccgCcccc gggcGgggg	-358 (L)	134861
ccccCtcac ggggGagtg	-362 (M)	134857
ccttCtcgc ggaaGagcg	-379 (M)	134840
ctcgCgcct gagcGcgga	-383 (L)	134836
gcccTgcag cgggAcgtc	-399 (L)	134820
tgcaGgttg acgtCcaac	-403 (M)	134816
aggtTgtgt tccaAcaca	-406 (M)	134813

ggttGtgtt ccaaCacaa	-407 (M)	134812
gttgTgttt caacAcaaa	-408 (M)	134811
gtgtTtccg cacaAaggc	-411 (H)	134808
tccgCctct aggcGgaga	-416 (L)	134803
tctaCcccg agatGgggc	-422 (L)	134797
ccccGcctc ggggCggag	-426 (M)	134793
tctcAgacc agagTctgg	-454 (M)	134765
acccGggct tgggCccga	-460 (H)	134759
cccgGgctc gggcCcgag	-461 (H)	134758
ccggGctct ggccCgaga	-462 (H)	134757
tctcGgtcc agagCcagg	-468 (M)	134751
gccgCttca cggcGaagt	-477 (M)	134742
ccgcTtcag ggcgAagtc	-478 (M)	134741
aggtCttgg tccaGaacc	-485 (L)	134734
tggcGcagc accgCgtcg	-491 (L)	134728
ggcgCagcc ccgcGtcgg	-492 (L)	134727
cctcAgaga ggagTctct	-499 (M)	134720
agagTtggc tctcAaccg	-505 (L)	134714
gagtTggcg ctcaAccgc	-506 (L)	134713
tctgTgttg agacAcaac	-519 (L)	134700
gtgtTgacc cacaActgg	-522 (L)	134697
gttgAccaa caacTggtt	-524 (L)	134695
accaAacct tggtTtgga	-528 (L)	134691
ggagCccgg cctcGggcc	-556 (M)	134663
gcccCgccc cgggGcggg	-564 (L)	134655
caggGttat gtccCaata	-578 (L)	134641
ggttTgcca ccaaAcggt	-599 (L)	134620
ctgcAgctg gacgTcgac	-620 (M)	134599

aggtAtgct tccaTacga	-640 (L)	134579
ggcgAgggg ccgcTcccc	-683 (M)	134536

1711 (pA) ----- 49 sites in total
C, 20/49 (41%); T, 11/49 (22%); G, 10/49 (20%); A, 8/49 (16%)

(b) RIPs determined using pC

Nucleotide sequence	Position (relative to ATG)	Position (relative to BAC clone #AC003083)
caatGtgca gttaCacgt	+311	135529
caatgTgca gttacAcgt	+312	135530
gatcCggga ctagGccct	+320	135538

Supplementary Methods

Cell culture and synchronization. HeLa cells were grown in RPMI-1640 medium containing 10% (v/v) fetal bovine serum (Hyclone), 2.05 mM L-glutamine, 100 μg ml^{-1} streptomycin and 100 units ml^{-1} penicillin. HEK293 and 293T cells were grown in Dulbecco's modified Eagle's medium (DMEM) containing 10% (v/v) fetal bovine serum and antibiotics. Cells were maintained at 37°C in a humidified atmosphere with 5% CO_2/95% air.

To synchronize cells in mitosis, exponentially growing cells were incubated for 18 h in complete medium containing 50 ng ml^{-1} nocodazole (Cytoskeleton Inc.). Cells arrested at mitosis were collected by shake-off. Synchronization of HeLa cells at the G1/S border was achieved by double-thymidine block. Briefly, exponentially growing cells were incubated for 18 h in the complete medium containing 2.0 mM thymidine (Sigma). Cells were then washed once with 1× PBS (2.7 mM KCl, 1.0 mM KH_2PO_4, 137 mM NaCl, 10 mM Na_2HPO_4, pH 7.5), and maintained for 11 h in complete medium, followed by a second thymidine block for 14 h.

Subsequently, the cells were washed with 1× PBS, and then released into complete medium. Synchronization was confirmed by flow cytometry of propidium iodide stained cells.

Flow cytometry. Cells were harvested, washed with 1× PBS, and then fixed in 70% (v/v) ethanol for >18 h at –20°C. The cells were then washed twice with 1× PBS and incubated in propidium iodide staining solution (0.1% (w/v) sodium citrate, 0.3% (v/v) NP-40, 100 μg ml^{-1} propidium iodide, 100 μg ml^{-1} RNase A) for 1 h at room temperature. Samples were analyzed using a Beckman Coulter Epics Elite flow cytometer.

Chromatin immunoprecipitation (ChIP) assay. ChIP assays were performed following Abcam's X-ChIP protocol (http://www.abcam.com/ps/pdf/protocols/x_chip_protocol.pdf) with some modifications. Briefly, exponentially growing HeLa cells on three 10-cm dishes were washed with 1× PBS and cross-linked for 10 min at room temperature with 1% (w/v) formaldehyde in serum-free medium. Cells were then washed and scraped into 2 ml of 1× PBS containing 500 μM PMSF. Cell suspensions were pooled together and centrifuged at 700 *g* for 10 min at 4°C. If necessary, protease inhibitors and PMSF were added to the cell pellets which were stored at –80°C. Cell pellets were resuspended in 1.5 ml of X-ChIP Lysis buffer (50 mM HEPES-KOH pH 7.5, 140 mM NaCl, 1 mM EDTA pH 8.0, 1% (v/v) Triton X-100, 0.1% (v/v) sodium deoxycholate, 0.1% (w/v) SDS, 1 mM PMSF and protease inhibitors). DNA was fragmented by sonication to an average size of ~300 bp (6 pulses of 30 s each, at setting 5, Sonic Dismembrator Model 100, Fisher Scientific). Sonicated lysates were centrifuged at maximum speed in a microcentrifuge for 1 min at 4°C. The supernatant was transferred to a fresh tube and stored at –80°C until use. An aliquot of the sonicated cell lysate was used to determine the size of the fragmented chromatin and the final DNA concentration by the

following procedure. 450 µl of fresh Elution buffer (1% (w/v) SDS, 100 mM $NaHCO_3$) and 5 µl of Proteinase K (20 mg ml^{-1}) were added to 50 µl of lysate, followed by overnight incubation at 65°C to reverse the cross-links. After phenol-chloroform extraction and ethanol precipitation, DNA was resuspended in 100 µl of sterile water and quantified by spectrophotometry (Beckman DU-530). DNA size was estimated by agarose gel electrophoresis. This DNA preparation was also used for Q-PCR as an "Input" DNA.

Twenty five micrograms of DNA from cell lysates (above) were used in each immunoprecipitation reaction. Cell lysates were diluted 10-fold using ChIP Dilution buffer (1% (v/v) Triton X-100, 2 mM EDTA pH 8.0, 150 mM NaCl, 20 mM Tris-HCl pH 8.0, protease inhibitors) and pre-cleared with 50 µl of a 1:1 mixture of Protein A-agarose beads and Protein G-agarose beads (Upstate), followed by incubation for 3 h at 4°C. Immunoprecipitation was carried out for 4 h at 4°C using 4 µg of rabbit polyclonal anti-ORC4 antibodies (sc-20634, Santa Cruz Biotechnologies) or 4 µg of normal rabbit IgG as a negative control (sc-2027, Santa Cruz Biotechnologies). The complexes were recovered by incubating overnight at 4°C with 50 µl of a 1:1 mixture of Protein A- and Protein G-agarose beads. Beads were washed four times with Washing buffer (0.1% (w/v) SDS, 1% (v/v) Triton X-100, 2 mM EDTA pH 8.0, 150 mM NaCl, 20 mM Tris-HCl pH 8.0) and once with Final Washing buffer (0.1% (w/v) SDS, 1% (v/v) Triton X-100, 2 mM EDTA pH 8.0, 500 mM NaCl, 20 mM Tris-HCl pH 8.0). Elution was done with 450 µl of freshly made Elution buffer for 15 min at room temperature. Following the reversal of cross-links by incubating samples for 5 h at 65°C, DNA was purified by phenol-chloroform extraction and ethanol precipitation. An equal volume for each DNA preparation was used as template for Q-PCR. Values generated by Q-PCR were normalized by subtracting the amounts recovered with pre-immune IgG control. Finally, relative abundance was

calculated by comparing the value of each DNA segment recovered from immunoprecipitation to that of the sample taken prior to immunoprecipitation ("Input" DNA) and normalizing it to the value obtained for the Prom3 primer pair.

SUPPLEMENTARY REFERENCES

Abdurashidova, G., Deganuto, M., Klima, R., Riva, S., Biamonti, G., Giacca, M., and Falaschi, A. (2000). Start sites of bidirectional DNA synthesis at the human lamin B2 origin. Science *287*, 2023-2026.

Dimitrova, D.S., Giacca, M., Demarchi, F., Biamonti, G., Riva, S., and Falaschi, A. (1996). In vivo protein-DNA interactions at human DNA replication origin. Proc. Natl Acad Sci U. S. A *93*, 1498-1503.

Wu, X., and Lee, H. (2002). Human Dbf4/ASK promoter is activated through the Sp1 and MluI cell-cycle box (MCB) transcription elements. Oncogene *21*, 7786-7796

One-way PCR-based mapping of a replication initiation point (RIP)

Julia Romero[1,2] & Hoyun Lee[1–3]

[1]Department of Biochemistry, Microbiology and Immunology, the Faculty of Medicine, the University of Ottawa, 451 Smyth Road, Ottawa, Ontario K1H 8M5, Canada. [2]Tumour Biology Group, Northeastern Ontario Regional Cancer Centre/Sudbury Regional Hospital, 41 Ramsey Lake Road, Sudbury, Ontario P3E 5J1, Canada. [3]Department of Medical Sciences, the Northern Ontario School of Medicine, 935 Ramsey Lake Road, Sudbury, Ontario P3E 2C6, Canada. Correspondence should be addressed to H.L. (hlee@hrsrh.on.ca).

Published online 16 October 2008; doi:10.1038/nprot.2008.173

The *lamin B2* locus is the only mammalian origin whose replication initiation points (RIPs) have been mapped. Although this paper was published 8 years ago, no further mammalian RIP-mapping studies have been reported, largely due to technical difficulties of ligation-mediated (LM)-PCR used by the authors. Here, we report the development of a simple, one-way PCR-based protocol that allows one to accurately determine RIPs at mammalian origins. The procedure can be completed within 48 h from the time of cell lysis in the agarose gel. Nascent DNA is then isolated from the same gel after DNA is separated by alkaline gel electrophoresis. Subsequently, RIPs are determined by one-way PCR-based primer extension using labeled primers. Using this protocol, we have successfully mapped RIPs in the human *DBF4* locus. As one-way PCR is routinely used by many scientists, this protocol will provide a powerful new tool for studying DNA replication in many organisms including mammalian cells.

INTRODUCTION

Faithful replication of the entire genome once per cell cycle is critical to maintain genetic integrity. Because the origin of DNA replication (ori) plays a key role in this process, understanding the initiation mechanism is essential, for which mapping the replication initiation point (RIP) is the first step. Several RIPs have been mapped in unicellular organisms such as bacteria and budding yeast. However, RIP mapping in mammalian cells turned out to be very challenging, mainly due to technical difficulties. The *lamin B2* locus is thus far the only mammalian ori whose RIPs have been mapped[1]. The authors determined RIPs at this locus using the technically demanding ligation-mediated (LM) PCR[1]. Although this paper was reported in 2000, no one has reported another mammalian RIP using LM-PCR or any other methods, underscoring the need to develop more effective and user friendly RIP-mapping techniques.

Using one-way PCR-based primer extension, Gerbi and Bielinsky successfully identified the precise transition point between continuous (leading strand) and discontinuous (lagging strand) DNA synthesis at the *Saccharomyces cerevisiae* ARS1 ori[2–4]. The original method developed by these authors[2] may be divided into the following six steps:

- *Isolation of bulk genomic DNA* by cesium chloride gradient centrifugation.
- Benzoylated naphthoylated DEAE (BND) chromatography to *enrich replication intermediates.*
- *Treatment with T4 polynucleotide kinase* to phosphorylate DNA molecules lacking an RNA primer (i.e., to remove randomly nicked DNA).
- *λ-exonuclease digestion* to degrade phosphorylated DNA and enrich for RNA-primed nascent DNA, which is resistant to the enzyme.
- *Primer extension* by one-way PCR with a radioactive-labeled oligonucleotide primer to map the 5′-end of a nascent DNA strand.
- *Sequencing gel electrophoresis and detection* of the amplified primer extension products.

The authors rationalized that a RIP of the leading strand coincides with the 5′ end of the shortest strand detected by the primer extension because successively larger strands are presumably generated by the ligation of Okazaki fragments to the leading strand. Since it was first reported in 1997, this RIP-mapping technique has been used to determine precise RIPs at several ori loci in *Schizosaccharomyces pombe*[5] and the fly *Sciara cropophila*[6]. RIP mapping has also been used in archaea, such as *Pyrococcus abyssi*[7], *Sulfolobus solfataricus*[8] and *Haloarcula* sp[9]. However, the original RIP-mapping protocol was not considered sensitive enough to map RIPs in mammalian cells. Therefore, LM-PCR was used to map RIP at the human *lamin B2* origin[1].

We postulated that one-way PCR-based primer extension could be directly used for mapping RIPs in mammalian oris without the need of LM-PCR, if nascent strands can be effectively harvested. We could indeed achieve this goal by bypassing several steps required for enriching short nascent DNA strands. Importantly, this streamlined protocol did not compromise the quality and purity of the nascent DNA sample. The steps that we could omit include cesium chloride gradient centrifugation, BND chromatography, treatment with T4 polynucleotide kinase to phosphorylate DNA molecules lacking an RNA-primer and λ-exonuclease treatments. Using this new approach, we successfully mapped RIPs at the human *DBF4* locus, where we have recently identified a strong ori[10]. This new RIP-mapping method combines one-way PCR-based primer extension with several previously reported techniques to effectively isolate nascent DNA fragments as described below (see **Table 1** for comparison of different methods).

