ISAS - INTERNATIONAL SCHOOL
FOR ADVANCED STUDIES

Precise localization of the start sites of bidirectional DNA replication at the human lamin B2 origin

Thesis submitted for the Degree of
Doctor Philosophiae
(Academic Year 1998/99)

Candidate: Marta Deganuto

Supervisor: Prof. Arturo Falaschi

SISSA - SCUOLA
INTERNAZIONALE
SUPERIORE
DI STUDI AVANZATI

TRIESTE
Via Beirut 2-4

TRIESTE

CONTENTS

INTRODUCTION

DNA replication is an essential process to conserve genetic information through generations and appears to be a particularly complex phenomenon in metazoan organisms. Billions of base pairs ($3x10^9$ in the case of the human genome), subdivided in several linear chromosomes need to be copied just once, with absolute accuracy and just at the right time in the cell cycle. To accomplish this task a powerful enzymatic machinery was evolved, defined as replisome, in which the catalytic ad structural components are able to synthetize a new DNA chain using one of the parental strands as template with an incredibly low error frequency ($< 1 \times 10^{-9}$).
DNA replication does not start randomly along the genome since replisome assembly is focused at particular regions of the chromosomes, defined as origins, from where DNA synthesis specifically initiates. Origin definition in metazoans is probably a multifactorial process in which DNA sequence, chromatin structure and local chromosomal architecture establish the binding site of a protein complex that marks the origin. The replisome is then assembled onto this complex, leading to initiation of DNA replication specifically at the origin site. Thousands of replication origins are present in the metazoan genomes, and their time of activation is regulated along S-phase (Huberman and Riggs, 1968; Rivin and Fangman, 1980). Developmental signals, cell-cycle and check point regulatory networks are integrated in determining where and when replication has to start in a genome. Defining the sites of replication initiation and regulating origin activity are crucial steps in cell cycle progression, and appear to be very complex processes which still await to be fully clarified at the molecular detail in higher eukaryotes. Nonetheless, in the past few years, an increasing number of data is starting to shed light on this phenomenon, which plays a key role in ensuring that cells accomplish correctly their major task: the faithful transmission of the genetic information to the progeny.

INTIATION OF DNA REPLICATION: A GENERAL DESCRIPTION

How DNA replication works: the replication fork model

DNA is a double helix composed by two polynucleotide strands associated by the hydrogen bonds between the nitrogenous bases linked to the backbone of the chain, which consists in an alternating series of deoxyribose and phosphate molecules. The DNA strand has a defined polarity and two antiparallel polinucleotide chains are associated in the double helix. In their usual form guanines can form hydrogen bonds only with cytosines

and adenines only with thymines (Watson and Crick, 1953). The specificity of complementary base pairing is the base of the semiconservative DNA replication mechanism. DNA polymerases infact synthesize a new polynucleotide chain using one of the parental strands as template and inserting in the growing daughter strand nucleotides which are complementary to the ones present in the parental strand.

A common feature of all the DNA polymerases is that they cannot initiate synthesis of a chain of DNA from free nucleotides. They require a primer molecule bearing a free 3'-OH end, to which further nucleotides are covalently linked (Kornberg and backer, 1992). A second common feature is that they are able to synthesize DNA only in the 5' to 3' direction using the strand of opposite polarity as a template. This fact leads to a paradox since the two DNA strands are replicated simultaneously by the same enzymatic machinery proceeding in one direction on the double helix, moving in the 5' to 3' direction on one strand and in 3' to 5' the direction in the other strand (Kornberg et al., 1992). The answer to this paradox came from the discovery of the so called Okazaki fragments, which are short segments of DNA, initiated by an RNA primer, synthesised in a 5' to 3' direction and then ligated together forming a strand which grows in the 5' direction (Okazaki et al., 1968; Kornberg and Baker, 1992). Studies in *E. coli* demonstrated that DNA synthesis starts from a precise region of the genome, and proceeds in both directions on the chromosome, giving rise to structures defined as replication forks, travelling in opposite direction along the molecule. At the fork, DNA synthesis proceeds in a continuous manner on one template strand, generating the leading strand, which grows in the 5' to 3' direction. Instead the synthesis on the complementary template strand is discontinuous, generating the lagging strand, at which the ligation of Okazaki fragments results in a overall growth of the strand in the 3' to 5' direction. The site where replication initiation occurs, indicated as *Origin of Bidirectional Replication (OBR)* is also the site at wich the replication on one template strand changes from the continuos mode at the fork moving in one direction to the discontinuous mode at the fork moving in the opposite direction. The nucleotide at which the lagging strand synthesis ceases and the leading strand synthesis begins represents the *transition point (TP)* between continuos and discontinuous DNA synthesis (fig 1.1). Further studies demonstrated that the replication is bidirectional and semidiscontinuous in bacteria, in some eukaryotic DNA viruses (polyomavirus, papillomavirus and herpesvirus (Fields et al., 1996)) and in eukaryotes.

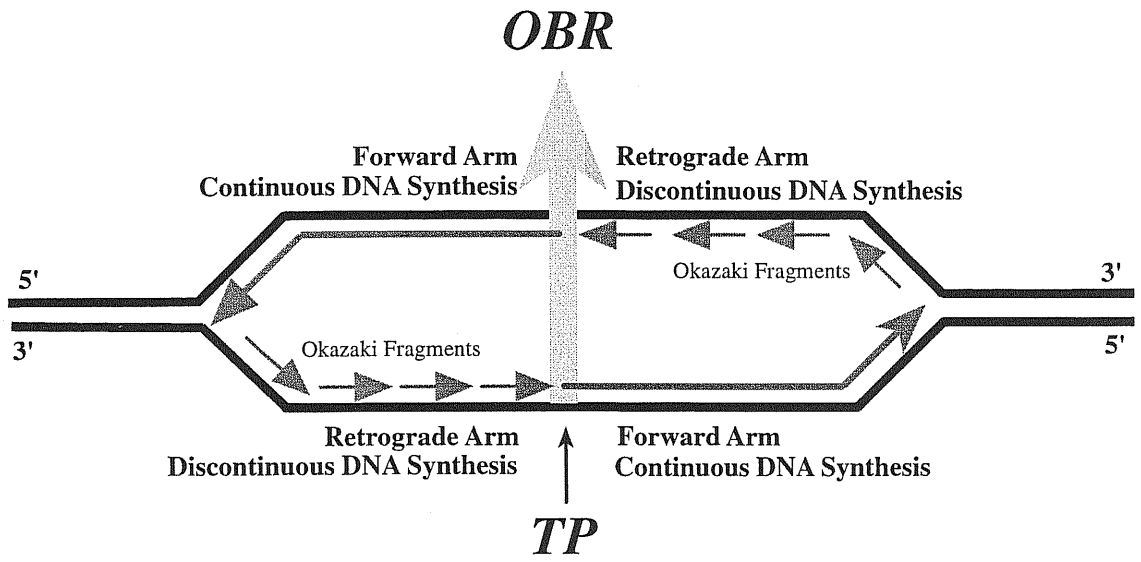

Figure 1.1: The replication fork model

How DNA replication is regulated: the replicon model

As the biochemical mechanism of DNA replication was elucidated, a model was proposed to explain how the process could have been regulated. It was known that, when fragments of bacterial chromosome were inserted in recipient cells, they were able to replicate only after integration in the host chromosome. Likewise, it was observed that, in bacteria, independent genetic elements such as phages or sex factors could loose by mutation their ability to multiply autonomously. Jacob, Brenner and Cuzin proposed the existence of a genetic element able to replicate autonomously inside a cell, defined as *replicon* (Jacob et al., 1963). It was assumed that the capacity of a sequence to behave as a replicon was dependent on the presence of at least two specific functional determinants: a) a cis-acting sequence at which replication begins defined as *replicator* and b) a trans-acting genetic element codifying for a protein that acts on the replicator determining the starting of the replication process, defined as *initiator* (fig 1.2). This model was superbly confirmed in its basic structure by the subsequent studies on bacterial and phage DNA replication, and a link was established between DNA replication and cell growth and division (Kornberg and Baker, 1992).
Later on, Bramhill and Kornberg proposed an important extension of the replicon model in which the initiator protein serves three essential functions: 1) origin recognition, 2) DNA unwinding, 3) recruitment of others replication enzymes by protein-protein interactions (Bramhill and Kornberg, 1988). In other words, replication begins at or near the sites where the protein interacts

with specific DNA sequences to initiate DNA unwinding. This process will allow the entrance of specific helicases which extend the unwound DNA tract and enable the assembly of the replication machinery. The cis-acting site where the nascent DNA polymerization actually begins can be biochemically defined as *origin of replication (ori).* Data obtained at nucleotide definition confirm the validity of the model in describing DNA replication intiation at the chromosome of *E. coli* and at the chromosomes of some eukaryotic DNA viruses (Kornberg and Baker, 1992; DePamphilis, 1993). These genomes are characterized by the presence of a single replicon per DNA molecule and the replication origin is invariably a precisely defined site.

The situation at eukaryotic genomes appears to be far more complex. *S. cerevisiae* genome is replicated from 250-400 origins whereas higher eukaryotic genomes harbour over 10^5 origins.

Each origin drives the replication of a limited region of the genome leading to the formation of tandemly arranged replication units (Huberman and Riggs, 1968, Rivin and Fangman, 1980) which are considered on first approximation functionally analogous to the bacterial replicon. Since the bacterial replicon is an independent unit, the eukaryotic counterpart is expected to replicate autonomously once cloned in an episome and inserted in a cell that produces the appropriate initiator molecule. This assumption is true for yeast in which specific genomic sequences were isolated thanks to their ability to increase the efficiency of a plasmid to transform yeast cells by a factor of 100-1000 (Stinchomb et al., 1979). These sequences, defined as ARS, (Autonomous Replicating Sequences) behave as replicators promoting the initiation of replication either when inserted in an extrachromosomal element or in a chromosomal location (Newlon and Theis, 1993). Recently the start sites of nascent DNA were mapped at the best defined ARS sequence, ARS1, in proximity of the region at which the yeast initiator complex interacts (see following paragraphs). Therefore ARS1 ori in yeast appears to be analogous to the origins in *E. coli* and SV40, and to behave in accordance with the replicon model.

It is therefore tempting to consider the replicon model as a general paradigm to explain origin function and regulation in all the organisms, from bacteria to mammals.

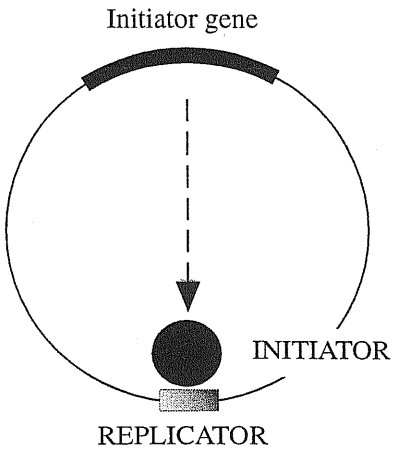

Figure 1.2: The replicon model according to Jacob Brenner and Cuzin

ORIGIN OF REPLICATION IN SIMPLE GENOMES

Anatomy of replication origin in simple organisms.

Presently, in the light of the available information on the best understood replication systems (*E. coli*, SV40, *S. cerevisiae*) we can draw the common anatomy of origins of DNA replication in simple organisms. At the SV40 origin, which was subjected to a fine molecular characterization, three elements can be recognised: a sequence bound by initiator proteins also referred as Origin Recognition Element (ORE), a DNA Unwinding Element (DUE) required for initial unwinding of the double helix and one or more binding sites for specific transcription factor(s) with an auxiliary function in origin activity (fig. 1.3). ORE and DUE constitute what is commonly referred as the ori core. Spacing, orientation and arrangement of these three components are usually critical for ori function. Sequences functionally analogous to the SV40 ORE and DUE are present also at the *E. coli* and yeast origins.
Examples of classical recognition proteins are represented by dnaA in *E.coli*, Large T-antigen in SV40 and EBNA1 in Epstein-Barr virus. All these proteins are behaving as initiators, binding a specific DNA sequence and promoting local unwinding of the DNA as predicted by the replicon model (one or more binding site for these proteins are normally present in the ORE's). On the contrary, the DUE is not a well defined sequence like ORE. The origin unwinding can in fact be favoured by different DNA sequences depending on base-stacking interactions forces. DUE sequences are generally AT reach elements with continuos thymines stretches on one strand and adenines stretches on the other. Ori auxiliary (AUX) components

consist of transcription factors-binding sites that facilitate ori core activity. The binding of transcription factors to the auxiliary elements can facilitate either binding of the initiator protein at the origin, or DNA unwinding, in alternative it may alleviate chromatin mediated repression of origin activity (see for review DePamphilis et al., 1993; Donovan et al., 1996).

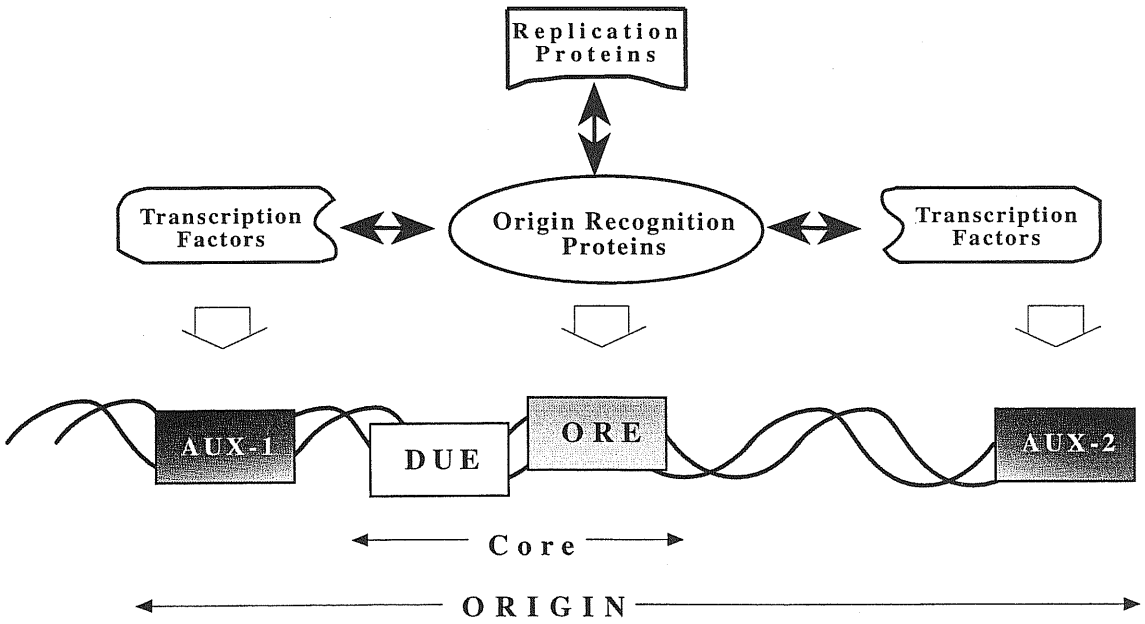

Figure 1.3: Origin structure in simple genomes.

Mapping the start sites of bidirectional DNA replication at viral chromosomes

In 1982 Hay and DePamphilis were able to map the start site of bidirectional replication, defined as the point of switching between continuos and discontinuous DNA synthesis at the SV40 chromosome. They demonstrated that the start sites for the leading strands complementary to the upper and to the lower parental strands overlap by one nucleotide (Hay and DePamphilis, 1982). The same approach was used to map at nucleotide level the start site of bidirectional DNA replication of the polyoma virus chromosome. In this case the two leading strand start sites overlap by 18 nucleotides (Hendrickson et al. 1987). The mapping of the 5' ends of nascent DNA at ori P, the origin of replication located at the episomal form of Epstein-Barr virus, revealed that the two leading strand start sites do not overlap, leaving a two nucleotide gap between them (Niller et al., 1995).
The position of the leading strand start sites on both parental strands defines the Origin of Bidirectional Replication. In the cases

above mentioned, the viral OBRs are small sequences mapping inside the genetically defined ori region, which for SV40 and polyoma virus coincides with the DUE element (Guo et al., 1991, Hay and DePamphilis, 1982; Hay et al., 1984), adjacent to the binding sites for the replication specific proteins. Since viruses rely on the cellular replication machinery to duplicate their DNA and the only required viral specific protein is the initiator that binds at ORE, it was thought that in eukaryotic cells the mechanism of replication initiation could have been the same. Due to the greater complexity of eukaryotic genomes, clarifing the mechanism of replication initiation has proven a much tougher task. The first level of complexity is due to the difference in structure between origins in prokaryotes and eukaryotic DNA viruses and origins in eukaryotic organisms. Viral origins exhibit a fixed molecular anatomy, in which the sequence elements require a specific spacing and orientation with respect of one another, initiate replication many times per cell cycle and function independently from the DNA contest in which are inserted. On the contrary, eukaryotic origins have a more complex structure, sensitive to the chromosomal contest and subjected to cell cycle regulation which ensures that origins fire only once in any S-phase. The second level of complexity is given by the multiplicity of origins per genome. Prokaryotic and eukaryotic viral chromosomes are circular molecules replicated from a single origin, whereas in higher eukaryotes the genome is subdivided in linear chromosomes, each consisting of of thousands of tandemly organized replicons which are duplicated in an ordered fashion along S-phase.

ORIGINS OF REPLICATION IN YEAST

ARS: the replicator sequence

As mentioned above, ARS elements were identified thanks to their capacity to increase the efficiency of plasmids to transform yeast cells by a factor of 100-1000. Further studies based on 2-D gel techniques (discussed in the following sections of this introduction), demonstrated that bidirectional DNA replication actually initiates specifically within the ARS elements, either when they are present on a plasmid (Brewer and Fangman, 1987; Huberman et al., 1987) or at their chromosomal location (Brewer and Fangman, 1991). Furthermore, mutations that abolish ARS function abolish also chromosomal origin activity (Deshpande and Newlon, 1992). These data conferm the hypothesis that ARSes

constitute the replicator sequences in yeast. However not all ARS elements isolated in plasmid transformation assays are functioning as origins of replication in their chromosomal environment. Thus, differently than in eukaryotic viruses, yeast replicator elements are sensitive to the contest which can determine the efficiency of origin function and the time in S phase in which an origin is used (Brewer et al., 1993).
Deletion studies demonstrated that ARS sequences have a modular structure, composed by three elements denominated A, B, C (Celniker et al., 1984), which were genetically dissected by linker substitution mutation analysis (Marahrens and Stillman, 1992). This approach proved that the A element, essential for ARS function, overlaps with the ARS Consensus Sequence (ACS), a stretch of eleven base pair conserved among all the isolated ARSes. Any mutation at this sequence suppresses origin activity. The B element can further be separated in three sub-elements: B1, B2 and B3. B2 has the characteristic of a DUE, whereas B3 is the binding site for ABF1 (Diffley and Stillman, 1988) a transcription factor that may act as auxiliary factor in enhancing origin function (Fig. 1.4). The presence of any two of the three functional B elements is sufficient for origin activity of ARS1 (Marahrens and Stillman, 1992). While A and B3 elements are conserved in sequence among ARSes, B1 and B2 are not, but function interchangeably at different ARS element, demonstrating that they play the same functional role. (Rao et al., 1994).

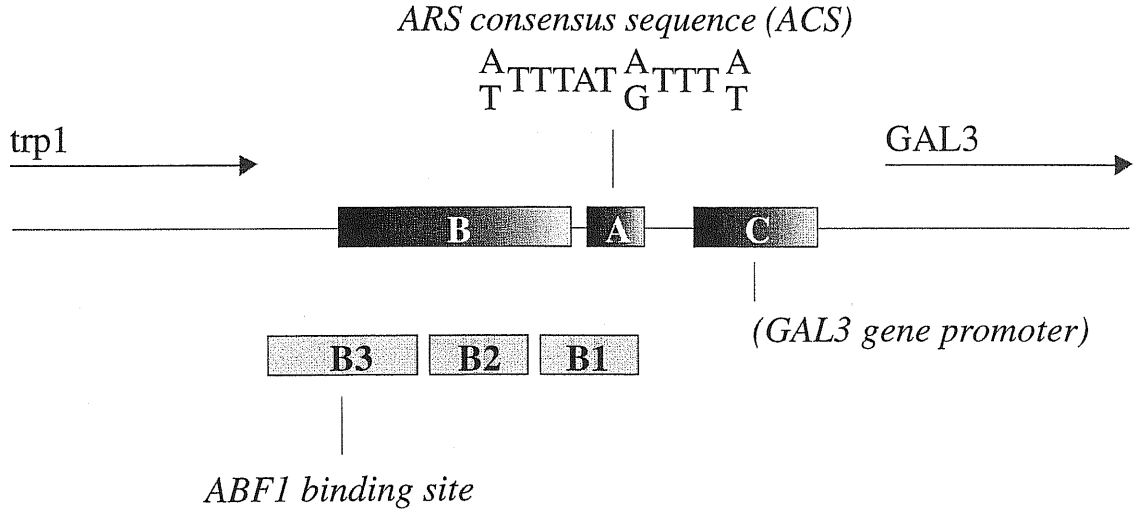

Figure 1.4: The *S. cerevisiae* Autonomous Replicating Sequence 1 (ARS1). Earlier studies showed that three regions were important for ARS activity (A, B, and C). Later, other studies demonstrated that was possible to subdivide the B region in three functionally distinct subregions.