We validated our one-way PCR-based RIP mapping using the *lamin B2* locus model (see Supplementary Fig. 2 in ref. 10; http://www.nature.com.proxy.bib.uottawa.ca/nsmb/journal/v15/n7/

TABLE 1 | Summary and comparison of different approaches in determining RIPs.

Steps	Gerbi and Bielinsky[a]	LM-PCR[b]	This protocol[c]
Nascent DNA isolation, purification	• Genomic DNA extraction by CsCl centrifugation or DNAzol • BND chromatography • T4 polynucleotide kinase treatment • λ-exonuclease digestion	• Phenol:chloroform:isoamyl alcohol extraction and ethanol precipitation • Sucrose gradient centrifugation • T4 polynucleotide kinase treatment • λ-exonuclease digestion	• In-gel cell lysis and alkaline gel electrophoresis • Gel extraction
Primer extension	• 30 cycles of one-way PCR amplification using radioactive-labeled oligonucleotides and Vent (exo$^-$) DNA polymerase	• Second strand DNA synthesis • Ligation to double-stranded asymmetric linkers and PCR amplification (LM-PCR) • Radioactive primer extension using Vent (exo$^-$) DNA polymerase	• 30 cycles of one-way PCR amplification using nonradioactive primers and IDPol DNA polymerase
Sequencing gel electrophoresis and detection[d]			

[a]The protocol described by Gerbi and Bielinsky[2] is as follows: bulk genomic DNA is isolated using CsCl gradient centrifugation or DNAzol extraction[2], followed by BND chromatography, T4 kinase treatment and λ-exonuclease digestion to enrich nascent DNA. 5′-end mapping of the nascent strands is achieved by 30 cycles of standard one-way PCR amplification. [b]The method used by Abdurashidova and colleagues to map the RIPs at the human lamin B2 ori[1]. This protocol incorporates an LM-PCR step to increase sensitivity. [c]The modified one-way PCR-based RIP-mapping procedure described in this protocol was used to map the RIPs at the human *DBF4* locus[10]. [d]All three RIP-mapping protocols include the analysis of amplified DNA by sequencing gel electrophoresis.

suppinfo/nsmb.1439_S1.html). In this validation experiment, however, we had to omit betaine for the reason discussed in the section "The use of betaine in PCR" below.

This new one-way PCR-based protocol is certainly sensitive enough to determine mammalian RIPs, even relatively weak oris such as the one at the *lamin B2* locus. Nevertheless, this mapping method may not be ideal for the detection of RIPs at extremely weak mammalian ori loci.

Finally, this new protocol can also be very useful in determining RIPs in nonmammalian organisms because it can bypass several rather tedious steps. Thus, the RIP-mapping protocol described here can greatly facilitate research in the regulation of DNA replication of many different organisms including mammalian cells.

Experimental design

Cell type/preparation. To carry out RIP mapping at the *DBF4* ori locus, we used HeLa cells as the source of nascent DNA. HeLa cells were maintained in RPMI-1640 medium containing 10% (vol/vol) fetal bovine serum (Hyclone, Logan, UT), 2.05 mM L-glutamine, 100 μg ml^{-1} streptomycin and 100 U ml^{-1} penicillin. Although we could easily detect *DBF4* RIPs from exponentially growing HeLa cells, the results were much more pronounced when nascent DNA fragments were isolated from synchronous HeLa cells at 1 h in S phase[10]. This was achieved by releasing HeLa cells synchronized at the G_1/S border by a double-thymidine treatment into complete medium for 1 h as described previously[10].

Emetine treatment. One of the most serious problems of determining a RIP is that ligation of Okazaki fragments to a leading strand complicates the interpretation of the true initiation point of leading strand synthesis. In yeast cells, one strategy to circumvent this problem is to compare RIPs in wild-type and DNA ligase I-deficient mutants[3]. Since this method is difficult to apply in mammalian cells, we and others used emetine, which inhibits new production of lagging strands without affecting leading strand synthesis[1,11]. Although concern was raised previously regarding drug treatments and the generation of potential artifactual results[2], we did not observe any notable problem when we determined *DBF4* RIPs with cells maintained in the presence and absence of emetine[10]. However, we cannot completely rule out the potential effects of emetine on RIP-mapping results because we have not systematically compared RIP patterns of cells grown in the absence and presence of emetine.

Preparation of nascent DNA strands by alkaline gel electrophoresis. Several different approaches have been developed to isolate nascent DNA without contamination by randomly nicked DNA. Isolation of DNA molecules containing RNA primer has been a preferred method for RIP mapping[2]. However, yielding nascent strands by this approach may not be sufficiently high because many tedious steps are required. We found that the simplest and most promising approach of isolating nascent DNA strands is: (i) to lyse cells directly in the wells of alkaline agarose gels; (ii) to separate DNA by electrophoresis under denaturing conditions and (iii) to isolate short single-stranded DNA by gel extraction[12]. This procedure minimizes sample handling and has been previously employed by others to map oris in mammalian cells[13,14]. This simple method also appears to minimize DNA damage because origin enrichment was not found with nonreplicating control DNA[10]. If substantial damage occurred during nascent strand preparation, we would have observed the enrichment of origin sequence from the control DNA.

The use of betaine in PCR. Another modification we made is the use of betaine in our one-way PCR. This was particularly important for the RIP mapping at the *DBF4* locus, where G+C content is very high (~70%). Betaine has been previously shown to prevent the formation of secondary structures in template DNA and promote polymerization of GC-rich sequences[15]. Interestingly, adding betaine to the RIP-mapping reaction for the human *lamin B2* ori resulted in a decrease in efficiency[10], perhaps because of the high A+T content in this locus. The protocol presented here is optimal for the detection of RIPs at the human *DBF4* locus where the G+C content is high. Because we had to omit betaine in the reaction for the RIP mapping at the lamin B2 ori[10], a slight adjustment for mapping protocol at different oris may be necessary.

Choice of DNA polymerases. Although Vent (exo$^-$) DNA polymerase has been the polymerase of choice for RIP mapping[1–3], this enzyme was not ideal under our experimental conditions. The

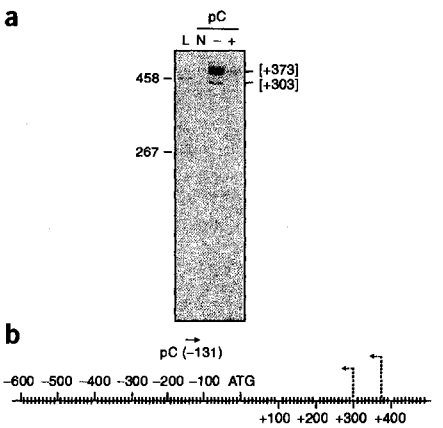


Figure 1 | RIP mapping of the human *DBF4* ori by one-way PCR-based primer extension. (**a**) Primer extension with primer pC (position is shown in **b**) identified two RIPs for 'antisense' (i.e., the opposite direction of *DBF4* transcription) leading strands emanating from the *DBF4* ori locus. The reaction was carried out with 1–2 kb nascent DNA fragments isolated from HeLa cells maintained for 1 h post-G_1/S arrest in the presence of emetine (lane '−'). The control cells were treated in the same manner, except 5 μg ml^{-1} of aphidicolin was also included along with emetine (lane '+'). L, digoxigenin-labeled molecular weight ladder (sizes in nucleotides are shown to the left of the panel); N, no template control. Numbers in brackets to the right of the panel indicate the approximate positions of RIPs determined based on band sizes. Note that the '−' and '+' before numbers denote, respectively, the upstream and downstream positions of the *DBF4* ATG codon. (**b**) A physical map of the human *DBF4* locus. The position of primer pC is shown. The position '−131' denotes 131 nucleotides upstream of the ATG translational start codon ('A' of the ATG is +1). The +100 denotes 100 nucleotides downstream from the ATG start codon. The vertical dashed lines are leading strand start sites at the *DBF4* ori. Arrows are the direction of replication deduced from data shown in the panel a.

possible reason for this is that the activity of this polymerase may not be compatible with betaine. In our hands, IDPol polymerase was the best for one-way PCR-based primer extension.

PCR primer design. The primers we use for RIP mapping are at least 19-bases long with a melting temperature close to ≥60 °C. As ~800 bases might be the maximum length that can be efficiently amplified by IDPol polymerase in the one-way PCRs[10], primer distance to the potential replication start sites should be <800 bases, preferably between 100 and 500 bases.

Choice of detection system. An important advantage of our approach is that this new protocol is designed to use nonradioactive labeling. Although radioactively labeled oligonucleotides have been the choice for primer extension reactions in previous RIP-mapping studies[1–3,5,6], we used digoxigenin-labeled primers, followed by immunoblotting with antidigoxigenin antibodies. We have found that this nonradioactive-labeling method is sensitive enough to detect RIPs in mammalian cells. Therefore, the entire work described in our protocol can be completed on the laboratory bench, eliminating the need for a specially shielded working area or disposing radioactive wastes.

Consideration of controls. To detect potential artifacts resulted by the reaction mix (including primers), each one-way PCR experiment should include a no-template control. In addition, a control RIP-mapping reaction should be carried out, in parallel with the main experiment, using nonreplicating DNA isolated from cells in mitosis[10] or cells treated with a high concentration of replication inhibitors such as aphidicolin (**Fig. 1a**). This is necessary to rule out the possibility that the products observed after primer extension are resulted from amplification of randomly nicked DNA, instead of replication intermediates (i.e., leading strands).

MATERIALS

REAGENTS

- 1-kb DNA molecular weight marker (New England BioLabs, cat. no. N3232)
- 100 mM dNTP set (Invitrogen, cat. no. 10297-018)
- Acrylamide (Sigma, cat. no. A8887) ! **CAUTION** Neurotoxin and potential carcinogen. Wear gloves at all times when handling acrylamide-containing solutions.
- Agarose (Invitrogen, cat no. 15510-027)
- Alkaline phosphatase-conjugated antidigoxigenin antibodies, Fab fragments (Roche Applied Science, cat. no. 11093274910 (1093274))
- Amersham Hybond N^+ nylon membrane (GE Healthcare Life Sciences, cat. no. RPN303B)
- Ammonium persulfate (Sigma, cat. no. A9164)
- Betaine (Sigma, cat. no. B2629)
- Blocking reagent (Roche Applied Science, cat. no. 11096176001 (1096176))
- Boric acid (Sigma, cat. no. B7901)
- Bromophenol blue (Sigma-Aldrich, cat. no. B8026)
- CDP-*Star* substrate (Roche Applied Science, cat. no. 11685627001 (1685627))
- Deionized formamide (Ambion, cat. no. AM9342) ! **CAUTION** Irritant. May be mutagenic and teratogenic.
- 5′-digoxigenin-labeled primers (Integrated DNA Technologies) ▲ **CRITICAL** Ideally, primers should have a T_m of ≥60 °C and be purified by HPLC.
- Disodium EDTA-$2H_2O$ (Sigma, cat. no. E5134)
- DNA molecular weight marker V, DIG-labeled (Roche Applied Science, cat. no 10821705001).
- Double-distilled water
- Emetine dihydrochloride hydrate (Sigma, cat. no. E2375) ! **CAUTION** Highly toxic. Emetine solutions should be handled with care and disposed of accordingly.
- Ethidium bromide (Sigma, cat. no. E7637) ! **CAUTION** Toxic and potent mutagen.
- Ficoll 400 (Fluka, cat. no. 46324)
- GIBCO 1× Trypsin-EDTA solution (Invitrogen, cat. no. 25200-056)
- Glacial acetic acid (Fisher Scientific, cat. no. A38 P212)
- Glycerol (Sigma-Aldrich, cat. no. G6279)
- IDPol DNA polymerase (ID Labs Biotechnology, cat. no. IDL007; supplied with 10× magnesium-free reaction buffer and 20 mM $MgSO_4$ solution) ▲ **CRITICAL** Although Vent (exo^-) DNA polymerase is commonly used in the primer extension reactions, it was not efficient under our experimental conditions. Although we used IDPol DNA polymerase in all of our experiments, other enzymes with similar characteristics may also work.
- KCl (Sigma, cat. no. P9541)
- KH_2PO_4 (Sigma, cat. no. P9791)
- Maleic acid (Sigma-Aldrich, cat. no. M153)
- Methylene bisacrylamide (Sigma, cat. no. M7279) ! **CAUTION** Irritant and potential teratogen.
- Na_2HPO_4 (Sigma-Aldrich, cat. no. 71496)
- NaCl (Sigma, cat. no. S3014)
- NaOH pellets (Sigma, cat. no. 221465)
- QIAquick gel extraction kit (Qiagen, cat. no. 28706)

- SDS (Sigma, cat. no. L4390) ! **CAUTION** Irritant.
- TEMED (Sigma, cat. no. T7024)
- Trizma base (Sigma, cat. no. T1503)
- Tween 20 (Sigma, cat. no. P9416)
- Urea (Sigma, cat. no. U6504)
- Xylene cyanol FF (Sigma-Aldrich, cat. no. X4126)

EQUIPMENT

- 1.5-ml Polypropylene micro tubes (Sarstedt, cat. no. 72.690.200)
- 0.5-ml PCR tubes (Sarstedt, cat. no. 72.735.002)
- 10-cm cell culture dishes (Sarstedt, cat. no. 83.1802.003)
- 15-ml conical tubes (Sarstedt, cat. no. 62.554.002)
- Amersham Hyperfilm ECL (GE Healthcare Life Sciences, cat. no. 28-9068-37)
- DNA-sequencing gel electrophoresis apparatus
- Horizontal DNA electrophoresis chamber (~14 cm × 10 cm gel-casting tray)
- Microcentrifuge (Eppendorf)
- PCR machine
- Trans-Blot SD Semi-dry electrophoretic transfer cell (Bio-Rad, cat. no. 170-3940)
- UV transilluminator (Chromato-Vue TM-36)
- X-ray film developer

REAGENT SETUP

0.5 M EDTA, pH 8.0 To prepare 100 ml solution, add 18.61 g of disodium EDTA-$2H_2O$ to ~70 ml distilled water. Adjust pH to 8.0 initially with NaOH pellets and then 10 N NaOH. Adjust the final volume to 100 ml and autoclave. The solution may be stored for several months at room temperature (~21 °C), provided that it remains free of contamination.