Proteins interacting with origins in yeast: the initiator elements

ORC complex, Cdc6 protein and the pre and post-replicative complex formation

ARS elements, the replicator sequences in yeast, are bound by the Origin Recognition Complex (ORC) which represent the strongest candidate for an eukaryotic initiator. ORC was identified as a six subunit complex that specifically binds to the ACS of ARS1 *in vitro* in the presence of ATP (Bell and Stillman, 1992). The six ORC components referred to as Orc1 protein to Orc6 protein (in order of decreasing mass) are all six essential for yeast viability (Bell et al., 1993; Bell et al., 1995; Li and Herskowitz, 1993; Loo et al., 1995). ORC-DNA binding is required for the essential function of ORC, since mutations in origin sequence that reduce or eliminate origin function *in vivo* also reduce or eliminate ORC binding *in vitro*. Whereas ACS only is required for ORC to recognise DNA, ORC also protects additional sequences in the B domain of ARS elements both in *vivo* and *in vitro* (Bell and Stillman, 1992; Diffley and Cocker, 1992).
When ARS1 region was analysed by *in vivo* footprinting in different phases of the cell cycle, S phase cells and M phase cells showed a pattern of protection closely resembling the one generated *in vitro* by the purified ORC complex, whereas a more extended DNA region was protected in G1 phase (Diffley et al., 1994). Two kinds of complexes are therefore assembled at the origin, a prereplicative complex (pre-RC) during G1 that shrinks to a post replicative one (post-RC) during the following S. The enlarged footprinting observed in G1 is due to the product of *CDC6*, a gene essential in yeast. *CDC6* mutants are infact unable to enter S phase and replicate DNA (Hartwell, 1976). In absence of Cdc6 protein, the footprint in G1 closely resembled the postreplicative complex and the ORC footprint *in vitro*, suggesting the hypothesis that the assembly of the prereplication complex at the origin is a two step mechanism, in which the sequence specific binding of ORC is followed by the binding of Cdc6 (Cocker et al., 1996). Entry into S phase requires the assembly of the pre-RC, and disassembly of the pre-RC correlates with the G1/S transition (Diffley et al., 1994; Santocanale and Diffley, 1996). The mechanisms that control cell cycle progression may provide the signal to initiate DNA replication and inhibit reformation of the pre-RC after initiation (Dahmann et al., 1995; Piatti et al., 1995) ensuring once per cycle

replication. Together these results suggest that ORC is bound to yeast origins throughout the cell cycle and that in G1 phase additional replication factors assemble at origins in preparation for initiation of replication.

The MCM complex

The minichromosome mainteinace proteins (MCM) proteins are also candidates for components of the pre-RC. The MCM protein family consist of six related proteins (Mcm2 protein to Mcm7 protein) that are conserved among eukaryotes and share a 240 aminoacid region with similarity with the DNA dependent ATPases (reviewed in Chong et al., 1996). Analysis of yeast strains with mutation in the *MCM* genes indicate a role of these gene products in DNA replication (Moir et al., 1982; Gibson et al., 1990; Hennessy et al., 1991; Yan et al., 1991). The Mcm proteins can form an eterohexamer and the human Mcm-4-Mcm6-Mcm7 complex shows an ATP-dependent helicase activity *in vitro* (Ishimi et al., 1997). It may be possible that Mcm proteins form an hexameric complex *in vivo* which is loaded at the DNA acting as helicase, in analogy to what is observed for SV40 TAg and *E. coli* DNA B (Leatherwood, 1998). MCM protein complex may therefore act as the replicative helicase in eukaryotes. The loading of the MCM complex onto the chromatin depends from Cdc6 which in turn is localized at the origin by its binding with ORC (Donovan et al., 1997; Tanaka et al., 1997). A further observation that hints at the MCM complex as the replicative helicase is that the recruitment of RPA (replication protein A) at the origin is MCM dependent. RPA interacts whit polα primase tethering it at the origin, and stabilizing the single strand DNA possibly unwound by the MCM complex (Tanaka and Nasmyth, 1998).
Recently the dynamics of protein association at yeast origins were studied *in vivo* by chemical cross-link with contradictory results. Whereas some data suggest that after origin firing the MCM complex dissociates from the chromatin (Tanaka et al., 1997), other observations suggest that the MCM complex might move along the chromatin with the replication fork (Aparicio et al., 1997). This behaviour would be consistent with the idea that the MCM complex has a helicase activity. The replicative helicases are in fact expected to start out at the replication origin and then to move outward from the origin with the replication fork. This experimental approach anyhow does not allow to distinguish whether the complex is really moving or it is present already at the chromatin, becoming detectable only after the cell enters S phase

because a changing in the chromatin structure occurs while the replication fork is approaching. The above described results are also compatible with a role of the MCM complex in chromatin remodelling to facilitate the fork passage (Newlon, 1997). Alternatively the MCM complex may play a different role in initiation of DNA replication. Nevertheless, the majority of observation support the model that the MCM complex is a helicase and this hypothesis underlies the discussion of the results described in the following sections.

The regulation of pre-RC assembly and origin activation

The formation of the pre-RC complex at the origins enables them to fire in response to activation signals. The activation leads to the conversion of the pre-replicative complex into the post-replicative one, which is unable to drive origin replication initiation. The coupling of origin activation with the inactivation of the initiator complex ensure that origins fire once and only once per cell cycle. Therefore pre-RC activation is subjected to a thight regulation both by specific factors and by the cell-cycle regulatory cyclin dependent kinase (CDK) complexes.

Specific regulation of pre-RC formation

Once ORC and Ccd6 have assembled Mcms onto the DNA, the resulting pre-RC complex is inactive until S-phase. Replication initiation requires the action of the kinase composed by the catalytic Cdc7 subunit and the activating Dbf4 protein which were found to interact specifically at yeast origin of replication (Dowell et al., 1994). It was demonstrated that specific alterations in *S. cerevisiae* Mcm5 protein allow the origins to fire without the activity of Dbf4-Cdc7 kinase (Jackson et al., 1993, Hardy et al., 1997a). The simplest interpretation of this finding is that Dbf4-Cdc7 overcomes an inhibitory effect of Mcm5 protein on DNA replication regulating the activity of the MCM complex by phosphorylation. Both *in vivo* and *in vitro* data support this hypothesis (Lei et al., 1997, Sato et al., 1997; Brown and Kelly, 1998). Further studies demonstrated that Cdc7 protein is needed throughout S phase for firing of individual replication origins. Limiting Cdc7 functions results in fact in a slow progression through S phase rather that in a delayed entry in S phase (Bousset et al., 1998; Donaldson et al., 1998).

A furher essential replication factor is the Cdc45 protein, known to interact genetically with many pre-RC components, including ORC, Mcm proteins and Cdc7 (Hopwoood and Dalton, 1996; Zou et al., 1997; Hardy, 1997b). Cdc45 acts late in the replication initiation process because it binds to chromatin only after activation of S-phase CDKs which probably alter pre-RC structure allowing Cdc45 to associate at the origins; after binding it seems to move outward from the origin with the replication fork (Aparicio et al., 1997). Cdc45 co-precipitates with Mcm proteins (Hopwood and Dalton, 1996; Zou and Stillman, 1998), interacts with chromatin only after MCM complex is bound (Zou and Stillman, 1998) and is required in conjunction with Dbf4-Cdc7 to trigger replication initiation (Owens et al., 1997). Since the function of Dbf4-Cdc7 appears to be phosphorylation of Mcm proteins, it is likely the Cdc 45 also acts on them.

Cell cycle regulation of origin firing

The proteins described in the previous paragraphs are probably all components of the prereplicative complex, and their assembly at the origins is cell cycle regulated through the action of the cyclins in conjunction with their active subunits: the cell cycle dependent kinases (CDKs). A critical role, in regulating origin activation is played by S-CDKs whose activity depends on the cyclin regulatory subunits Clb 5 and 6 (Diffley, 1998). There are data indicating a direct interaction between the ORC complex and the main cyclin-CDKs in the fission yeast *S. pombe* which could mean that the cell cycle regulation machinery is tethered at the origins by contacts with the origin binding complex (Leatherwood et al., 1996). One of the main targets of the CDKs activity is the Cdc6 protein. Ccdc6 is synthesized *de novo* at the beginning of each G1 and binds to ORC at the origins. Proceeding through G1, Cdc6 protein becomes a target for S-CDKs mediated phosphorylation, which leads to the inhibition of the preRC formation (Piatti et al., 1996; Detweiler and Li, 1998). Evidence obtained in *S. cerevisiae* and *S. pombe*, indicates that S-CDKs phosphorylation targets the protein to ubiquitin-mediated proteolysis (Drury et al., 1997; Kominami et al., 1997, Jallepally et al., 1997; Jallepalli et al., 1998).
The other major target of cyclin-CDKs regulation is the MCM complex. Evidence obtained in *Xenopus* and mammalian systems indicate that Mcm proteins are phosphorylated at the onset of the S-phase by S- CDKs, and phosphopylation reduces their chromatin binding ability (Kimura et al. 1995; Todorov et al., 1995; Lei et al., 1996, Hendrickson et al., 1996). Taken together, these findings

indicate that the S-CDKs prevent the formation of the pre-RC, blocking the loading of MCM complex. They can act directly on the complex components or affecting Cdc6 protein, from which the association of MCM complex at the origin depends.
From these observations appears that the pre-RC can be assembled onto the chromatin at the exit of M phase and in early G1, when the activity of CDKs is low. During the progression through G1, CDK activity increases, blocking pre-RC reassembly and consequent origin refiring during S, G2 and M phases. Thus the origin firing requires a period of low CDK activity which allows the origin to acquire the competence to replicate, followed by a period of high CDK activity which allows competent origins to fire. Because the second step in initiation of DNA replication also blocks the first, origins can fire only once per cell cycle.

Regulation of origin firing along S-phase

The firing of origins in yeast and in the higher eukaryotes is time regulated along S-phase. In *S. cerevisiae*, the time of origin firing appears to be established in late mitosis or G1 in response to the chromosomal contest. A temporal pattern of origin activation seems to take place at the chromosome, since origins are changing from early to late firing as one proceeds towards the telomeres (Raghuraman et al., 1997). Recently, the molecular mechanism that regulates the timing of origin activation begun to be elucidated. The S-phase promoting cyclins Clb5 and Clb6 in association with the Cdc28-CDK, cooperate in defining the time of firing, being late origin activation dependent from an active Clb5 (Donaldson et al., 1998; Diffley, 1998). The association of the of RPA with the origins provides also a mean by which origin firing can be regulated along the S-phase. In fact early and late origins differ not in the timing of Mcms recruitment, but in the timing of RPA recruitment. The recruitment of RPA at the origins is regulated by Rad53, the main S-phase checkpoint protein in yeast (Tanaka and Nashmith, 1998), which arrests replication in response of DNA damage or fork block (see for review Weinert, 1998). These results suggest that the origin firing programme is subjected to checkpoint controls, and are strengthened by the observation that, in presence of hydroxyurea, which causes replication fork stalling, only early replication origins can fire. If RAD53 is mutated, the block of late origins firing is removed. Rad53 protein is thus involved in monitoring successful fulfilment of the programme of DNA replication during S-phase, and in co-ordinating a controlled arrest if problems arise (Shirahige et al., 1998; Santocanale and Diffley, 1998).

Regulation of replication initiation in yeast: a comprehensive model

The data reported above, obtained mainly in *S. cerevisiae*, can be organized in a model which constitutes a sort of a paradigm in the description of the complex of interactions between replicator sequences and initiator elements in a eukaryotic system (fig. 1.5). ORC complex is constitutively expressed and bound to the origin throughout the cell cycle. In the period of low CDK activity that follows the exit from G1, the unstable Cdc6 protein binds to ORC . A recent report demonstrates that the Cdc6 proteins shares significative homologies with eukaryotic and prokaryotic clamp loader proteins, responsible for the loading of complexes around the DNA molecule, in reactions that requires the hydrolysis of ATP (Jonsson and Hubscher, 1997). The Cdc6 protein contains the two Walker box motives of a classical ATP/GTP-binding site (Walker et al., 1982) and can hydrolyze these nucleotides (Zwerschke et al., 1994) possibly promoting the load the MCM hexamer at the origin. Probably at this stage the Dbf4-Cdc7 complex activates the MCM complex which in turn will locally unwind the DNA. After the loading, the association of the MCM complex with the chromatin is stable even in absence of Cdc6 protein (Donovan et al., 1997; Hua and Newport, 1998), which is phosphorylated by CDKs and degraded during the progression toward S-phase. One of the final triggers determining origin firing is probably the association of Cdc45 protein at the pre-replication complex, which converts it in a preinitiation complex (preIC, Zou and Stillman, 1998). Experiments in Xenopus egg extracts have demonstrated that Cdc45 protein physically interacts and colocalizes with polα primase. One attractive model suggested by Baker and Bell to explain these findings (Baker and Bell, 1998) is that Cdc45 is a scaffold protein holding together the MCM helicase with the replicative polymerases much like in *E. coli* the tau protein links the DNA B helicase with the leading and lagging strand polymerases. In this scenario, Cdc45 and Dbf4-Cdc7 would be interdependent on each other for function, because Dbf4-Cdc7 would release or activate the MCM helicase and Cdc45 would connect the helicase with replicative polymerases and might further activate its function (Leatherwood, 1998). Another protein responsible for the recruitment of the polα primase at the preIC complex is RPA, whose association at the origin region depends on the activation of S-CDKs, Dbf4-Cdc7 kinase and Mcm proteins (Tanaka and Nasmyth, 1998). The chromosomal contest would influence the time of association of RPA to the origins, which, together with the concerted action of the S-phase promoting cyclins might define the time of origin firing in S-phase.

Once polα primase is associated to the pre-IC, and the synthesis of the leading strand starts, additional replication and processivity factors are recruited at the fired origin. The leading strand synthesis is catalyzed by polδ, which binds processively the DNA thanks to the interaction with the ring shaped protein PCNA, loaded on the chromatin by the RF-C factor (Baker and Bell, 1998). The lagging strand synthesis is promoted by polα primase which initiates the synthesis of RNA primed Okazaki fragments. It is not known whether ORC stays bound to single strand origin DNA. An hypothesis is that ORC is probably transiently displaced from its binding site in the process of unwinding and could rebind as soon as its binding site is duplicated (Bielinsky and Gerbi, 1999). The identity of a processive helicase that catalyzes DNA unwinding as the fork proceeds, is still an open question. The data indicating that the MCM complex moves together with the replication fork point to it as to the replicative helicase in yeast (Aparicio et al., 1997). Since the data about the helicase activity of the human MCM complex suggest that it is weakly processive (Ishimi et al., 1997), an alternative hypothesis is that MCM acts locally at the origin and the further unwinding is taking over by another, yet uncharacterized, processive helicase.

From the data presented above it appears evident that regulation of DNA replication initiation is the convergent target of cell-cycle and checkpoint controls. The reported studies elucidated how the information coming from the regulatory pathways that control cell growth and proliferation are integrated to trigger origin activation once per cell cycle, at the appropriate time in S-phase.

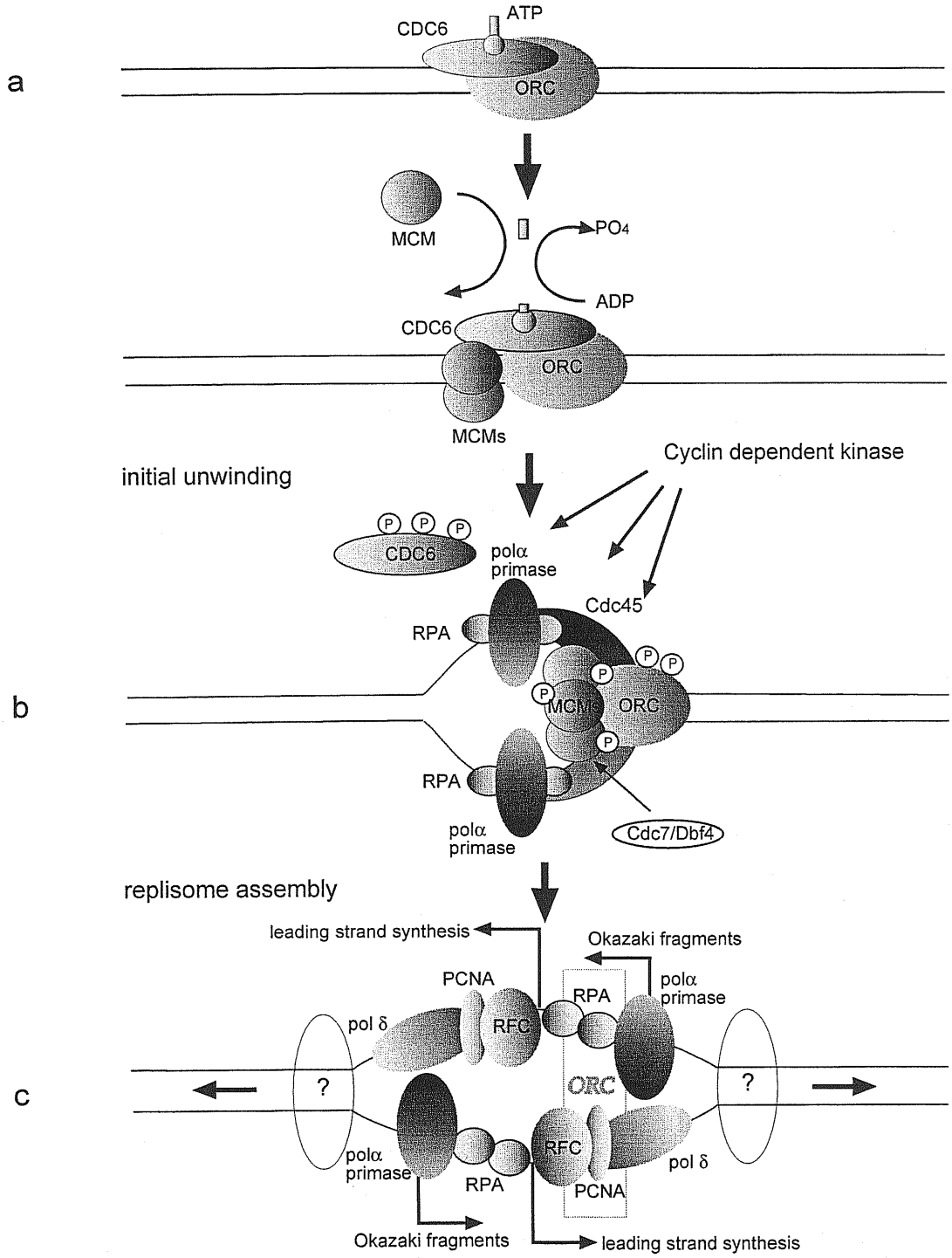

Figure 1.5: Model of replication initiation events taking place at the origins in yeast.

ORC and Cdc6 are associated at the origins in early G1 forming the pre-replicative complex. Cdc6 loads the Mcm proteins at the origin in an ATP dependent reaction (a). Progressing along the cell cycle, the cyclin dependent kinases phosphorylate Orc proteins and Cdc6 displacing it from the complex. The phosphorylation of Mcms by the cyclin dependent kinases and by the specific kinase Cdc7/Dbf4 triggers the initial unwinding of DNA at the origin (b). Cdc45 protein associates with Mcms and recruits polα at the origin. RPA stabilizes the unwound DNA and is also involved in the recruitment of polα, which initiates the leading strand synthesis. The parental double strand is further unwound in both directions (arrows), and single stranded DNA is stabilized by RPA (c). Upon the assembly of additional proteins (e.g. polδ, RF-C, and PCNA) constituting the replication machinery (replisome assembly), leading strand synthesis becomes processive on one parental strand, and Okazaki fragments are initiated by polα primase on the complementary strand. The helicase which triggers the progression of the fork is indicated by a question mark. Possibly, the Mcm complex acting at the origin site is further moving with the replication fork catalyzing the DNA unwinding as the fork proceeds (Aparicio et al., 1997).