10 N NaOH Dissolve 40 g of NaOH in the final volume of 100 ml distilled water. It may be stored for up to 6 months at room temperature.

10× PBS (27 mM KCl, 10 mM KH_2PO_4, 1.37 M NaCl, 100 mM Na_2HPO_4, pH 7.5) To make 1 l of solution, dissolve 80 g NaCl, 2 g KCl, 14.4 g Na_2HPO_4 and 2.4 g of KH_2PO_4 in ~800 ml in distilled water. Adjust pH to 7.4–7.5, and then the final volume to 1 l with distilled water, followed by autoclaving. The buffer may be stored for several months at room temperature.

1× PBS + 10% (vol/vol) glycerol Prepare freshly at the time of use by mixing 100 µl 10× PBS with 800 µl sterile water and 100 µl glycerol. Add 5 µl of 10% (wt/vol) bromophenol blue to serve as tracking dye.

1 M Tris-HCl, pH 8.0 To prepare 300 ml solution, dissolve 36.34 g Trizma base in ~250 ml distilled water. Adjust pH to 8.0, and the final volume to 300 ml, followed by autoclaving. It may be stored for several months at room temperature.

10% (wt/vol) SDS Dissolve 2 g of SDS in sterile water to make the final volume of 20 ml solution. It may be stored at room temperature for several months.

5× SDS lysis buffer (50 mM Tris-HCl, 5 mM EDTA, 0.5% (wt/vol) SDS) Mix 500 µl of 1 M Tris-HCl, pH 8.0, 100 µl of 0.5 M EDTA, pH 8.0, and 500 µl of 10% (wt/vol) SDS. Adjust the final volume to 10 ml solution with sterile water. It may be stored at 4 °C for up to 1 month.

20× TE buffer (200 mM Tris, pH 8.0, 20 mM EDTA) To prepare 100 ml buffer, mix 76 ml distilled water with 20 ml 1 M Tris-HCl, pH 8.0, and 4 ml of 0.5 M EDTA, pH 8.0. Autoclave and store at 4 °C for up to several months.

10× DNA-loading buffer Mix 2 g Ficoll 400, 2 ml 0.5 M EDTA, pH 8.0, 1 ml 10% (wt/vol) SDS and 250 µl 10% (wt/vol) bromophenol blue. Adjust the final volume to 10 ml solution with sterile water. Aliquot and store at −20 °C for up to several months.

1-kb DNA marker solution (50 µg ml^{-1} 1-kb DNA ladder in 1× TE buffer) Mix 20 µl of 500 µg ml^{-1} stock DNA ladder (New England BioLabs) with 10 µl 20× TE buffer, 20 µl 10× DNA-loading buffer and 150 µl sterile water. The diluted ladder may be stored for several months at −20 °C.

'DNA ladder mix' Prepare a fresh batch at the time of use by mixing 40 µl of 1-kb DNA marker solution, 19 µl sterile water, 3.3 µl 1 N NaOH (the final concentration is 50 mM) and 2.7 µl 0.5 M EDTA, pH 8.0 (the final concentration is 20 mM).

Alkaline running buffer (50 mM NaOH, 1 mM EDTA) Prepare a fresh batch at the time of use by mixing 993 ml of distilled water with 5 ml of 10 N NaOH and 2 ml 0.5 M EDTA, pH 8.0.

50× TAE buffer (2 M Tris, 1 M glacial acetic acid, 50 mM EDTA, pH 8.4) Dissolve 121 g of Trizma base, 9.31 g disodium EDTA-$2H_2O$ and 28.6 ml glacial acetic acid, and adjust the final volume to 500 ml with sterile distilled water. The pH of the solution should normally be between 8.4 and 8.5 without further adjustment. It may be stored at room temperature for several months.

Sequencing gel-loading buffer (98% (vol/vol) formamide, 10 mM EDTA pH 8.0, 0.1% (wt/vol) xylene cyanol, 0.1% (wt/vol) bromophenol blue) To prepare 10 ml solution, mix 9.8 ml deionized formamide, 200 µl 0.5 M EDTA, pH 8.0, 0.01 g xylene cyanol and 0.01 g bromophenol blue. Store at 4 °C in 1 ml aliquots up to several months.

5× TBE buffer (450 mM Tris, 450 mM boric acid, 10 mM EDTA, pH 8.4) Dissolve 27 g Trizma base and 13.75 g boric acid in ~400 ml sterile distilled water. Add 10 ml of 0.5 M EDTA pH 8.0. Adjust the final volume to 500 ml with sterile distilled water. The pH should be 8.2–8.5 without further adjustment. Store at room temperature for up to one month. Discard if a precipitate is visible.

40% Acrylamide solution Dissolve 95 g acrylamide and 5 g bisacrylamide, and adjust the final volume to 250 ml with distilled water. Filter and store at 4 °C protected from light for up to several months.

10% (wt/vol) Ammonium persulfate Dissolve 0.5 g of ammonium persulfate in sterile water, and then adjust the final volume to 5 ml. The solution may be stored at 4 °C for several weeks. However, freshly prepared solution should be used to ensure effective gel polymerization.

5% (wt/vol) Acrylamide, 7.7 M urea sequencing gel Mix 7.5 ml of 40% acrylamide solution, 27.72 g urea and 6 ml 5× TBE, and adjust the final volume to 60 ml. Heat slightly and stir until all the urea has dissolved (make sure the solution cools down to room temperature before proceeding to the next step). Add 600 µl of 10% (wt/vol) ammonium persulfate and 36 µl TEMED, mix gently and promptly cast the gel. Allow the gel to solidify overnight at room temperature before running samples.

Diluted DIG-labeled DNA molecular weight marker solution Mix 2 µl of DIG-labeled DNA molecular weight marker stock (Roche), 14 µl sterile water and 14 µl sequencing gel-loading buffer. Store at −20 °C for up to several weeks. Heat denature immediately before use.

Maleic acid buffer (100 mM maleic acid, 150 mM NaCl, pH 7.5) Dissolve 11.61 g maleic acid and 8.77 g NaCl in ~800 ml distilled water. Adjust pH to 7.5, initially with NaOH pellets, and then 10 N NaOH. Adjust the final volume to 1 l with distilled water, and then autoclave. The buffer may be stored at room temperature for several months.

10× Blocking solution (10% (wt/vol) blocking reagent in maleic acid buffer) Dissolve 5 g of blocking reagent in 40 ml of maleic acid buffer by gently heating it (do not allow the solution to boil). Once completely dissolved, adjust the final volume to 50 ml with maleic acid buffer. Store at 4 °C for up to a month and dilute 1/10 in maleic acid buffer at the time of use.

Washing solution 0.03% (vol/vol) Tween 20 in maleic acid buffer. Prepare a fresh batch at the time of use.

Detection buffer (100 mM Tris, 100 mM NaCl, pH 9.5) Dissolve 3.05 g Trizma base and 1.46 g NaCl in ~200 ml distilled water. Adjust pH to 9.5, and the final volume to 250 ml, and autoclave. Store at room temperature for up to several months.

PROCEDURE

Isolation of short DNA-leading strands by alkaline gel electrophoresis ● TIMING ~21 h

1| Dissolve 1.25 g of agarose in ~90 ml of sterile water.

2| Cool down while stirring until the solution reaches ~60 °C.

3| Adjust the volume to 94.8 ml with sterile water.

4| Add 200 µl of EDTA 0.5 M pH 8.0 and 5 ml of NaOH 10 N, in that order.

5| Mix well and quickly pour ~80 ml of the gel mixture into a 14 cm × 10 cm gel-casting tray with a preparative comb.

6| While the gel solidifies, pretreat cells on a 10-cm culture dish for 1 h at 37 °C with 2 μM emetine (diluted in fresh, prewarmed growth medium).

7| Carefully dispose of the culture medium, and wash the cells (still on the plate) twice with 1× PBS.

8| Trypsinize cells by adding 1 ml of trypsin-EDTA solution and incubating for 3 min at 37 °C (or until cells have detached from the surface of the culture dish).

9| Add 4 ml of 1× PBS, transfer the cell suspension to a 15-ml conical tube, and centrifuge at 500*g* for 5 min at room temperature.

10| Resuspend the cell pellet in 1 ml of 1× PBS and transfer it to a 1.5 ml microcentrifuge tube.

11| Centrifuge the cells in a microcentrifuge at 400*g* for 5 min at 4 °C.

12| Resuspend the cell pellet in 260 μl of 1× PBS + 10% (vol/vol) glycerol solution containing 0.025% (wt/vol) bromophenol blue.

13| Load 85 μl of the cell suspension into one preparative well of the alkaline gel.

14| Add 25 μl of 5× SDS lysis buffer to each well and mix gently.

15| Incubate for 10 min at room temperature.

16| Add to each well 10 μl of 0.25 M EDTA pH 8.0 and 6.5 μl of NaOH 1 N (in that order). Mix gently each time.

17| In a separate well, load 30 μl of the 'DNA ladder mix'.

18| Add alkaline running buffer on both sides of the electrophoresis chamber until the solution reaches the top edge of the gel, without covering it.
▲ **CRITICAL STEP** Carry out electrophoresis initially under semi-dry conditions to prevent samples from floating out of the wells.

19| Run at 15 volts for ~2 h or until the bromophenol blue dye has completely entered the gel.

20| Add alkaline running buffer so that the entire gel is submerged, and then carry out electrophoresis overnight at 15 volts.

21| Neutralize the gel by incubating for 45 min in 1× TAE buffer.

22| Cut the gel to separate the size marker from the sample lanes.

23| Stain the gel containing the molecular size marker for 15 min in 1× TAE buffer containing 0.6 μg ml^{-1} ethidium bromide. In the meantime, keep the rest of the gel (containing the nascent DNA) in 1× TAE (i.e., nonstained).

24| Visualize the stained gel under UV and, based on the migration of the ladder, cut out the portion of the nonstained gel containing 1–2 kb nascent strands from each sample lane.

25| Transfer the gel slices into a microcentrifuge tube, and extract nascent DNA molecules using the QIAquick Gel Extraction Kit (Qiagen), according to the manufacturer's instructions. Due to the limit of the gel amount that can be loaded into an individual extraction column (maximum 400 mg), it is typical to use three columns to extract the entire 1–2 kb fraction from a single agarose gel lane. Elute each column using 35 μl of sterile water and then pool the three elutes into one vial before proceeding to the next step.
■ **PAUSE POINT** The nascent DNA can be stored for several weeks at −20 °C.

One-way PCR-based primer extension ● TIMING ~4 h

26| Set up PCRs in a final volume of 30 μl, which contains (final concentrations): 1× magnesium-free PCR buffer, 3 mM $MgSO_4$, 400 nM digoxigenin-labeled primer, 200 μM each dNTP, 1 M betaine, 1 U of IDPol DNA polymerase and 5 μl of the nascent DNA preparation (use 5 μl of sterile water for no-template control). We used an equal volume of nascent DNA when analyzing the same DNA preparation with different primers. If comparing different DNA preparations, we used an equal amount of nascent DNA, which we determine by quantitative PCR analysis.

▲ **CRITICAL STEP** For RIP mapping at the *DBF4* locus, the addition of betaine is crucial for the success of one-way PCRs. This is probably because the *DBF4* ori is very G+C rich. However, one should consider omitting betaine if the DNA sequence is A+T rich.
? TROUBLESHOOTING

27| Centrifuge the tube briefly to collect all the reaction solution at the bottom, and then carry out PCR as follows: initial incubation for 5 min at 95 °C, 30 cycles of amplification (1 min at 94 °C; 1 min at the appropriate annealing temperature (usually 68 °C) and 1.5 min at 72 °C), and final extension for 7 min at 72 °C, followed by incubation for 10 min at 4 °C.

28| Add 20 μl of sequencing gel-loading buffer to each tube.

29| Heat the samples for 4 min at 90 °C, and quickly chill on ice.
■ **PAUSE POINT** Samples can be stored for several months at −20 °C. If this is the case, they should be re-denatured by heat treatment (Step 29) at the time of use.

Sequencing gel fractionation and digoxigenin detection ● TIMING **~20 h**
30| Separate the primer extension products (2 μl) by electrophoresis on a 5% (wt/vol) acrylamide gel prepared in 0.5× TBE and containing 7.7 M urea. Run the gel for ~2 h (or until the bromophenol blue has migrated ~20 cm from the top of gel) at 5 watts, using freshly prepared running buffer (0.5× TBE). In the same gel, run 2 μl of 'Diluted DIG-labeled DNA molecular weight marker'.

31| Transfer DNA from the gel onto a Hybond N+ membrane using a semi-dry transferring apparatus for 1 h at 400 mA with 0.5× TBE. Due to the limited dimensions of the commercially available transferring apparatus, only a portion of the gel may be transferred at a given time. This gel portion usually cannot be larger than 20-cm long and 10-cm wide. Despite this size limit, the 20×10 cm dimension is usually sufficient to detect all the digoxigenin-labeled (molecular weight) DNA bands at relatively good resolution. However, if higher resolution is desired, one may load two or more sets of samples on the same gel at different times. Different portions of gel (e.g., "top" and "bottom") can then be transferred separately onto membranes. Carefully remove the glass plate that is not attached to the gel, and then place a pre-cut piece of thin chromatography paper on top of the gel portion from which DNA will be transferred onto a membrane (note that the gel is still attached to the other glass plate). Carefully cut the gel around the contour of the paper to separate the sample-containing gel portion from the remainder of the gel. Grab the paper (with the gel still attached to it) by a corner, and slowly and carefully peel it off the glass plate. Note that the paper is still attached to the gel while DNA is transferred onto a Hybond N+ membrane that is placed on the opposite side of the gel.

32| Fix the membrane for 6 min at 302 nm wave length (i.e., 'high' setting on a TM-36 UV Chromato-Vue Transilluminator).

33| To reduce background, 'block' the membrane for 1 h at room temperature using freshly diluted 1× blocking solution in maleic acid buffer.