Mapping the start site of bidirectional DNA replication at the ARS1 ori

Recently the start site of bidirectional replication was mapped at nucleotide resolution at the ARS 1 origin in *S. cerevisiae* either when it is located on a plasmid or in its chromosomal position. This result was accomplished thanks to some critical technical improvements that allow to increase the sensitivity of detection of the 5' ends of nascent DNA much above the threshold reached by Hay and DePamphilis in the mapping of SV40 start sites of DNA synthesis. Susan Gerbi and collaborators (Gerbi and Bielinsky, 1997) developed a method defined as RIP (Replication Initiation Point) mapping (sketched in fig. 1.6) based on the isolation of replicative DNA structures (forks and bubbles) on benzoylated-naftoylated DEAE cellulose. The replication intermediates are treated with λ-exonuclease, an enzyme which digests DNA from its 5' end. Since the enzyme is not active on RNA, the newly synthesized nascent DNA bearing an RNA primer, is protected from its action. Therefore the λ-exonuclease treatment destroys the aspecifically sheared DNA and preserves selectively the nascent DNA, since it presents the unique feature of having an RNA primer at its 5' end.
The authors selected specific primers in the ARS1 origin region and performed multiple rounds of primer extension using the λ-exonuclease treated DNA as a template. The primer is extended along the template up to the junction with the RNA primer, which represent the start site of DNA synthesis. Due to the fact that the source of nascent DNA are asynchronously growing cells, there is polymorphism in the population of the isolated molecules depending on how many Okazaki fragments have been ligated to a replication start site for any given molecule, at the moment of the analysis. Therefore the extension products show a variety of lengths. There are no Okazaki fragment ends in the region of leading strand synthesis and so, in the case in which the primer is complementary to the leading strand, the nucleotide position at the 5' end of the smallest detectable extension product marks the transition point between continuos and discontinuous DNA synthesis, and the more 5' site of leading strand initiation.
RIP mapping technique was applied to analyze the switching point between continuos and discontinuous DNA synthesis at the ARS1 ori either on a plasmid or in its chromosomal location.
When the analysis was performed on the plasmidic ARS, the switching point between continuos and discontinuous DNA synthesis on the upper nascent strand overlaps the switching point

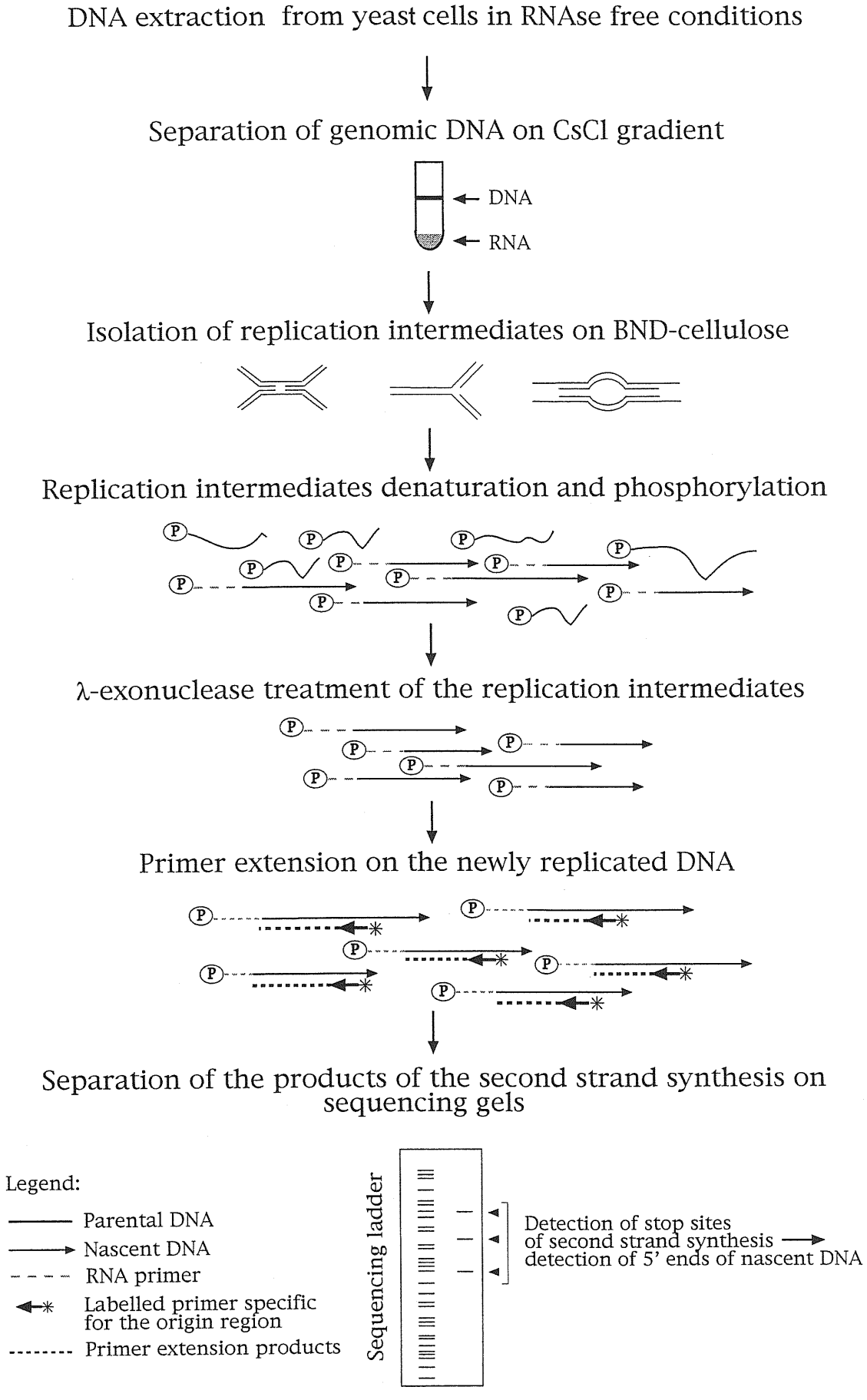

Figure 1.6: Scheme of the RIP mapping approach (Gerbi and Bielinsky, 1997).

on the lower nascent strand by 18 nucleotides (Bielinsky and Gerbi, 1998). Several other initiation events are taking place at the origin, but it was not determined in this study if they were due to Okazaki fragment synthesis or to polymorphism in the selection of the leading strand start site.
The same analysis applied on the chromosomal ARS1 gave a somewhat different result: the points of transition between continuous and discontinuous DNA synthesis are not overlapping but separated by two nucleotides, and initiation events occur less frequently at the origin area. Experiments performed on a yeast strain with a mutation in the DNA ligase, demonstrated that the nucleotide at the 5' end of the smallest extension product represent the start site of the leading strand synthesis. The other extension products are due to the elongation of the specific primer along a leading strand to which an increasing number of Okazaki fragments have been ligated, and disappear if the ligation doesn not take place. From this experiment it is possible to conclude that, at the chromosomal ARS1 ori there is only a single leading strand initiation site within the population of replicating yeast cells (Bielinski and Gerbi, 1999).
Both in the plasmid and in the chromosome the origin of bidirectional replication is located between the B1 element, protected *in vivo* by the ORC complex, and the B2 element, which is the putative DUE sequence of ARS1 (Lin and Kowalski, 1997). These results are in perfect agreement with the replicon model, which postulates that the replication starts in the region interacting with the initiator specific proteins (ORC in the case of yeast). The discrepancies between the results obtained in the plasmid and in the chromosome can be explained considering the different chromatin structure in chromosomal and episomal ARS1 copy. (Venditti et al., 1994). A possible explanation might be that the chromatin structure in the ARS1 flanking region is homomorfic in the chromosome, being constrained by other features of the chromosomal architecture, but etheromorphic in the episome, where such constraints might not exist. These differences in chromatin structure between plasmid and chromosome could affect the unwinding kinetics within ARS1. It is thus possible that more sites are available for leading strand initiation in the plasmid defining a broad origin of bidirectional replication, whereas only one is selectively exposed in the chromosome, in which the transition point between continuous and discontinuous DNA synthesis coincides with the OBR.
ARS1 represents the symplest eukaryotic ori known (Newlon and Theis, 1993). As will be discussed in the following sections, ori structure in metazoans appears to be more complex. However, given the high degree of phylogenetic conservation among the

proteins (e.g. ORC) (Dutta and Bell, 1997) crucial for replication initiation and among the components of the catalytic machinery (Baker and Bell, 1998), it is conceivable that the principle of replication initiation at the chromosomal ARS1 ori might also apply to chromosomal origins in metazoans. To date substantial evidence is availble that origins in metazoans are confined to discrete sites in the genome (DePamphilis, 1999). It remains to see whether the precision that governs the initiation at ARS1 is evolutionary conserved.

REPLICATION ORIGINS AT METAZOAN CHROMOSOME

Replicator sequences in higher eukaryotes

At any cell cycle metazoan cells have to accomplish the formidable task of accurately replicating their large genomes structured in several chromosome pairs each containing thousands of tandemly organized replicons. Each replicon harbours an origin which is expected to be analogous to the replication origins described in detail for yeast and animal viruses and to behave in accordance with the replicon model. Conversely the idea that replicators exist also in metazoan was dealt a serious blow by the experiments of Harland and Laskey (1980) which showed that any kind of DNA molecule is able to replicate efficiently and once per cell cycle when inserted in *Xenopus* egg extracts.
A possible explanation of this result might be that in *Xenopus* extracts the requirement for origin specification is overcome by the excess in initiator proteins present in egg extracts, and may reflect the ability of any DNA to interact aspecifically with replication complexes. Moreover in the early stage of the development, transcription does not take place. Chromatin is free from interactions with transcription factors and transcriptional machinery, which possibly tend to confine origin activity at the intergenic spaces. In this condition the replication initiator complex can be established anywhere in the genome. On the contrary, in the physiological chromosomal contest, chromatin repression may allow interaction with the replication complexes only in specific sites defined by a number of factors, including DNA sequence at the origin, chromatin organization and transcriptional state, and nuclear structure. In this scenario it is not surprising that the search for genetic elements which specify the initiation of replication by means of ARS assays were extremely inefficient and/or indiscriminate in metazoan (Gilbert, 1998). In fact, if the chromosomal contest plays a critical role in specifying origins of

DNA replication, they cannot be defined in an episomal contest in which the information carried by the chromosomal architecture is lost.

Chromosomal mapping of origin of DNA replication

Due to the lack of a functional assay analogous to the ARS assay in yeast to study origins of replication *in vivo* at metazoan chromosomes, origin identification has been addressed indirectly, thanks to the peculiar features of the DNA at the replication initiation site. The methods that were developed to map origins rely essentially on

a) analysis of the DNA structure at the replication initiation site, (replication intermediates)
b) analysis of the leading-lagging strand polarity,
c) chromosomal mapping of the nascent DNA strands (Fig. 1.7).

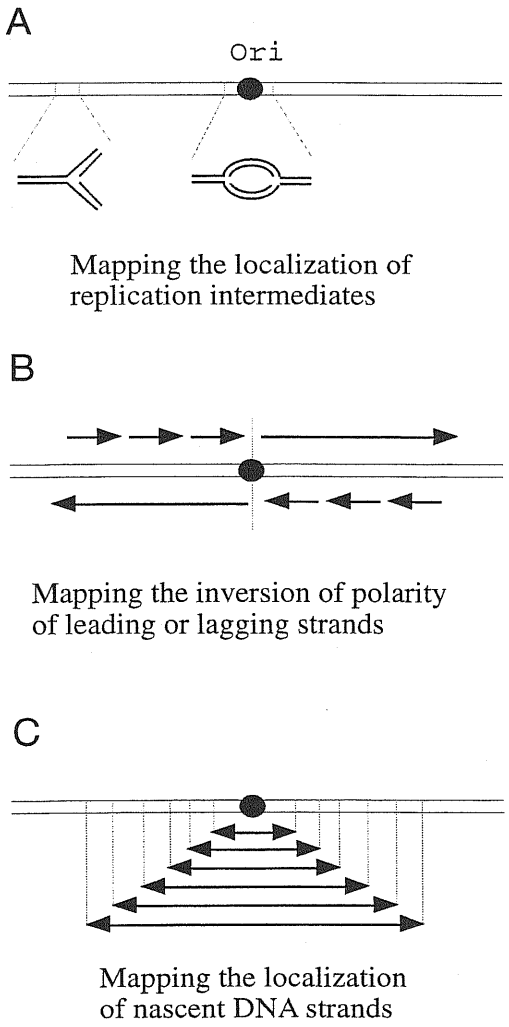

Figure 1.7: Techniques available for origin mapping in metazoan cells

Structural analysis of the replication intermediates

Two complementary two dimensional (2-D) gel electrophoretic mapping techniques were developed to obtain an intimate view of replication intermediates. Both techniques rely on the unique migration pattern of replication intermediates in agarose gels and utilize defined molecular probes to examine these intermediates in the genomic region of interest.

Neutral-neutral 2D gels

In the neutral-neutral 2-D gel method (Brewer and Fangman, 1987), replication intermediates are separated in the first dimension on the basis of the molecular mass (using low gel concentration), which varies from 1n (unreplicated) to 2n (fully replicated) and in the second dimension on the basis of both mass and shape (using high gel concentration). Characteristic arcs are traced by fragments containing either a single fork, an origin or a termination structure. The replication intermediates in the region of interest are then visualized by hybridization with appropriate radioactive probes (fig 1.8).

Neutral-alkaline 2-D gels

The neutral alkaline 2-D gel method (Nawotka and Huberman, 1988) determines the direction of the fork movement through a chromosomal locus. A restriction digest of replicating DNA is separated by mass in the first dimension. The nascent strands are then released from their templates with alkali and separated by size in the second dimension. The direction of fork movement through any given restriction fragment can be determined by sequentially probing the fragment with adjacent probes. An origin-proximal probe can recognize fragments of a broad range of sizes emanating from an origin, tracing a complete diagonal, whereas an origin distal probe detects only the longer nascent strands (fig. 1.9). Logically it can be deduced that a region between two divergent moving forks must contain an origin.

2-D gel based origin-mapping techniques have been successfully used to map initiation events in yeast, protozoa and slime mold to loci of 0.3 to 1kb (DePamphilis, 1997). By this mean, it was shown that some, but not all, ARS elements in yeast are working as chromosomal replication origins and that the same sequences

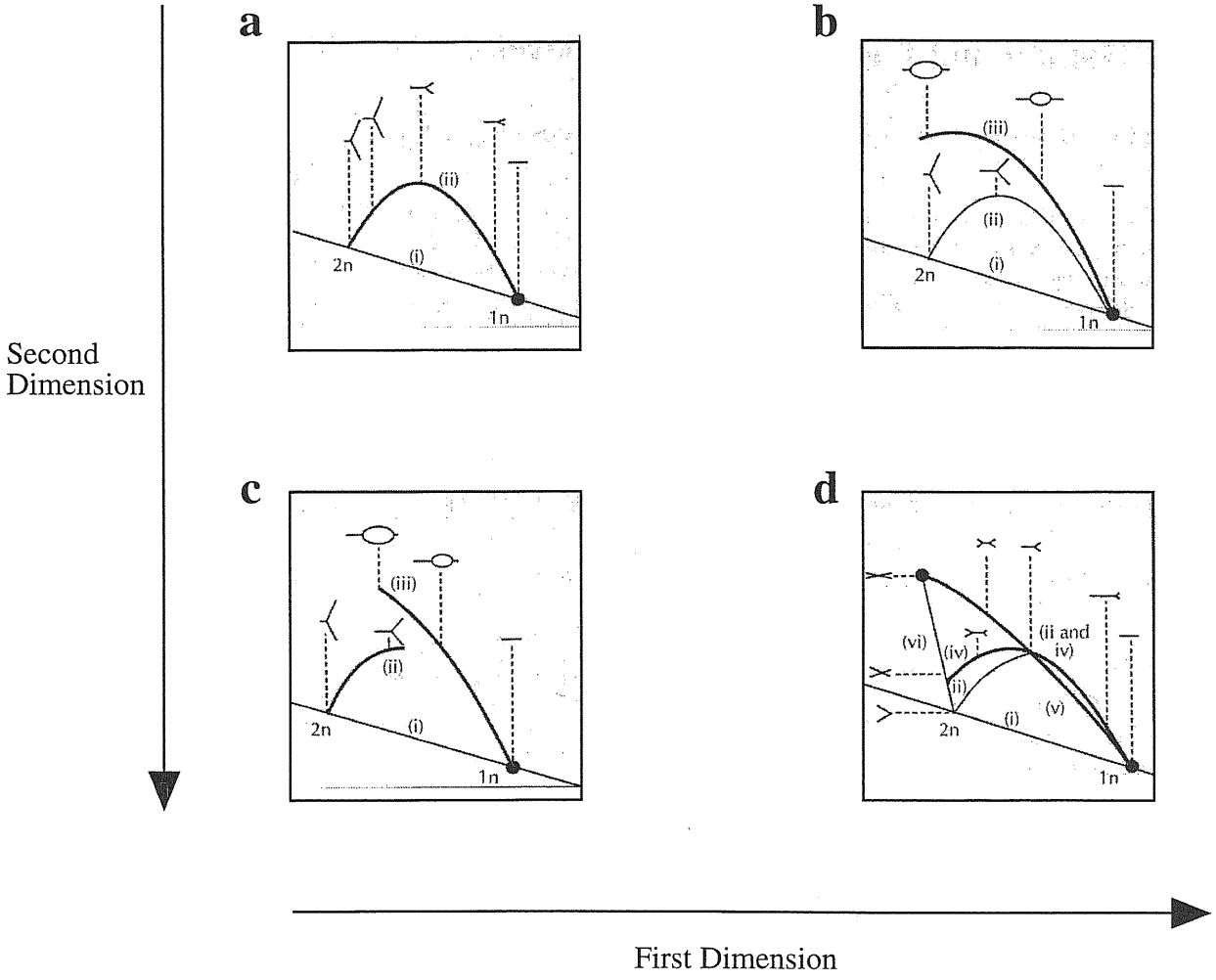

Figure 1.8: The principle of the neutral/neutral 2-D gel replicon-mapping technique. Idealized autoradiographic images obtained when a digest of replicating DNA is hybridized to probes for fragments that contain different intermediates. Curve (i) represents the diagonal of non-replicating fragments from the genome as a whole. a) A complete single fork or simple Y-arc (ii) from a fragment replicated passively from an outside ori. b) A fragment with a centered ori traces a curve (iii) that arches over the single fork arc (curve ii). c) A fragment containing an off-centered ori traces the beginning of a bubble arc (curve iii) when the replication fork crosses one restriction site. d) A fragment containing two approaching forks. If they meet asymmetrically or symmetrically, curves iv and v are obtained, respectively. If there is a fixed terminus in a fragment, the collected X-shaped structures would result in a concentrated spot somewhere on curve vi, on which any recombination structures present would fall too.

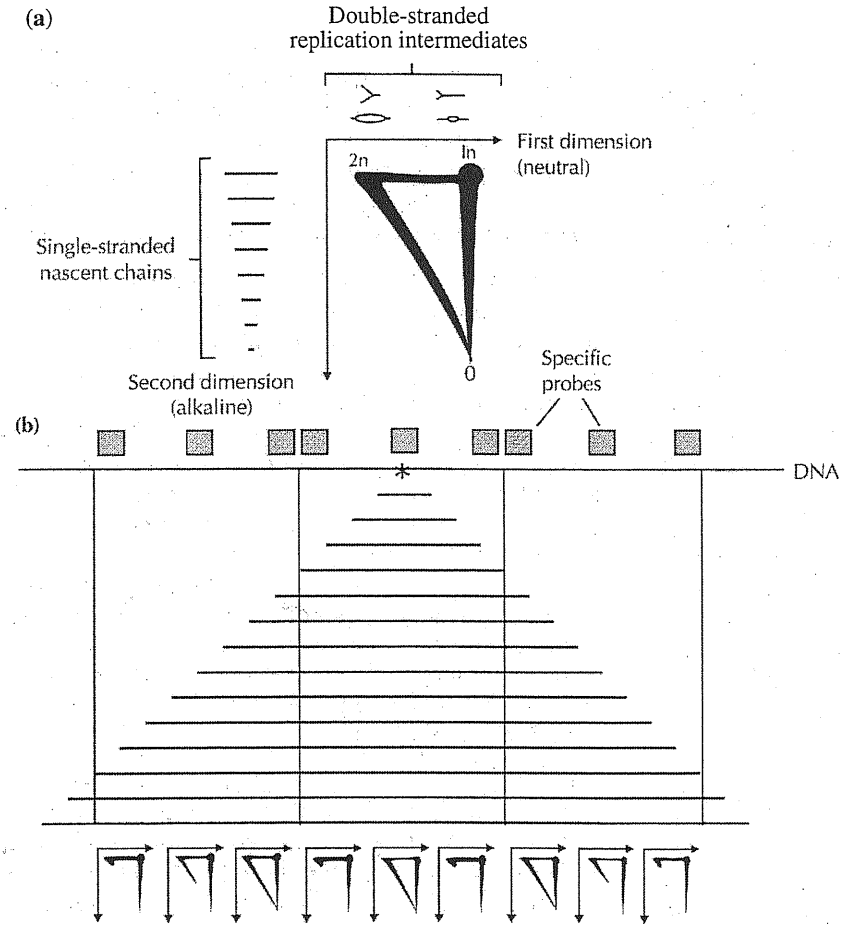

Figure 1.9: Principle of the neutral/alkaline 2-D gel replicon-mapping technique. (a) Idealized patterns of migration of double-stranded replication intermediates in the first (neutral) dimension and of single stranded nascent chains in the second (alkaline) dimension. (b) Three adjacent restriction fragments are marked by the vertical lines. An ori is positioned in the central fragment (asterisk), and nascent strands of different sizes are indicated. The panels below show the autoradiograms that would be obtained when each of the indicated probes (shaded boxes) is used to detect its cognate restriction fragment.

required for ARS activity are also required for chromosomal origin activity (Deshpande and Newlon, 1992; Newlon et al., 1993). These methods appear to work best with well defined, efficient origins that are active each cell cycle. In more complex organisms, S-phase origins appear to be less efficient (bubble arcs are superimposed on strong fork arcs in neutral-neutral 2-D gels), which complicates the quantification of the data. Moreover the hybridization pattern are usually more complex to analyze than those expected from the model, suggesting the possible presence of unusual DNA structures responsible for the unexplained hybridization spots. Applications of these methods to cells from mammalian, frogs and flies has revealed broad "initiation zones" in which replication bubbles appear to be distributed throughout DNA regions of 6 to 55kb and in which replication forks appear to migrate with similar frequency in both directions, suggesting that initiation can occur at many different sites whithin these zones at similar frequencies (Dijkwel and Hamlin, 1995).