34| Incubate the membrane overnight at 4 °C in 1× blocking solution containing a 1/10,000 dilution of the antidigoxigenin antibody. To minimize the amount of antibody, use a heat-sealed plastic bag to carry out this incubation step.

35| Wash the membrane twice for 15 min each with freshly prepared washing buffer at room temperature.

36| Incubate the membrane for 1 min in detection buffer.

37| Place the membrane in a new heat-sealed plastic bag and cover with CDP-*Star* substrate diluted 1/100 in detection buffer. Incubate for 5 min at room temperature.
▲ **CRITICAL STEP** To prevent contamination, the vial of CDP-*Star* substrate should be opened only under a sterile environment.

38| Remove excessive solution from the membrane, and then expose it to a film. The approximate size of one-way PCR-based primer extension products can be determined by comparing the position of a sample band and those of the digoxigenin-labeled molecular weight markers. Subsequently, a RIP position can be estimated by combining the primer position and the band size determined by gel electrophoresis[10].
? TROUBLESHOOTING

● TIMING
Steps 1–25, 21 h (including overnight gel electrophoresis)
Steps 26–29, 4 h
Steps 30–38, 20 h (including overnight antibody incubation)

? TROUBLESHOOTING

Troubleshooting advice can be found in **Table 2.**

TABLE 2 | Troubleshooting table.

Problems	Possible reasons	Solution
Spotted background observed on the membrane (Step 38)	Precipitation of antidigoxigenin antibodies in the vial	Quick spin the antibody vial immediately before use
No signal in any lanes, including molecular weight marker (Step 38)	DNA transfer to the membrane was not successful	Ensure that the DNA transfer equipment is working properly
	DNA was not properly fixed to the membrane after transfer	Fix with UV immediately after transfer to a membrane
No signal in one-way PCR samples (Steps 26, 38)	PCR conditions may be too stringent	Eliminate/reduce betaine from the PCRs and/or reduce the annealing temperature
Double bands observed in molecular weight ladder (Step 38)	Incomplete sample denaturation	Heat samples for at least 4 min at 90 °C and quickly chill on ice before gel electrophoresis

ANTICIPATED RESULTS

Figure 1a shows a typical result of RIP mapping at the *DBF4* ori using the protocol presented here. The size of the top band shown in **Figure 1a** (lane '−') is estimated to be 504 nucleotides. Since the start site of primer pC is −131 (i.e., 131 nucleotide upstream from the ATG translation start site), the RIP of this leading strand was determined to be the +373 position (**Fig. 1b**). This position was later confirmed by sequencing gel electrophoresis[10]. As expected, no bands were observed in the control sample lacking template DNA (**Fig. 1a**, lane 'N'). Similarly, primer extension signal was greatly reduced when the nascent template DNA was isolated from aphidicolin-treated cells (**Fig. 1a**, lane '+').

By carrying out detailed RIP mapping with many different primers, we discovered that *DBF4* ori contains two replication initiation zones that are regulated by an asymmetric bidirectional replication mechanism[10]. This finding was possible because we could construct a fine RIP map at this locus using the one-way PCR-based primer extension method described here. It should be noted that a nascent strand abundance assay alone, which is commonly used for mapping replication ori, could not detect two replication initiation zones[10]. Therefore, the protocol presented here is a very powerful tool in unraveling the regulation of replication oris.

Published online at http://www.natureprotocols.com/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Abdurashidova, G. *et al.* Start sites of bidirectional DNA synthesis at the human lamin B2 origin. *Science* **287**, 2023–2026 (2000).
2. Gerbi, S.A. & Bielinsky, A.K. Replication initiation point mapping. *Methods* **13**, 271–280 (1997).
3. Bielinsky, A.K. & Gerbi, S.A. Chromosomal ARS1 has a single leading strand start site. *Mol. Cell* **3**, 477–486 (1999).
4. Bielinsky, A.K. & Gerbi, S.A. Discrete start sites for DNA synthesis in the yeast ARS1 origin. *Science* **279**, 95–98 (1998).
5. Gomez, M. & Antequera, F. Organization of DNA replication origins in the fission yeast genome. *EMBO J.* **18**, 5683–5690 (1999).
6. Bielinsky, A.K. *et al.* Origin recognition complex binding to a metazoan replication origin. *Curr. Biol.* **11**, 1427–1431 (2001).
7. Matsunaga, F., Norais, C., Forterre, P. & Myllykallio, H. Identification of short 'eukaryotic' Okazaki fragments synthesized from a prokaryotic replication origin. *EMBO Rep.* **4**, 154–158 (2003).
8. Robinson, N.P. *et al.* Identification of two origins of replication in the single chromosome of the archaeon *Sulfolobus solfataricus*. *Cell* **116**, 25–38 (2004).
9. Sun, C., Zhou, M., Li, Y. & Xiang, H. Molecular characterization of the minimal replicon and the unidirectional theta replication of pSCM201 in extremely halophilic archaea. *J. Bacteriol.* **188**, 8136–8144 (2006).
10. Romero, J. & Lee, H. Asymmetric bidirectional replication at the human DBF4 origin. *Nat. Struct. Mol. Biol.* **15**, 722–729 (2008).
11. Burhans, W.C. *et al.* Emetine allows identification of origins of mammalian DNA replication by imbalanced DNA synthesis, not through conservative nucleosome segregation. *EMBO J.* **10**, 4351–4360 (1991).
12. Kamath, S. & Leffak, M. Multiple sites of replication initiation in the human β-globin gene locus. *Nucleic Acids Res.* **29**, 809–817 (2001).
13. Gerhardt, J. *et al.* Identification of new human origins of DNA replication by an origin-trapping assay. *Mol. Cell Biol.* **26**, 7731–7746 (2006).
14. Liu, G., Bissler, J.J., Sinden, R.R. & Leffak, M. Unstable spinocerebellar ataxia type 10 (ATTCT*(AGAAT)) repeats are associated with aberrant replication at the ATX10 locus and replication origin-dependent expansion at an ectopic site in human cells. *Mol. Cell Biol.* **27**, 7828–7838 (2007).
15. Henke, W. *et al.* Betaine improves the PCR amplification of GC-rich DNA sequences. *Nucleic Acids Res.* **25**, 3957–3958 (1997).

Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
European Journal of Cell Biology 84 (2005) 927–938
European Journal of Cell Biology
www.elsevier.de/ejcb

High levels of Cdc7 and Dbf4 proteins can arrest cell-cycle progression

Baoqing Guo[a], Julia Romero[a,b], Byung-Ju Kim[a,b], Hoyun Lee[a,b,*]

[a]*Department of Research, Northeastern Ontario Regional Cancer Centre, 41 Ramsey Lake Road, Sudbury, Ont., Canada P3E 5J1*
[b]*Department of Biochemistry, Microbiology and Immunology, the University of Ottawa Medical School, Ottawa, Ont., Canada*

Received 23 April 2005; received in revised form 7 September 2005; accepted 8 September 2005

Abstract

Cdc7-Dbf4 serine/threonine kinase is essential for initiation of DNA replication. It was previously found that overexpression of certain replication proteins such as Cdc6 and Cdt1 in fission yeast resulted in multiple rounds of DNA replication in the absence of mitosis. Since this phenomenon is dependent upon the presence of wild-type Cdc7/Hsk1, we hypothesized that high levels of Cdc7 and/or Dbf4 could also cause multiple rounds of DNA replication, or could facilitate entry into S phase. To test this hypothesis, we transiently overexpressed hamster Cdc7, Dbf4 or both in CHO cells. Direct observations of individual cells by fluorescence microscopy and flow cytometric analysis on cell populations suggest that overexpression of Cdc7 and/or Dbf4 does not result in multiple rounds of DNA replication or facilitating entry into S phase. In contrast, moderately increased levels of Dbf4, but not Cdc7, cause cell-cycle arrest in G2/M. This G2/M arrest coincides with hyperphosphorylation of Cdc2/Cdk1 at Tyr-15, raising the possibility that high levels of Dbf4 may activate a G2/M cell-cycle checkpoint. Further increase in Cdc7 and/or Dbf4 by 2–4 fold can arrest cells in G1 and significantly slow down S-phase progression for the cells already in S phase.

Keywords: Cdc7; Dbf4; Replication; Cell-cycle arrest; Checkpoint; Transfection; GFP; Cdc2/Cdk; Tyr-15 phosphorylation

Introduction

Cdc7-Dbf4 serine/threonine kinase is essential for entry into and traversing through S phase. Although Cdc7-Dbf4 kinase was initially thought to be a global regulator for the G1/S transition (Chapman and Johnston, 1989; Jackson et al., 1993; Johnston and Thomas, 1982), more recent data from *Saccharomyces cerevisiae* have shown that it directly functions at the individual origins of DNA replication (*oris*) throughout S phase (Bousset and Diffley, 1998; Donaldson et al., 1998). The protein levels of yeast Cdc7 and its mammalian homologues are relatively constant throughout the cell cycle. However, the Cdc7 kinase activity is observed only from late G1 through S phase (Yoon and Campbell, 1991; Kitada et al., 1992; Sclafani et al., 1988; Yoon et al., 1993; Guo and Lee, 1999, 2001), coinciding with expression of the regulatory subunit Dbf4 (Jackson et al., 1993; Oshiro et al., 1999; Nougarede et al., 2000; Jares et al., 2004; Kumagai et al., 1999). The catalytic Cdc7 kinase subunit is bound and activated by Dbf4, which, in turn, is regulated by transcription and protein turnover (Chapman and Johnston, 1989; Oshiro et al., 1999; Ferreira et al., 2000; Weinreich and Stillman, 1999; Kumagai et al., 1999).

*Corresponding author. Department of Research, Northeastern Ontario Regional Cancer Centre, 41 Ramsey Lake Road, Sudbury, Ont., Canada P3E 5J1. Tel.: +1 705 522 6237x2703; fax: +1 705 523 7326.
E-mail address: hlee@hrsrh.on.ca (H. Lee).

0171-9335/$ - see front matter © 2005 Elsevier GmbH. All rights reserved.
doi:10.1016/j.ejcb.2005.09.016

Yeast Mcm2 and its related proteins from *Xenopus* and mammals are substrates of Cdc7-Dbf4 kinase in vitro (**Jiang et al., 1999; Jares and Blow, 2000; Lei et al., 1997; Sato et al., 1997; Brown and Kelly, 1998; Kumagai et al., 1999; Kihara et al., 2000; Masai et al., 2000; Takeda et al., 2001; Ishimi et al., 2001**). Furthermore, genetic analysis showed that the phosphorylation of Mcm2 in yeast is dependent upon wild-type Cdc7 function (**Jiang et al., 1999; Lei et al., 1997; Takeda et al., 2001; Snaith et al., 2000**), suggesting that Mcm2 is a critical physiological substrate of Cdc7-Dbf4 kinase. The phosphorylation of Mcm2 by Cdc7-Dbf4 kinase is essential for *ori* activation; however, it alone is not sufficient (**Jares and Blow, 2000; Mimura and Takisawa, 1998; Owens et al., 1997; Tercero et al., 2000; Treuner et al., 1998; Walter and Newport, 2000; Zou and Stillman, 1998, 2000**). The activation of replication initiation also requires the Cdc45 activity, for which both Cdk2 and Cdc7-Dbf4 kinases are essential (**Jares and Blow, 2000; Mimura and Takisawa, 1998; Owens et al., 1997; Tercero et al., 2000; Walter and Newport, 2000; Zou and Stillman, 1998, 2000; Masai and Arai, 2002**). Thus, Cdc7-Dbf4 kinase is involved in the activation of chromosome replication at the final two steps, suggesting that it functions as a critical molecular switch for DNA replication (**Masai and Arai, 2002; Masai et al., 1999**).

In addition to its role in *ori* activation, Cdc7-Dbf4 is also involved in the S-phase checkpoint control. Budding yeast Dbf4, its fission yeast homologue Dfp1/Him1, and Hsk1 (fission yeast Cdc7 homologue) are hyperphosphorylated in response to hydroxyurea (HU) (**Weinreich and Stillman, 1999; Brown and Kelly, 1998; Takeda et al., 1999; Snaith and Forsburg, 1999**). Although Dbf4 under these conditions is more stable, Cdc7-Dbf4 kinase activity is attenuated (**Naeger et al., 1999**). Furthermore, analysis of temperature-sensitive hsk1 mutants suggests that Cdc7 may also be involved in the control of sister chromatid cohesion (**Kihara et al., 2000; Takeda et al., 2001; Snaith et al., 2000; Bailis et al., 2003**). These authors thus suggest that Cdc7/Hsk1 kinase has a second execution point after replication initiation, which is required to maintain chromosome integrity. Cdc7/Hsk1 interacts with and phosphorylates the heterochromatin protein 1 (HP1), which is separable from its function in the activation of replication initiation (**Bailis et al., 2003**).

The phosphorylation of Cdc2/Cdk1 at Tyr-15 by Wee1 (or related protein kinases) can result in cell-cycle arrest at G2, and the dephosphorylation of the Tyr-15 residue by Cdc25 (and related phosphatases) is the rate-limiting step for Cdc2 activation and mitotic entry (**Krek and Nigg, 1991; Norbury et al., 1991; Gautier and Maller, 1991; Borgne and Meijer, 1996; O'Connell et al., 1997**). Irradiation results in stabilization of Wee1 by phosphorylation, which, in turn, contributes to G2 arrest by phosphorylation of Cdc2 at Tyr-15 (**Raleigh and O'Connell, 2000**). The phosphorylation rate of Tyr-15 is also increased by the downregulation of Cdc25 phosphatase (**Kumagai and Dunphy, 1991; Sanchez et al., 1997; Peng et al., 1997**). While the phosphorylation of Cdc2 at Tyr-15 negatively regulates cell-cycle progression, Thr-161 phosphorylation is required for mitotic entry (**Ducommun et al., 1991; Gould et al., 1991**).