Such conclusion is affected from the fact that the application of these techniques to the analysis of higher eukaryotic genomes presents some critical drawbacks. Due to the low sensitivity of 2-D gel techniques, they have been applied to highly amplified loci or to single copy loci utilizing synchronization procedures, tipically with aphidicolin. Aphidicolin, an inhibitor of the processive eukaryotic DNA polymerases, blocks the progression but not the initiation of DNA synthesis (Wang et al., 1991). When the fork is stalled at the origin, reinitiation phenomena may occur, leading to multiple branched initiation intermediates and to branch migration of nascent DNA (Dijkwel et al, 1991). To rule out this potential artifacts, mimosine, another synchronizing agent which is thought to block cells in late G1 before origin firing (Lalande et al. 1990, Lin et al., 1996), was used in studies that still confirm the delocalization of origins (Dijkwel and Hamlin, 1992, Wang et al., 1998). However the need of cell synchronization, imposing the use of metabolic blockers, might also in the case of mimosine induce a non physiological state in which initiation events could be perturbed. The use of of the 2-D gel approach in metazoan cells requires as well the enrichment of replication intermediates by DNA purification procedures that include nuclear matrix isolation (Dijkwel et al., 1991) and benzoylated-naphtoylated(BND)-cellulose chromatography (Levine et al., 1970). These procedures can easily introduce breaks in the bubbles structures, eventually masking an initiation event.

Analysis of the leading-lagging strand polarity

Since DNA polimerase synthesized DNA only in the 5' to 3' direction an Origin of Bidirectional Replication is the site where continous synthesis on the leading strand and discontinous synthesis on the lagging strand switch template and invert polarity. This switch can be detected following the distribution of selectively isolated Okazaki fragments or leading strands.

Okazaki fragment polarity

Labelled Okazaki fragments are hybridized with cloned, strand specific DNA from the region of interest. The hybridization signal is restricted to the clones that represent the lagging strand template and disappears at the switching point where the same parental strand becomes the leading strand template (Burhans et al, 1990). The main drawback of this methods is that it requires synchronization and harvesting of the cells in very early S phase, followed by cell permeabilization to allow radioactive deoxyribonucleotides to enter the nucleus and label Okazaki fragments synthesized in the origin area.

Imbalanced DNA synthesis

In the complementary approach, emetine, a protein synthesis inhibitor (Roufa and Marchionni, 1982) is used, since it is proven to block selectively the lagging strand synthesis, leaving unaffected the leading strand synthesis (Burhans et al., 1991). Labelled isolated nascent leading strands are hybridized with strand specific clones and also in this case the OBR is indicated by the change in the hybridization signal, at the point in which continuos DNA synthesis switches template. The main drawback of this approach is the lack of knowledge on the exact mode of emetine action and whether such treatment, which considerably alters cellular metabolism, also affects the physiological pattern of DNA replication initiation.

Chromosomal mapping of nascent DNA strands

Several mapping strategies were developed based on isolation of nascent DNA strands and determination of their genomic localization. The rationale of these techniques is that the nascent DNA emanating from an origin consists in a population of molecules centred at the origin itself. The analysis of nascent DNA is performed either by hybridization with genomic fragments

spanning the putative origin containing region or by PCR analysis with specific primers. The representation of defined markers within a certain genomic area in samples of short DNA fragments is inversely correlated with their distance from the origin. The stronger is the signal given by a marker, the closer it is to the origin.

Early Labelled Fragment Assay

In this approach, cells synchronized at the G1/S border are permeabilized and released in S phase in presence of radioactive deoxyribonucleotides. The first labelled DNA is used as a probe to assay by hybridization an array of genomic fragments spanning the putative origin region (Burhans et al., 1986a). Alternatively, specific initiation sites in nuclei isolated from G1 phase mammalian cells can be radiolabelled in *Xenopus* egg extract (Wu et al., 1997). In both cases, radiolabelled replication bubbles can be prevented from migrating far from the origin region by introducing crosslinks with psoralen before releasing cells in to S phase (replication bubble trap; Anachkova et al., 1989). This method is limited by the low resolution that allows to map OBRs only in amplified regions of the genome upon cell synchronization. Therefore, also in this case the analysis might be affected by the non-physiological state imposed by the synchronization procedure. Nevertheless, more recently, a modification of this method that implies the gradient purification of BrdU substituted nascent DNA followed by hybridization with labelled fragments covering the region of interest, allowed to map an origin of replication at the human ribosomal gene cluster, without the need for cell synchronization and permeabilization (Yoon et al., 1995).

PCR on nascent DNA strand

To overcome the resolution problems arising from the early labelled fragment assay, a PCR analysis method was developed in which nascent DNA is amplified by specific primers scattered along the region of interest; the closer a PCR marker is to the origin, the higher is its abundance relative to the others (fig.1.10). This approach allowed to map origins in single copy region without the need of cell synchronization (Vassilev et al., 1990; Vassilev and Johnson, 1990). The reliability of this method is based on the precise quantification of the relative abundance of the PCR markers selected for the analysis, and it has been greatly improved by the development of the competitive PCR technique (Diviacco et al., 1992). In this approach, a fixed amount of nascent DNA is coamplified with a known amount of competitor, an artificial

molecule that shares the same primer recognition sites of the genomic target, but migrates differently in electrophoresis. The evaluation of the nascent DNA/competitor ratio for several amplified fragments spanning a defined genomic area, allows to define the enrichment in abundance of the marker closer to the origin in comparison to the distal ones and to narrow the mapping of the OBR in the region centred around it. In the past years this technique became the method of choice to map origin of replication at high resolution at mammalian chromosomes (Giacca et al., 1994; Pelizon et al., 1996; Delgado et al., 1998; Kobayashi et al., 1998; Araujo et al., 1999).

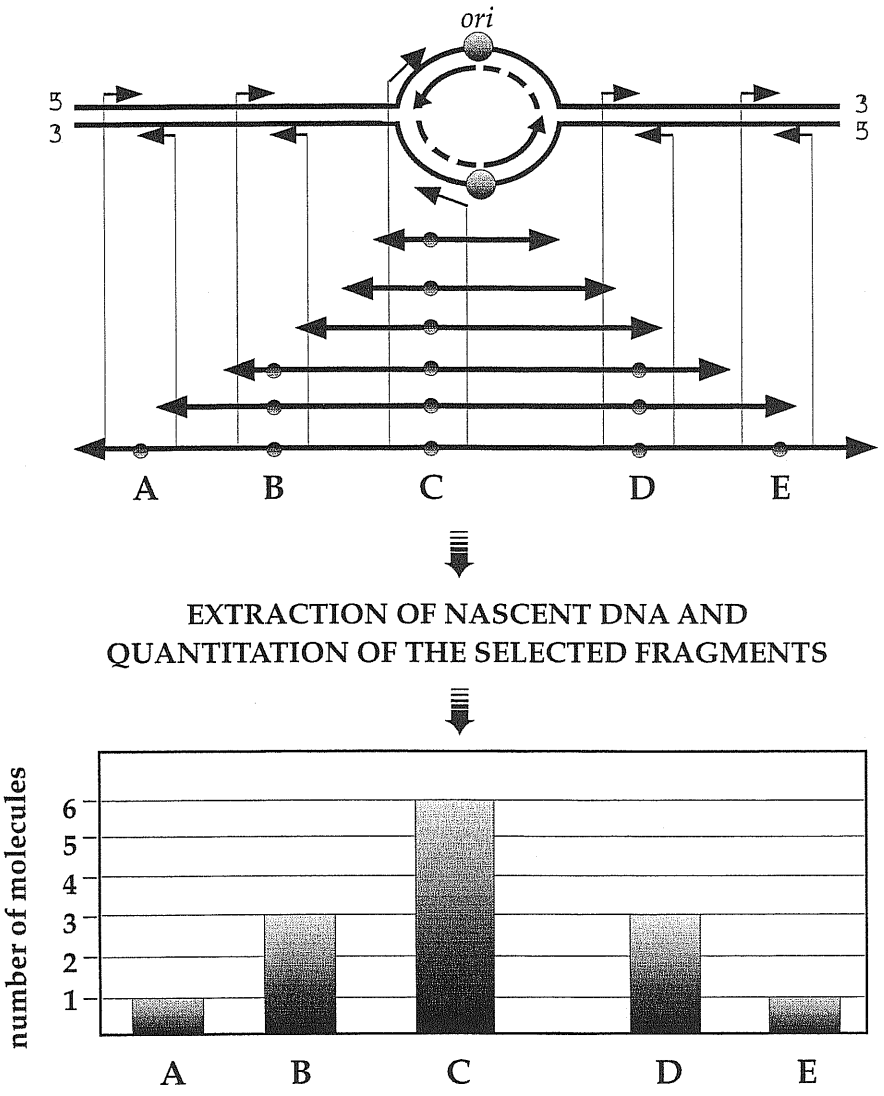

Figure 1.10: Ori- mapping procedure based on nascent DNA strand analysis. The short fragments of nascent DNA can be isolated and separated from the bulk DNA. Whithin this population of newly sinthetyzed DNA molecules, the abundance of selected markers scattered whithin a genomic region (A through E) identifies the active origin of replication as the fragment that shows the higher enrichment.

Metazoan origins: puzzles and paradigms

Thanks to the methods described above, a number of origins were mapped at mammalian chromosomes (see table 1, from Todorovic et al., in press). While all the data are consistent with bidirectional replication involving classical replication bubbles and forks, a complex and sometimes contradictory view of replication origins has emerged. While some studies conclude that most initiation events occur at specific sites analogous to those found in *S. cerevisiae* and eukaryotic viruses, other studies conclude that most initiation events are distributed throughout large DNA regions, with no preference for one site over another. This paradox is best illustrated by the studies on the mammalian rRNA and dihydrofolate reductase (DHFR) gene regions where several different mapping methods have been applied at the same genomic loci.

When nascent DNA is isolated for analysis, the data suggest that most initiation events occur at specific DNA sequence which constitutes the origin of bidirectional replication in metazoans. In the human rRNA gene region, the earliest labelled DNA fragments have been identified (Coffman et al., 1993; Gencheva et al., 1996) and the growth and relative abundance of nascent DNA chains have been determined (Yoon et al., 1995; Gencheva et al., 1996; Gogel et al., 1996). These studies revealed two primary initiation sites: a 1 to 6kb locus upstream of the rRNA gene promoter that is >10-fold more active than distal sites and a 3kb locus, with weaker activity, downstream to the 3'end of the gene. The major site appears to be conserved among species (Hernandez et al., 1993).