To gain further insight into the role of mammalian Cdc7 and Dbf4 in cell-cycle control, we transiently overexpressed Chinese hamster homologues of yeast Cdc7 and Dbf4 in Chinese hamster ovary (CHO) cells (**Guo and Lee, 1999, 2001**). It was previously found that overexpression of Cdt1 and Cdc6/18 in fission yeast resulted in multiple rounds of replication in the absence of mitosis in a wild-type Cdc7/Hsk1-dependent manner (**Gopalakrishnan et al., 2001; Yanow et al., 2001; Takeda et al., 1999, 2001**). Therefore, we anticipated that overexpression of Cdc7 and/or Dbf4 would result in similar multiple rounds of DNA replication in the absence of cell division, or facilitate entry into S phase. To our surprise, however, we have found that overexpression of hamster Cdc7 (ChCdc7) and/or Dbf4 (ChDbf4) causes cell-cycle arrest. In particular, moderately increased levels of Dbf4 result in G2/M arrest, which is correlated with elevated phosphorylation of Cdc2 at Tyr-15.

Materials and methods

Cell culture and synchronization

Chinese hamster ovary cells were maintained in minimal essential medium (MEM) supplemented with 10% Fetal Clone II serum (HyClone, Logan, Utah) in a humidified atmosphere containing 5% CO_2 and 95% air at 37 °C. To synchronize cells in G0, CHO cells (usually 30–40% confluent) were maintained for 45 h in isoleucine-omitted MEM supplemented with 10% dialyzed fetal bovine serum (Invitrogen). To achieve synchronization at the G1/S border or in the beginning of S phase, the cells trapped in G0 were released into cell cycle by maintaining them for 14 h in the complete medium containing 200 μM mimosine or 3 μg/ml aphidicolin, respectively (**Dijkwel and Hamlin, 1992; Lee et al., 1997**).

Plasmid constructs and transfection

The cloning of ChCdc7 and ChDbf4 has been described previously (**Guo and Lee, 1999, 2001**). To determine the subcellular localization of ChCdc7 and ChDbf4 proteins and their effects on DNA replication

in a living cell, we generated pEGFP-ChCdc7 and pEGFP-ChDbf4 fusion constructs by cloning the entire ChCdc7 or ChDbf4 open reading frame into the SmaI site of pEGFP (−N1, −C1), pEYFG (−N1, −C1), or pECFP (−N1, −C1) (Clontech). Transient transfection of the GFP-tagged ChCdc7 and ChDbf4 into CHO cells was carried out using LipofectAMINE™ (Invitrogen) according to the manufacturer's specifications with some modifications. Briefly, to prepare samples for transfection, cells were grown for 24–30 h in a 10-cm dish containing 5–6 sterile glass coverslips. Thirty microliters of LipofectAMINE™ was mixed with 120 μl H_2O in a microtube. Two micrograms of plasmid DNA diluted in 150 μl H_2O were prepared in a separate microtube. The two solutions were then mixed together and incubated at room temperature for 15 min (i.e., liposome–DNA mixture). The cells were washed twice with pre-warmed serum-free medium and then 5 ml serum-free MEM was added to each 10-cm culture dish. The liposome–DNA mixture was added to the culture medium in a dropwise fashion. After incubation for 5 h at 37 °C, fetal bovine serum was added to the cell culture to achieve a final serum concentration of 10%. Additional complete medium was added to the culture medium the following day to adjust the total volume to be 10–12 ml. Subsequently, the transfection efficiency and localization of the GFP-ChDbf4 and GFP-ChCdc7 fusion proteins were analyzed by fluorescence microscopy at scheduled timepoints.

Fluorescence-activated cell sorting (FACS) analysis

To prepare samples for FACS analysis, cells growing in a 10-cm culture dish were rinsed with PBS (2.7 mM KCl, 1.0 mM KH_2PO_4, 137 mM NaCl, 10 mM Na_2HPO_4, pH 7.4) and trypsinized, followed by washing once with cold PBS by centrifugation. Cells were then fixed with 75% ethanol, followed by staining with propidium iodide solution which contained 0.1% (w/v) sodium citrate, 0.3% (v/v) Nonidet P-40, 100 μg/ml RNase A, and 150 μg/ml propidium iodide (Sigma). For each sample, approximately 2×10^4 cells were analyzed using a Beckman Coulter Epic Elite flow cytometer. Fluorescence intensity upon stimulation with an argon-ion laser at 488 nm was measured between 500 and 540 nm and plotted against cell number. The fluorescence emission of total cells, transfected cells with low- or high-level intensity was analyzed by selecting different gates as described in the legend to Fig. 4A.

Western blot analysis

Asynchronously growing CHO cells were transfected with pEGFP vector, pEGFP-ChCdc7 or pEGFP-ChDbf4 construct. Cells were lysed 24 h post-transfection in RIPA buffer (50 mM Tris–HCl (pH 8.0), 150 mM NaCl, 0.5% sodium deoxycholate, 1.0% NP-40, and 0.1% SDS) plus 1.0 mM PMSF, 1 × protease inhibitor mixture (Roche), 1.0 mM sodium orthovanadate and 50 μM sodium fluoride. Total cell extract (15 μg/lane) was separated by 10% SDS-PAGE. Proteins in the gel were transferred to a PVDF membrane, after which the membrane was treated with 3.75% BSA solution containing 50 mM sodium fluoride and 5 mM sodium orthovanadate. The primary antibodies used were: antibodies generated against phosphorylated Cdc2 at Tyr-15 (Santa Cruz, #SC-7989) and Cdc2 at Thr-161 (Santa Cruz, #SC-12341), and anti-GFP antibodies (Santa Cruz, #SC-9996). The detailed procedure for immunoblot analysis was described previously (Naryzhny and Lee, 2003).

Luminescence spectrometry

To determine the association of GFP-ChCdc7 and GFP-ChDbf4 with chromatin and/or the nuclear matrix, the presence of GFP in different subcellular fractions was measured using an LS50 Luminescence Spectrometer (Perkin Elmer). The method of subcellular fractionation into cytosol, nucleoplasm and chromatin/nuclear matrix was as described previously (Naryzhny and Lee, 2001, 2004) with minor modifications. Briefly, cells in monolayer were rinsed twice with cold PBS and eluted by trypsinization, to which 5 ml PBS was added. The cells were then collected by centrifugation for 5 min at 1000 *g* and the cell pellet was resuspended in 1 ml cytosol extraction buffer (10 mM Tris–HCl (pH 6.8), 100 mM NaCl, 3.0 mM $MgCl_2$, 1 mM EDTA, 10% glycerol, 0.5% Triton X-100, 1.0 mM PMSF, and protease inhibitor cocktail (Roche)). The cell suspension was thoroughly mixed and incubated at room temperature for 10 min, and then centrifuged at 2000 *g* for 5 min.The supernatant, which is defined as cytosol fraction, was collected. The remaining nuclear pellet was resuspended in 1 ml nuclear extraction buffer (10 mM Tris–HCl (pH 6.8), 100 mM NaCl, 3.0 mM $MgCl_2$, 1.0 mM EDTA, 10% glycerol, 250 mM ammonium sulfate, 2.0 mM vanadyl ribonucleoside complex, 1.0 mM PMSF, protease inhibitor cocktail). This nuclear fraction was then centrifuged at 4000 *g* for 10 min, and the supernatant was collected (nucleoplasm). The remaining "insoluble fraction," which contains the chromatin and the nuclear matrix, was resuspended in 10 mM Tris–HCl (pH 7.4), 5.0 mM $MgCl_2$, 0.5% Triton X-100, 2.0 mM vanadyl ribonucleoside complex, 1.0 mM PMSF, protease inhibitor cocktail. Luminescence spectrometry was carried out at 488 nm excitation and 507 nm emission.

Detection of the DNA replication activities by 5-bromo-2′-deoxy-uridine (BrdU) incorporation

The CHO cells grown on glass coverslips placed in a 6-cm culture dish were transfected with plasmid. The DNA labeling using a BrdU-Labeling and Detection Kit (Roche Diagnostics) was carried out according to the manufacturer's specifications with minor modifications. Briefly, the growth medium was replaced with labeling medium and incubated for 1–24 h at 37 °C in a cell culture incubator. Upon completion of the labeling, cells on a coverslip were washed three times with PBS and then fixed for 20 min with methanol at −20 °C. Subsequently, the cells were washed three times with PBS containing 5% BSA, after which they were incubated at 37 °C for 30 min with 30 μl of diluted anti-BrdU antibodies. The cells on the coverslip were washed three more times with PBS containing 5% BSA, and then immunostained using 30 μl of diluted rhodamine-conjugated anti-mouse-IgG antibodies. After incubation for 30 min at 37 °C with the antibodies, the cells were washed three times with PBS and mounted on a slide.

Fluorescence microscopy

The BrdU-labeled cells on a coverslip were examined using a Zeiss Axiovert 100 inverted microscope. The filters (Chroma, Brattleboro, USA) used in the experiments were Cyan GFP (exciter D436/20x, emitter D480/40 m), CFP/YFP (exciter 51017x, emitter 51017 m), Endow GFP Bandpass (exciter HQ470/40x, emitter HQ525/50 m), and Rhodamine (exciter D540/25x, emitter D605/55 m). Images of the cells were taken with several filters consecutively at 63 × magnification. The image data were processed using the Northern Eclipse software (Empix Imaging Inc., Mississauga, Ontario, Canada).

Results and discussion

Transient expression of ChCdc7 and ChDbf4 fusion proteins in CHO cells

Since Cdc7-Dbf4 kinase is essential for *ori* activation, we wished to learn whether high protein levels of Cdc7 and Dbf4 could facilitate chromosome replication. We therefore overexpressed Cdc7, Dbf4 or both by transfecting asynchronous CHO cells with pECFP-ChCdc7, pEYFP-ChDbf4 or both constructs. Immediately after the transfection, chromosome replication of the transfected cells was examined by BrdU incorporation as described in Materials and methods. As shown in Fig. 1, single- or co-expression of Cdc7 and Dbf4 did not appear to affect chromosome replication, since both the transfected and non-transfected cells (i.e., cells with and without "green," respectively) could efficiently incorporate BrdU into their genome. The data from this initial experiment appeared to suggest that overexpression of Dbf4 and Cdc7 does not necessarily facilitate replication activation, which is consistent with previously published data (**Sato et al., 2003**). However, we noted that a few cells with elevated levels of Cdc7 or Dbf4 appeared to be less effective in BrdU incorporation.

Overexpression of ChCdc7 and ChDbf4 arrests cell-cycle progression

To further examine the effects of Cdc7 and Dbf4 overexpression on cell-cycle progression, we pulse-labeled CHO cells with BrdU after the transfected cells were first incubated for 24 h. As shown in Fig. 2, non-transfected cells showed strong BrdU incorporation (e.g., arrows in Fig. 2). The BrdU incorporation rates were not significantly affected in the cells with moderately increased levels of Cdc7 and/or Dbf4 protein ("a" in Fig. 2C, F and J). However, the majority of the cells with high levels of Cdc7 and/or Dbf4 did not effectively incorporate BrdU (e.g., "b" in Fig. 2C, F and J). These results suggest that high levels of Cdc7 and Dbf4 proteins prior to BrdU labeling can slow down or inhibit DNA replication, instead of facilitating it. However, we could not rule out the possibility that the cells with high levels of Cdc7 and/or Dbf4 could not replicate DNA during the 3-h pulse period because they were not in S phase. To rule out this possibility, CHO cells were first synchronized in G0 by incubating them for 25 h in isoleucine-omitted medium. The cells were then transfected with ChCdc7 and/or ChDbf4 recombinant plasmid, after which they were incubated for another 20 h in the isoleucine-omitted medium. Under these conditions, most of the CHO cells are trapped in the resting stage (G0 phase) (Fig. 3A, G0). Subsequently, the cells were released into cell cycle by incubating them in complete medium containing either aphidicolin or mimosine for 14 h, by which time most cells reached the G1/S boundary (Fig. 3A, G1/S). BrdU was included in the culture medium along with mimosine or aphidicolin to detect replication activities during the synchronization of the cells at the G1/S interface (Fig. 3B). Alternatively, the cells trapped in G1/S were further incubated for 3 h in complete medium to allow DNA replication prior to BrdU pulse labeling (Fig. 3A, 3 h; Fig. 3C). As previously shown under similar conditions (Dijkwel and Hamlin, 1992), the cells treated with mimosine for 14 h showed very little DNA replication (Fig. 3B, V, XII). In contrast, the cells treated with aphidicolin for 14 h incorporated a significant amount of BrdU (Fig. 3B, II and VIII), suggesting low levels of

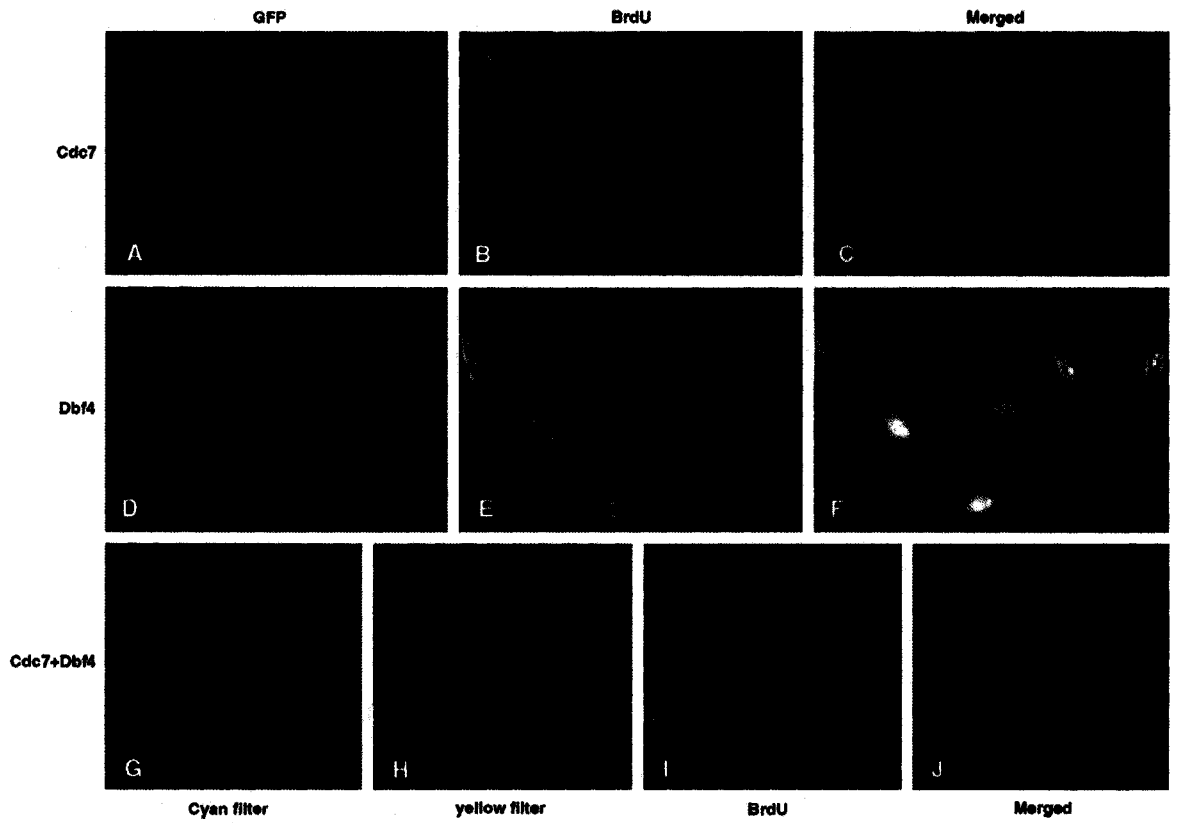


Fig. 1. Transient expression of ChCdc7 and ChDbf4 fusion proteins in CHO cells. Asynchronously growing CHO cells were transfected as described in Materials and methods with CFP-ChCdc7 (Cdc7), GFP-ChDbf4 (Dbf4) or both (Cdc7+Dbf4), immediately followed by BrdU labeling. The expression of the GFP-fusion proteins, their subcellular localization, and DNA replication (i.e., BrdU-labeled DNA) were observed at 24 h post-transfection. The Cdc7 was cloned into pECFP-C1, which shows in blue with Cyan GFP filter (A, G). Dbf4 was cloned into either pEGFP-C1 or pEYFP-C1, which can be seen with both GFP Bandpass (green) and CFP/YFP (Yellow) filters (D, H). Note that both cyan and yellow (CFP and YFP) can be observed using a CFP/YFP filter (H). Since anti-BrdU antibodies are rhodamine-conjugated, the cells that had incorporated BrdU (i.e., replicated DNA) are shown as red with a Rhodamine filter. (C, F, J) Merged images.

DNA replication continue to occur under aphidicolin treatments. Consistent with the data in Fig. 2, cells with high levels of Cdc7 (arrows in Fig. 3B, I–VI), Dbf4 (arrows in Fig. 3B, VII–IX), or both proteins (arrows in Fig. 3B, X–XIII) showed very little DNA replication activities.

Even when DNA replication was allowed to actively occur for 3 h after release from the G1/S arrest, cells expressing high levels of exogenous Cdc7 (arrows in Fig. 3C, I–III), Dbf4 (arrows in Fig. 3C, IV–VI), or both (arrows in Fig. 3C, VII–XIV) did not effectively replicate DNA. Thus, the high protein levels of ChCdc7 and/or ChDbf4 appear to arrest cells in G1 or result in significantly reduced replication activities for the cells already in S phase. The cells with moderately increased levels of Cdc7 or Dbf4 appear to replicate their chromosomes almost normally (e.g., "a" in Fig. 3B and C). To gain a better understanding about the overall effects of high levels of Cdc7 and Dbf4 on cell-cycle progression, we carried out a population-based study using flow cytometry. In these experiments, the cells arrested and transfected in G0 as described above (also, legend to Figs. 3 and 4) were released into cell cycle by incubating them in complete medium without adding any drug (time 0 in Fig. 4). At each scheduled timepoint, the cells were analyzed by flow cytometry. For the purpose of this analysis, we arbitrarily divided the CHO cells expressing low-to-moderate or high levels of Cdc7 and Dbf4 as shown in Fig. 4A. The cells classified as high levels contained 2–4 fold more GFP-Cdc7 or GFP-Dbf4 fusion proteins than those classified as low-to-moderate levels. (2–10 folds more GFP proteins were present in the high level when GFP vector alone was transfected, Fig. 4A). Exogenously expressed GFP alone

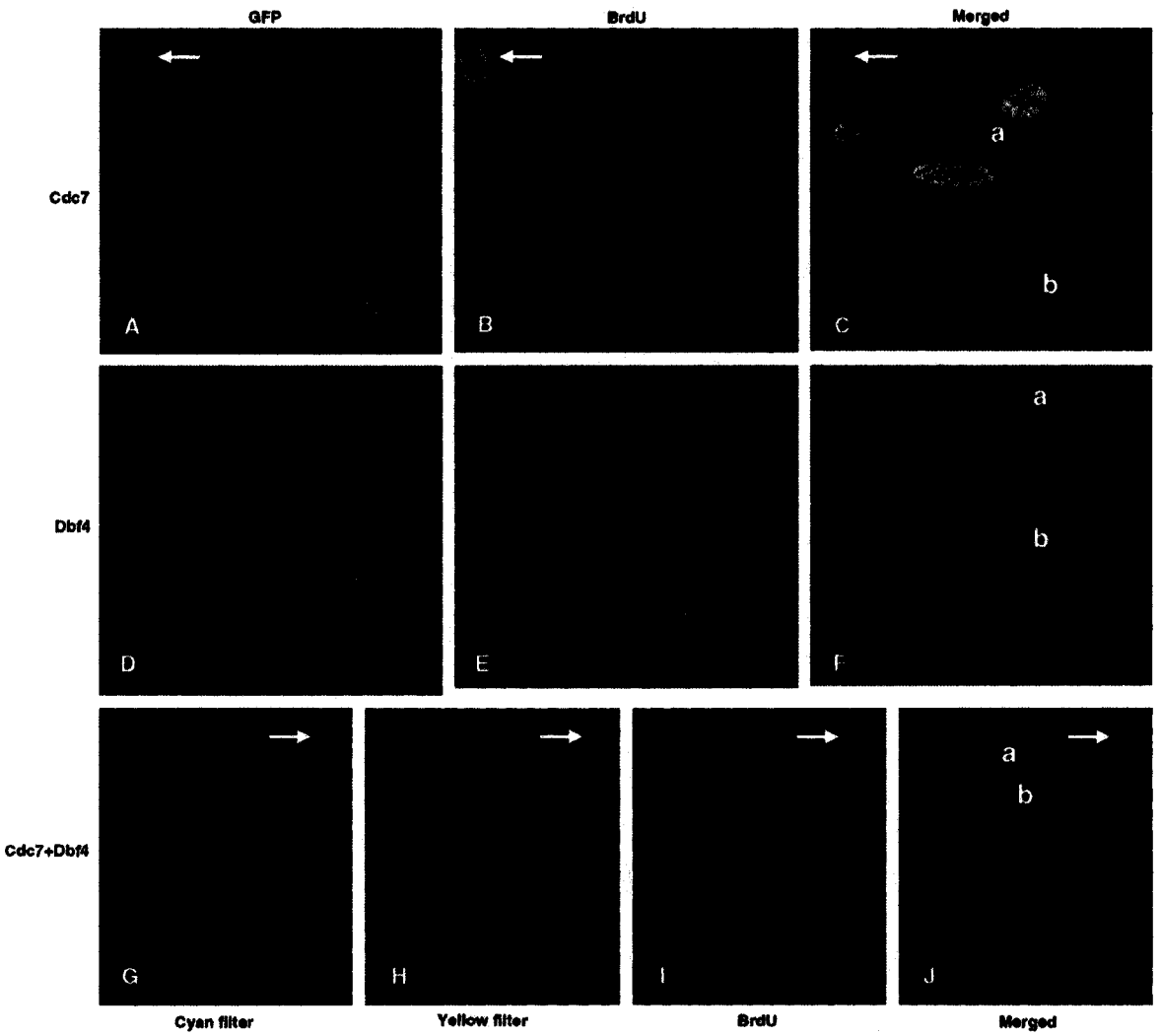
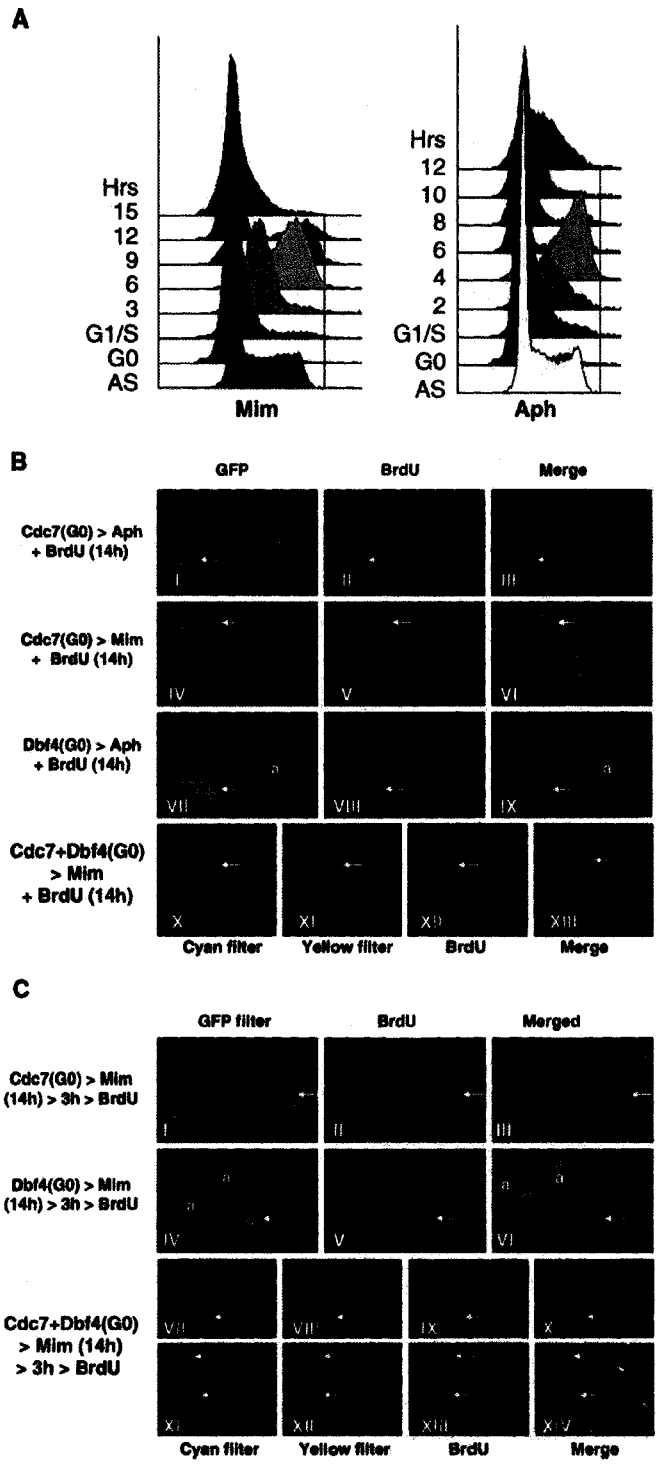


Fig. 2. High levels of Cdc7 and Dbf4 may inhibit DNA replication. Asynchronously growing CHO cells were transfected with GFP-Cdc7, GFP-Dbf4 or both constructs. Unlike the experiments described in Fig. 1, a 3-h pulse labeling was carried out 24 h post-transfection. ChCdc7 was cloned into pEGFP-C1 (A) or pECFP-C1 (G, H), while ChDbf4 was cloned into pEGFP-C1 (D) or pEYFP-C1 (H). "a" and "b" in (C, F, J) indicate cells with moderately increased and highly increased levels of Cdc7 and/or Dbf4 fusion protein, respectively. Note that cells marked "a" replicated their DNA, while the cells marked "b" apparently did not. "Cyan" and "Yellow" are Cyan GFP and CFP/YFP filters, respectively. Arrows indicate non-transfected cells.

Fig. 3. High expression of Cdc7 and Dbf4 in G1 can inhibit cells from entry into S phase. (A) CHO cells were transfected with GFP-Cdc7 or GFP-Dbf4 constructs after the cells were maintained in isoleucine-omitted medium for 25 h, after which they were maintained in the same medium for another 20 h. Under these conditions, most of the cells are in the resting stage (G0). Subsequently, the cells were released into complete medium containing mimosine (Mim) or aphidicolin (Aph). Under these conditions, cells traverse from G0 through G1 to the G1/S border. AS, asynchronous cells. (B) Cells in G0 were transfected with GFP-Cdc7 alone (I–VI), GFP-Dbf4 alone (VII–IX), or CFP-Cdc7 plus GFP-Dbf4 (X–XIII) as described above. The cells were then incubated in complete medium containing BrdU plus aphidicolin (I–III, VII–IX) or BrdU plus mimosine (IV–VI, X–XIII) for 14 h. For filter specifications, see legend to Figs. 1 and 2. (C) Cells synchronized and transfected in G0 were maintained in complete medium containing mimosine for 14 h to trap them at the G1/S interface. Subsequently, the cells were released into complete medium for 3 h to allow DNA replication to actively occur. Subsequently, the cells were pulse labeled with BrdU for 3 h. The panels VII–X and XI–XIV are two different sets of ChCdc7 and ChDbf4 co-transfection experiments. Arrows show cells with high levels of Cdc7, Dbf4 or both.



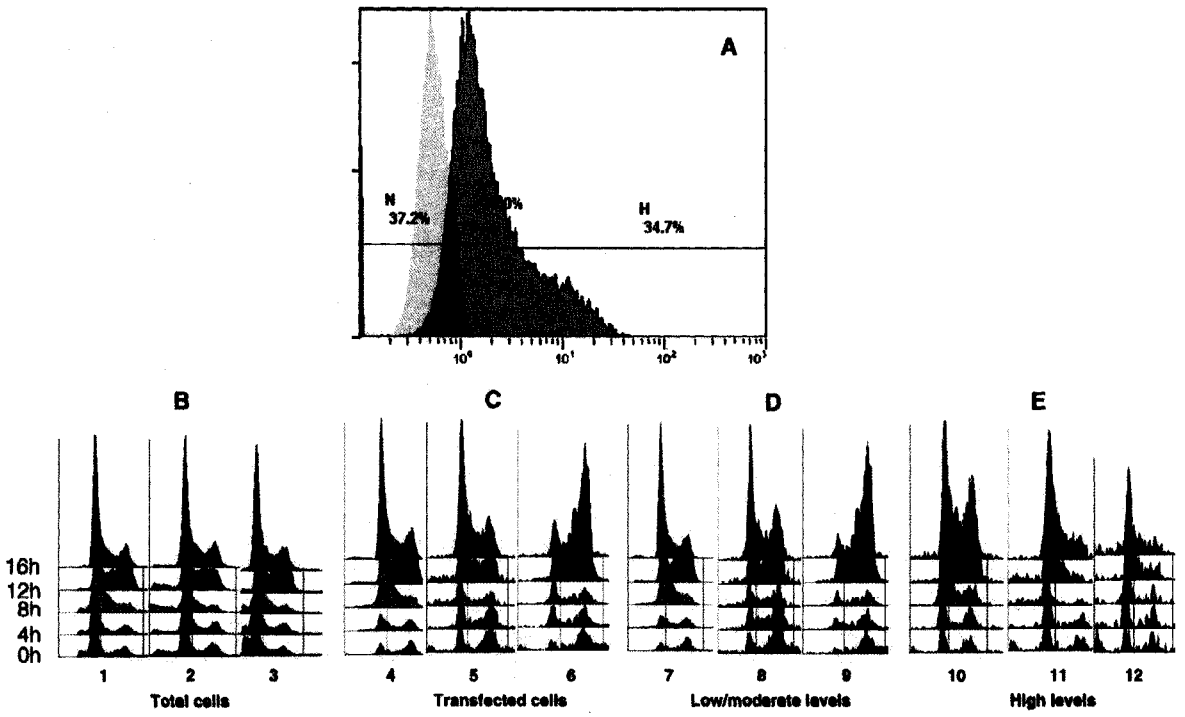


Fig. 4. Moderately increased levels of Dbf4 can arrest cells in G2 while high levels of Cdc7 or Dbf4 can arrest cells in G1 phase. CHO cells in G0 were transfected with either empty vector, GFP-Cdc7 or GFP-Dbf4. The cells were then released into cell cycle by growing them in the complete medium. Subsequently, the cells were determined for their cell-cycle positions using flow cytometric analysis. (A) To determine the effects of different levels of GFP-fusion proteins on the cell-cycle progression, the cells were "virtually" divided into the groups of transfected cells (which are included in the peak on the right) using a computer program built in the flow cytometer (Beckman Coulter). "L", "H", and "N" denote, respectively, low-to-moderately increased levels of fusion proteins, high levels of fusion proteins, and non-transfected cells. The low-to-moderate (1.2 to 2.2) and high (2.3 and above) expression levels were arbitrarily defined by the levels of GFP "fluorescence units." The example shown here is CHO cells transfected with GFP vector alone. The peak on the left represents non-transfected cells. (B–E) Cells were virtually sorted into the following four groups according to the levels of fluorescence: (i) total cells (i.e., transfected plus non-transfected cells; lanes 1–3), (ii) transfected cells (i.e., cells with any level of fluorescence; lanes 4–6), (iii) cells containing low-to-moderate levels of GFP proteins (lanes 7–9), and (iv) cells with high levels of GFP (lanes 10–12). The cell-cycle positions of these four groups were analyzed at different timepoints after release from G0 (timepoints are indicated on the left of lane 1). Lanes 1, 4, 7 and 10, transfected with GFP vector alone; lanes 2, 5, 8, and 11, transfected with GFP-Cdc7 fusion construct; lanes 3, 6, 9, and 12, transfected with GFP-Dbf4 fusion construct. It should be noted that relatively small numbers of cells contain green fluorescence at early timepoints (i.e., small peaks).

did not significantly affect cell-cycle progression (**Fig. 4B**, lanes 1, 4, 7 and 10). Similarly, low-to-moderate increases in GFP-Cdc7 proteins did not significantly affect cell-cycle progression (lanes 7 versus 8 in **Fig. 4D**). In contrast, moderate increases in Dbf4 protein levels effectively arrested cell cycle at G2/M (compare the 12- and 16-h timepoints in **Fig. 4D**, lanes 7 and 9). Furthermore, high levels of either Cdc7 or Dbf4 effectively arrested the cell cycle in G1 and S phase (**Fig. 4E**, lanes 11 and 12). As shown in **Table 1**, the data obtained from cell fractionation and luminescence spectrometry suggest that essentially all the GFP-Cdc7 and GFP-Dbf4 proteins in the nucleus are associated with either the chromatin or the nuclear matrix, indicating that they are likely functional (**Jares et al., 2004**; **Duncker et al., 2002**).

Moderately increased levels of Dbf4 result in hyperphosphorylation of Cdc2 at Tyr-15

To gain insight into the mechanisms of the G2/M arrest by moderately increased Dbf4 protein, we examined the phosphorylation status of Cdc2 after CHO cells were transfected with pEGFP vector, pEGFP-ChCdc7 or pEGFP-ChDbf4 recombinant

Table 1. Subcellular localization of Cdc7 and Dbf4 fusion proteins measured by luminescence spectrometry

	GFP vector (%)	GFP-Cdc7 (%)	GFP-Dbf4 (%)
Cytosol fraction	99.4	27.1	39.2
Nuclear fraction			
Nucleoplasm	0.4	0.0	0.0
Insoluble fraction	0.2	72.9	60.8

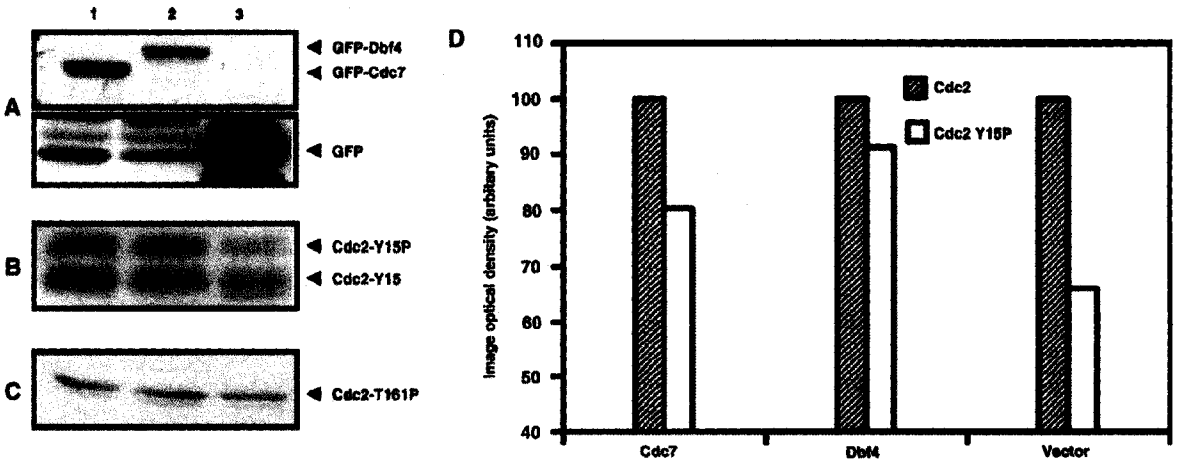


Fig. 5. High levels of exogenous Dbf4 may result in G2 arrest through Cdc2 phosphorylation at Tyr-15. (A–C) Asynchronous CHO cells were transfected with pEGFP-Cdc7 (lane 1), pEGFP-Dbf4 (lane 2), or pEGFP vector (lane 3). Cell extracts isolated at 24 h post-transfection were subjected to separation by SDS-PAGE and Western blot analysis for different antigens as indicated on the right. (D) Quantification of the signals in (B) using the Qscan program (Quanti Scan Version 3.0, Biosoft).

construct. Transfection efficiencies for Cdc7, Dbf4 and GFP vector were approximately 45%, 35% and 75%, respectively (the expression levels of these proteins are shown in **Fig. 5A**). The relative amount of phosphorylated Cdc2 at Tyr-15 was compared to the amount of non-phosphorylated Cdc2 protein (set as 100%). As shown in **Fig. 5B and D**, the phosphorylated Cdc2 at Tyr-15 was 80.3%, 91.5% or 66.0%, respectively, for the cells transfected with Cdc7, Dbf4 or vector. Considering relatively low transfection efficiencies, the increase in phosphorylated Cdc2 at Tyr-15 of the cells transfected with Dbf4 is quite significant. Unlike the phosphorylation at Tyr-15, the levels of phosphorylated Cdc2 at Thr-161 were not significantly different among the cells transfected with Cdc7, Dbf4 or vector alone (**Fig. 5C**). These data thus suggest that the G2/M arrest shown in the cells with moderately increased levels of Dbf4 may be caused, at least in part, by high levels of phosphorylated Cdc2 at Tyr-15.

The elevated activity of Wee1 (and related kinases) and/or decreased activity of Cdc25 (or related phosphatases) may contribute to the increased phosphorylation of Cdc2 at Tyr-15. It should be noted that phosphorylation upregulates Wee 1 kinase, but downregulates Cdc25 phosphatase. Therefore, phosphorylation of Wee1 and/or Cdc25 can increase the level of phosphorylated Cdc2 at Tyr-15, resulting in cell-cycle arrest at G2/M phase (**O'Connell et al., 1997; Raleigh and O'Connell, 2000**). Although both Wee1 and Cdc25 are substrates for Chk1 upon DNA damage (**O'Connell et al., 1997; Raleigh and O'Connell, 2000**), it may be possible that Cdc7 kinase can also phosphorylate either or both of them, particularly when Cdc7 kinase activity is unusually high due to the elevated levels of Dbf4. This raises an interesting possibility that high levels of Dbf4 protein can activate the G2/M cell-cycle checkpoint in a damage-independent manner.

The observation that moderately elevated levels of Dbf4, but not Cdc7, resulted in G2 arrest (**Fig. 4**) is consistent with the published data that Dbf4 is the rate-limiting factor for the Cdc7 kinase activity (**Yoon and Campbell, 1991; Kitada et al., 1992; Sclafani et al., 1988; Yoon et al., 1993; Guo and Lee, 1999, 2001**). However, we cannot completely rule out that Cdc7-independent mechanisms are involved in the G2/M arrest by moderately increased levels of Dbf4 protein (Hardy

and Pautz, 1996; James et al., 1999; Snaith and Forsburg, 1999).

We anticipated that overexpression of Dbf4 (and Cdc7) could facilitate DNA replication. Therefore, it was a surprise to find that high levels of Cdc7 and/or Dbf4 can apparently prevent cells from entering into S-phase or slow down S-phase progression (**Figs. 2, 3 and 4**). However, this result may not be completely unexpected, since overexpression of both Cdc7 and Dbf4 in yeast is lethal (**Nougarede et al., 2000**). In contrast, **Sato et al. (2003)** reported that overexpression of both huCdc7 and huDbf4/ASK does not cause significant effects on cell-cycle progression in human cells. Our data may reconcile these two conflicting conclusions. As shown in **Fig. 1**, interruption of cell-cycle progression is not clearly detectable when asynchronously growing cells are transfected with ChCdc7 or ChDbf4. Even when synchronized cells are analyzed in a similar way, the interruption of cell-cycle progression is still not appreciable with the entire cell population (**Fig. 4B**). However, when cells are analyzed according to the different levels of GFP-fusion proteins, the deregulation of cell-cycle progression by high levels of Cdc7 and/or Dbf4 is obvious (**Fig. 4D, E**). Taken together, it is possible that different levels of threshold for cell-cycle interruption by Cdc7 and Dbf4 may exist in yeast and human cells.

What would be the possible mechanism for the G1 cell-cycle arrest by high levels of Cdc7 and Dbf4? There is a line of evidence that Cdc7-Dbf4 kinase may be regulated by cyclin-dependent kinases (**Masai et al., 2000; Brown and Kelly, 1998**). **Masai et al. (2000)** found that human Cdk2 can phosphorylate Cdc7 in vitro at the threonine 376. Significantly, the replacement of threonine 376 to alanine dramatically reduces the Cdc7 kinase activity (**Masai et al., 2000**). If a similar regulation mechanism exists in vivo, it is possible that an excess amount of Cdc7 protein can titrate out the Cdk2 molecules in the cell. As a result, other Cdk substrates the phosphorylation of which is required for the cell-cycle progression into S phase may be hampered due to the lack of regulatory kinase, resulting in G1 arrest. We also cannot rule out the possibility that excessive Dbf4 proteins may phosphorylate certain proteins (and thus activate or inactivate them) that are not Cdc7-Dbf4 substrates under normal conditions, resulting in G1 arrest.

In summary, our data obtained from both fluorescence microscopy and flow cytometric analysis suggest that a moderate increase in Dbf4 protein results in a strong cell-cycle arrest at G2/M phase, probably due to elevated levels of phosphorylated Cdc2 at Tyr-15. Our data also suggest that high levels of Cdc7 and/or Dbf4 can arrest cells in G1. These data raise the possibility that Cdc7-Dbf4 kinase may be involved in the regulation of entry into S phase as previously suggested (**Chapman and Johnston, 1989; Jackson et al., 1993; Johnston and Thomas, 1982; Lee et al., 1997**), perhaps by controlling the activation of the earliest-firing origins of DNA replication.

Acknowledgments

This work was supported by operating grants from the Canadian Institutes of Health Research (MOP 57706) and Natural Sciences and Engineering Research Council of Canada (203528-2002) to H. Lee.

References

Bailis, J.M., Bernard, P., Antonelli, R., Allshire, R.C., Forsburg, S.L., 2003. Hsk1-Dfp1 is required for heterochromatin-mediated cohesion at centromeres. Nat. Cell Biol. 5, 1111–1116.

Borgne, A., Meijer, L., 1996. Sequential dephosphorylation of p34(cdc2) on Thr-14 and Tyr-15 at the prophase/metaphase transition. J. Biol. Chem. 271, 27847–27854.

Bousset, K., Diffley, J.F., 1998. The Cdc7 protein kinase is required for origin firing during S phase. Genes Dev. 12, 480–490.

Brown, G.W., Kelly, T.J., 1998. Purification of Hsk1, a minichromosome maintenance protein kinase from fission yeast. J. Biol. Chem. 273, 22083–22090.

Chapman, J.W., Johnston, L.H., 1989. The yeast gene, DBF4, essential for entry into S phase is cell cycle regulated. Exp. Cell Res. 180, 419–428.

Dijkwel, P.A., Hamlin, J.L., 1992. Initiation of DNA replication in the dihydrofolate reductase locus is confined to the early S period in CHO cells synchronized with the plant amino acid mimosine. Mol. Cell. Biol. 12, 3715–3722.

Donaldson, A.D., Fangman, W.L., Brewer, B.J., 1998. Cdc7 is required throughout the yeast S phase to activate replication origins. Genes Dev. 12, 491–501.

Ducommun, B., Brambilla, P., Felix, M.A., Franza Jr., B.R., Karsenti, E., Draetta, G., 1991. cdc2 phosphorylation is required for its interaction with cyclin. EMBO J. 10, 3311–3319.

Duncker, B.P., Shimada, K., Tsai-Pflugfelder, M., Pasero, P., Gasser, S.M., 2002. An N-terminal domain of Dbf4p mediates interaction with both origin recognition complex (ORC) and Rad53p and can deregulate late origin firing. Proc. Natl. Acad. Sci. USA 99, 16087–16092.

Ferreira, M.F., Santocanale, C., Drury, L.S., Diffley, J.F., 2000. Dbf4p, an essential S phase-promoting factor, is targeted for degradation by the anaphase-promoting complex. Mol. Cell. Biol. 20, 242–248.

Gautier, J., Maller, J.L., 1991. Cyclin B in *Xenopus* oocytes: implications for the mechanism of pre-MPF activation. EMBO J. 10, 177–182.

Gopalakrishnan, V., Simancek, P., Houchens, C., Snaith, H.A., Frattini, M.G., Sazer, S., Kelly, T.J., 2001. Redundant control of rereplication in fission yeast. Proc. Natl. Acad. Sci. USA 98, 13114–13119.

Gould, K.L., Moreno, S., Owen, D.J., Sazer, S., Nurse, P., 1991. Phosphorylation at Thr167 is required for *Schizosaccharomyces pombe* p34cdc2 function. EMBO J. 10, 3297–3309.

Guo, B., Lee, H., 1999. Cloning and characterization of Chinese hamster CDC7 (ChCDC7). Somat. Cell Mol. Genet. 25, 159–171.

Guo, B., Lee, H., 2001. Cloning and characterization of Chinese hamster homologue of yeast DBF4 (ChDBF4). Gene 264, 249–256.

Hardy, C.F., Pautz, A., 1996. A novel role for Cdc5p in DNA replication. Mol. Cell. Biol. 16, 6775–6782.

Ishimi, Y., Komamura-Kohno, Y., Arai, K., Masai, H., 2001. Biochemical activities associated with mouse Mcm2 protein. J. Biol. Chem. 276, 42744–42752.

Jackson, A.L., Pahl, P.M., Harrison, K., Rosamond, J., Sclafani, R.A., 1993. Cell cycle regulation of the yeast Cdc7 protein kinase by association with the Dbf4 protein. Mol. Cell. Biol. 13, 2899–2908.

James, S.W., Bullock, K.A., Gygax, S.E., Kraynack, B.A., Matura, R.A., MacLeod, J.A., McNeal, K.K., Prasauckas, K.A., Scacheri, P.C., Shenefiel, H.L., Tobin, H.M., Wade, S.D., 1999. *nimO*, an *Aspergillus* gene related to budding yeast Dbf4, is required for DNA synthesis and mitotic checkpoint control. J. Cell Sci. 112, 1313–1324.

Jares, P., Blow, J.J., 2000. *Xenopus* cdc7 function is dependent on licensing but not on XORC, XCdc6, or CDK activity and is required for XCdc45 loading. Genes Dev. 14, 1528–1540.

Jares, P., Luciani, M.G., Blow, J.J., 2004. A *Xenopus* Dbf4 homolog is required for Cdc7 chromatin binding and DNA replication. BMC Mol. Biol. 5, 5.

Jiang, W., McDonald, D., Hope, T.J., Hunter, T., 1999. Mammalian Cdc7-Dbf4 protein kinase complex is essential for initiation of DNA replication. EMBO J. 18, 5703–5713.

Johnston, L.H., Thomas, A.P., 1982. A further two mutants defective in initiation of the S phase in the yeast *Saccharomyces cerevisiae*. Mol. Gen. Genet. 186, 445–448.

Kihara, M., Nakai, W., Asano, S., Suzuki, A., Kitada, K., Kawasaki, Y., Johnston, L.H., Sugino, A., 2000. Characterization of the yeast Cdc7p/Dbf4p complex purified from insect cells. Its protein kinase activity is regulated by Rad53p. J. Biol. Chem. 275, 35051–35062.

Kitada, K., Johnston, L.H., Sugino, T., Sugino, A., 1992. Temperature-sensitive cdc7 mutations of *Saccharomyces cerevisiae* are suppressed by the DBF4 gene, which is required for the G1/S cell cycle transition. Genetics 131, 21–29.

Krek, W., Nigg, E.A., 1991. Mutations of p34cdc2 phosphorylation sites induce premature mitotic events in HeLa cells: evidence for a double block to p34cdc2 kinase activation in vertebrates. EMBO J. 10, 3331–3341.

Kumagai, A., Dunphy, W.G., 1991. The cdc25 protein controls tyrosine dephosphorylation of the cdc2 protein in a cell-free system. Cell 64, 903–914.

Kumagai, H., Sato, N., Yamada, M., Mahony, D., Seghezzi, W., Lees, E., Arai, K., Masai, H., 1999. A novel growth- and cell cycle-regulated protein, ASK, activates human Cdc7-related kinase and is essential for G1/S transition in mammalian cells. Mol. Cell. Biol. 19, 5083–5095.

Lee, H., Larner, J.M., Hamlin, J.L., 1997. A p53-independent damage-sensing mechanism that functions as a checkpoint at the G1/S transition in Chinese hamster ovary cells. Proc. Natl. Acad. Sci. USA 94, 526–531.

Lei, M., Kawasaki, Y., Young, M.R., Kihara, M., Sugino, A., Tye, B.K., 1997. Mcm2 is a target of regulation by Cdc7-Dbf4 during the initiation of DNA synthesis. Genes Dev. 11, 3365–3374.

Masai, H., Arai, K., 2002. Cdc7 kinase complex: a key regulator in the initiation of DNA replication. J. Cell. Physiol. 190, 287–296.

Masai, H., Sato, N., Takeda, T., Arai, K., 1999. CDC7 kinase complex as a molecular switch for DNA replication. Front. Biosci. 4, D834–D840.

Masai, H., Matsui, E., You, Z., Ishimi, Y., Tamai, K., Arai, K., 2000. Human Cdc7-related kinase complex. In vitro phosphorylation of MCM by concerted actions of Cdks and Cdc7 and that of a criticial threonine residue of Cdc7 bY Cdks. J. Biol. Chem. 275, 29042–29052.

Mimura, S., Takisawa, H., 1998. *Xenopus* Cdc45-dependent loading of DNA polymerase alpha onto chromatin under the control of S-phase Cdk. EMBO J. 17, 5699–5707.

Naeger, L.K., Goodwin, E.C., Hwang, E.S., DeFilippis, R.A., Zhang, H., DiMaio, D., 1999. Bovine papillomavirus E2 protein activates a complex growth-inhibitory program in p53-negative HT-3 cervical carcinoma cells that includes repression of cyclin A and cdc25A phosphatase genes and accumulation of hypophosphorylated retinoblastoma protein. Cell Growth Differ. 10, 413–422.

Naryzhny, S.N., Lee, H., 2001. Protein profiles of the Chinese hamster ovary cells in the resting and proliferating stages. Electrophoresis 22, 1764–1775.

Naryzhny, S.N., Lee, H., 2003. Observation of multiple isoforms and specific proteolysis patterns of proliferating cell nuclear antigen in the context of cell cycle compartments and sample preparations. Proteomics 3, 930–936.

Naryzhny, S.N., Lee, H., 2004. The post-translational modifications of proliferating cell nuclear antigen: acetylation, not phosphorylation, plays an important role in the regulation of its function. J. Biol. Chem. 279, 20194–20199.

Norbury, C., Blow, J., Nurse, P., 1991. Regulatory phosphorylation of the p34cdc2 protein kinase in vertebrates. EMBO J. 10, 3321–3329.

Nougarede, R., Della, S.F., Zarzov, P., Schwob, E., 2000. Hierarchy of S-phase-promoting factors: yeast Dbf4-Cdc7 kinase requires prior S-phase cyclin-dependent kinase activation. Mol. Cell. Biol. 20, 3795–3806.

O'Connell, M.J., Raleigh, J.M., Verkade, H.M., Nurse, P., 1997. Chk1 is a wee1 kinase in the G2 DNA damage checkpoint inhibiting cdc2 by Y15 phosphorylation. EMBO J. 16, 545–554.

Oshiro, G., Owens, J.C., Shellman, Y., Sclafani, R.A., Li, J.J., 1999. Cell cycle control of Cdc7p kinase activity through regulation of Dbf4p stability. Mol. Cell. Biol. 19, 4888–4896.

Owens, J.C., Detweiler, C.S., Li, J.J., 1997. CDC45 is required in conjunction with CDC7/DBF4 to trigger the initiation of DNA replication. Proc. Natl. Acad. Sci. USA 94, 12521–12526.

Peng, C.Y., Graves, P.R., Thoma, R.S., Wu, Z., Shaw, A.S., Piwnica-Worms, H., 1997. Mitotic and G2 checkpoint control: regulation of 14-3-3 protein binding by phosphorylation of Cdc25C on serine-216. Science 277, 1501–1505.

Raleigh, J.M., O'Connell, M.J., 2000. The G(2) DNA damage checkpoint targets both Wee1 and Cdc25. J. Cell Sci. 113, 1727–1736.

Sanchez, Y., Wong, C., Thoma, R.S., Richman, R., Wu, Z., Piwnica-Worms, H., Elledge, S.J., 1997. Conservation of the Chk1 checkpoint pathway in mammals: linkage of DNA damage to Cdk regulation through Cdc25. Science 277, 1497–1501.

Sato, N., Arai, K., Masai, H., 1997. Human and *Xenopus* cDNAs encoding budding yeast Cdc7-related kinases: in vitro phosphorylation of MCM subunits by a putative human homologue of Cdc7. EMBO J. 16, 4340–4351.

Sato, N., Sato, M., Nakayama, M., Saitoh, R., Arai, K., Masai, H., 2003. Cell cycle regulation of chromatin binding and nuclear localization of human Cdc7-ASK kinase complex. Genes Cells 8, 451–463.

Sclafani, R.A., Patterson, M., Rosamond, J., Fangman, W.L., 1988. Differential regulation of the yeast CDC7 gene during mitosis and meiosis. Mol. Cell. Biol. 8, 293–300.

Snaith, H.A., Forsburg, S.L., 1999. Rereplication phenomenon in fission yeast requires MCM proteins and other S phase genes. Genetics 152, 839–851.

Snaith, H.A., Brown, G.W., Forsburg, S.L., 2000. *Schizosaccharomyces pombe* Hsk1p is a potential cds1p target required for genome integrity. Mol. Cell. Biol. 20, 7922–7932.

Takeda, T., Ogino, K., Matsui, E., Cho, M.K., Kumagai, H., Miyake, T., Arai, K., Masai, H., 1999. A fission yeast gene, him1(+)/dfp1(+), encoding a regulatory subunit for Hsk1 kinase, plays essential roles in S-phase initiation as well as in S-phase checkpoint control and recovery from DNA damage. Mol. Cell. Biol. 19, 5535–5547.

Takeda, T., Ogino, K., Tatebayashi, K., Ikeda, H., Arai, K., Masai, H., 2001. Regulation of initiation of S phase, replication checkpoint signaling, and maintenance of mitotic chromosome structures during S phase by Hsk1 kinase in the fission yeast. Mol. Biol. Cell 12, 1257–1274.

Tercero, J.A., Labib, K., Diffley, J.F., 2000. DNA synthesis at individual replication forks requires the essential initiation factor Cdc45p. EMBO J. 19, 2082–2093.

Treuner, K., Eckerich, C., Knippers, R., 1998. Chromatin association of replication protein A. J. Biol. Chem. 273, 31744–31750.

Walter, J., Newport, J., 2000. Initiation of eukaryotic DNA replication: origin unwinding and sequential chromatin association of Cdc45, RPA, and DNA polymerase alpha. Mol. Cell 5, 617–627.

Weinreich, M., Stillman, B., 1999. Cdc7p-Dbf4p kinase binds to chromatin during S phase and is regulated by both the APC and the RAD53 checkpoint pathway. EMBO J. 18, 5334–5346.

Yanow, S.K., Lygerou, Z., Nurse, P., 2001. Expression of Cdc18/Cdc6 and Cdt1 during G2 phase induces initiation of DNA replication. EMBO J. 20, 4648–4656.

Yoon, H.J., Campbell, J.L., 1991. The CDC7 protein of *Saccharomyces cerevisiae* is a phosphoprotein that contains protein kinase activity. Proc. Natl. Acad. Sci. USA 88, 3574–3578.

Yoon, H.J., Loo, S., Campbell, J.L., 1993. Regulation of *Saccharomyces cerevisiae* CDC7 function during the cell cycle. Mol. Biol. Cell 4, 195–208.

Zou, L., Stillman, B., 1998. Formation of a preinitiation complex by S-phase cyclin CDK-dependent loading of Cdc45p onto chromatin. Science 280, 593–596.

Zou, L., Stillman, B., 2000. Assembly of a complex containing Cdc45p, replication protein A, and Mcm2p at replication origins controlled by S-phase cyclin-dependent kinases and Cdc7p-Dbf4p kinase. Mol. Cell. Biol. 20, 3086–3096.