In the DHFR gene region in Chinese hamster ovary (CHO) cells, the earliest labelled DNA fragments have been identified (Anachkova and Hamlin, 1989, Burhans et al., 1986a, Burhans et al., 1986b; Heintz and Stillman, 1986; Leu et al;, 1989), and the region of switch between continuos and discontinuous DNA synthesis has been determined by measuring both Okazaki fragments and leading strand biases in hybridization experimets with strand specific probes (Handeli et al., 1989; Burhans et al., 1990, Burhans et al., 1991). Moreover the growth and relative abundance of nascent DNA chains have been determined (Vassilev et al. 1990, Pelizon et al., 1996). All these studies agree in revealing two OBRs, a strong one (oriβ) mapped within a 0.5 to 3kb sequence located 17 kb downstream of the 3' end of the DHFR gene, and a weaker one (oriγ) located 22 kb further downstream.

In contrast, when replication intermediates are isolated from mammalian cells, fractionated by 2-D gel electrophoresis and then hybridized with sequence specific probes, initiation events appear to be distributed almost randomly throughout large DNA regions

(initiation zones). Neutral-neutral 2-D gel electrophoresis has detected replication bubbles throughout the 31 kb intergenic spacer in rRNA gene repeats and the 55 kb region between the DHFR gene and the 2BE2121 gene located downstream, while neutral-alkaline 2-D gel analysis has detected equivalent numbers of forks traveling in both directions (Little et al., 1993; Dijkwel and Hamlin, 1995).

How can this discrepancy be reconciled? Recently a 12 kb region encompassing oriβ was studied by nascent DNA abundance analysis via competitive PCR (Kobayashi et al., 1998). Two peaks of nascent DNA abundance were shown: one corresponding to the oriβ locus as previously defined by early labelled fragment assay, nascent strand abundance and leading-lagging strand polarity switch studies; the other one, less prominent, mapped in a region located 5kb downstream to oriβ and defined as oriβ'. These results demonstrate that the DHFR initiation zone defined by 2-D gel analysis harbours at least three primary initiation sites (oriβ, β' and γ) probably surrounded by multiple secondary sites that are not detected efficiently in nascent strand abundance analysis, but are visible in 2-D gel electrophoresis. The DHFR origin appears to behave in accordance with the "Jesuitic model", which stipulates that, in a chromatin region permissive for replication initiation, there are many potential start sites of leading strand synthesis among which some are preferred, in accordance with the jesuitic dictum: "Many are called, but few are chosen" (DePamphilis, 1999). This interpretation brings in closer agreement the 2-D gel data in favour of the existence of a broad initiation zone in which replication may start randomly at any site with some preference for the oriβ region (Wang et al., 1998), and the data that localize a maximum of origin activity in a 2 kb region encompassing oriβ.

The discrepancy in the results given by these techniques may relay in fact on the peculiar features of the techniques themselves which are looking at the same phenomenon from two different points of view.

In the first place, the methods based on the selective isolation of nascent DNA, either by metabolic labelling or by size fractionation, are detecting exclusively origins active at the moment in which the analysis is performed. In contrast, 2D gel patterns represent a photograph of the total population of DNA structures that exist within a particular DNA region at a particular time, regardless of the activity of these structures at the time in which the photograph was taken. In fact at metazoan origins, which are used at low frequency at any cell cycle, neutral-neutral 2-D gels analysis reveals invariably a weak bubble arc (the locus is actively replicating) superimposed to a strong fork arc (the locus is passively replicated from an external ori). A second consideration

is that, if the experiments are performed in synchronized cells, nascent DNA is analyzed in the first minutes after the releasing of the cells in S-phase, whereas 2-D gel analysis are commonly carried out in cells released for 1 to 2 hours in S-phase, when the signal from replication bubbles is greatest. The nascent DNA analysis methods and the 2-D gels techniques may thus reveal selectively a different population of replication events, the former focusing on the most active origin at the beginning of S phase, and the latter focusing on the most prominent structure in S-phase (DePamphilis, 1997).
It has to be considered that the 2-D gel analysis is performed on replication intermediates purified by nuclear matrix isolation and BND-cellulose purification. If origins are not colocalized with the nuclear matrix, a defined time or distance will elapse before replication forks associate with matrix. This phenomenon may exclude the isolation of "newborn" replication bubbles in which the origin has just fired, causing a reduction in the origin mapping resolution. This fact could explain the failure of 2D gels in detecting primary initiation zones at mammalian chromosomes, considering also that, when nuclear matrix isolation was omitted, as in the case of the analysis of an origin of DNA amplification in *S. coprophyla*, 2-D gel analysis revealed a primary initiation zone (Liang et al., 1993).

Genetic analysis of OBRs in Metazoans

To test the hypothesis that oriβ contains a metazoan replicator sequence, the region was subjected to genetic analysis. The ori β region was knocked out by homologous recombination and replication studies at the mutated locus demonstrated that replication activity is not affected, and replication initiates at the DHFR intergenic region with the same efficiency and the same timing than in wild type cells. These data argue against the existence of essential replicator elements in metazoans and might be explained in view of the Jesuitic model, hypothesizing the existence of a hierarchy of non essential replicators distributed at frequent intervals throughout the intergenic region. When the replicators that are somehow preferred are deleted, the replication starts from secondary initiation sites in a region in which the chromatin structure allows the assembly of the metazoan prereplication complex.
A different situation seems to take place at the human β-globin origin, the best genetically characterized metazoan origin so far. This origin was mapped by imbalanced strand synthesis in a 2kb

region located upstream to the β-globin gene. A deletion at the origin region, as occurs in haemoglobin Lepore syndrome, cancels bidirectional DNA synthesis at this site and leads to the passive replication of the region from a locus positioned upstream (Kitsberg et al., 1994). Thus, in contrast to the results obtained for the DHFR locus, these data indicate that the 2kb β-globin origin region has the features of a replicator sequence, since its deletion correlates with the loss of replication activity at the locus, and represent the first genetic proof of the existence of genetically definable origin of replication in animal cells. The entire β-globin genomic cluster is controlled by a region localized 50kb upstream of the cluster defined as locus control region (LCR) (Forrester et al., 1990). A naturally occurring deletion at the LCR changes deeply the chromatin organization at the locus, preventing β-globin gene expression in erythroid cells, shifting replication timing from early to late S phase and making all the locus DNAseI resistant (Tuan et al., 1985; Crossley and Orkin, 1993; Engel, 1993). Analysis of this deletion mutant demonstrated that initiation at the β-globin locus is abrogated and that the region is passively replicated from a downstream located origin. Further studies showed that the origin of replication from which the β-globin region is passively replicated in haemoglobin Lepore cells is also shut off, probably because it represents a cryptic initiation site activated if the β-globin locus is inactivated but affected from the change in the chromatin structure imposed by the LCR deletion (Aladjem et al., 1995). A similar phenomenon seems to take place also at the DHFR locus. In this case, if a 13.5 region located at the 3' end of the DHFR gene is deleted, the replication initiation activity in the intergenic space between DHFR and the 2BE2121 gene is completely abolished (Kalejta et al., 1998). These data indicate that metazoan replicators are sensitive to the chromosomal contest, defined by control elements which act at a distance and modulate the activation state of entire chromatin domains (Wolffe, 1998).
Recently a genetic dissection of the putative replicator present at the human β–globin locus was performed. A site specific recombination strategy was used to transfer a 8kb long region harbouring the β-globin origin in other chromosomal locations. This approach is alternative to the ARS assay which is not feasible in metazoans. Putative replicators are transferred in other location of the same genome so that the information related to genome complexity, nuclear structure and general chromosomal organization are not lost as it happens when the putative OBR regions are cloned in an episomal contest. The transferred β-globin is able to sustain replication at ectopic chromosomal location, independently from the presence of the locus control region,

probably because the integration occurs at open chromatin sites. In view of these results, replicator activity in mammalian cells may be defined in terms of the elements needed to confer initiation of DNA replication within a particular chromosomal contest. Deletion analysis of the transferred region allowed to identify a core replicator surrounded by auxiliary regions that might serve to facilitate DNA unwinding or loading of the metazoan pre-replication complex at the core replicator (Stillman, 1993; Diffley, 1996). These findings strongly support the hypothesis that replicator elements exist in metazoans. If this is true, the origin of bidirectional replication, defined biochemically as the sequence at wich DNA replication starts, is supposed to colocalize with the genetically defined replicators, in analogy to what has been observed in simpler organisms.

Factors that influence determination of origins in metazoans

At the moment is difficult to address which are the cis-elements that are defining an origin in metazoans, since no significative common features are shared by already mapped origins, and just few of them are defined at high resolution (see table 1 for origins mapped at mammalian chromosomes). Drawing homologies among them awaits the setting of new analysis methods that will allow the screening and the comparison of a large number of origins.
The complexity in metazoan origin definition relies on the fact that, differently than in the case of simple organisms (yeast and viruses), the site specificity of the initiation of replication is not dictated by the sequence but is a multifactorial process. Several parameters are involved in origin establishment in metazoans, among which key roles are played by nuclear organization, chromatin structure, transcriptional activity of the region, and epigenetic modifications at the origin sequence (DePamphilis, 1999).

Nuclear organization

Nuclear structure is constituted by a network of filaments, that are defined as nuclear matrix, nuclear scaffold or nucleoskeleton in dependence to their biochemical and functional properties (Wolffe, 1998). Several studies demonstrated that nascent DNA is preferentially associated with this network, which creates nuclear sub-compartments where replication forks are condensed, defined

TABLE I. Origin mapping studies in mammalian cells

Region studied	Organism	Method	Main conclusions	References
β-globin gene	Human	Imbalanced DNA synthesis PCR of nascent DNA strands	Origin mapped in a 2 kb fragment	(Kitsberg et al., 1994; Aladjem et al., 1995)
Lamin B2 gene	Human	Quantitation of nascent DNA strands	Origin mapped in a ~500 bp region at the end of the gene, at a site in which cell-cycle regulated protein-DNA interactions occur	(Giacca et al., 1994; Kumar et al., 1996)
HSP70 gene	Human	PCR of nascent DNA strands	Origin region mapped in a ~400 bp fragment in the promoter region	(Taira et al., 1994)
rRNA genes	Human Mouse Rat	Analysis of nascent strand abundance 2-D gel electrophoresis	Multiple initiation sites in the 31 kb non transcribed spacer, with a preferred region upstream of the rDNA transcription start site	(Little et al., 1993; Yoon et al., 1995; Gencheva et al., 1996)
DNA-methyltransferase gene	Human	Quantitation of nascent DNA strands	Origin mapped within the gene	(Araujo et al., 1999)
c-myc gene	Human	PCR of nascent DNA strands	Origin mapped within a 2.5 kb region upstream of c-myc gene	(McWhinney and Leffak, 1990; Vassilev et al., 1990; Beberich et al., 1995)
DHFR gene	Hamster	Analysis of Okazaki fragments Earliest replicated DNA Imbalanced DNA synthesis, PCR of nascent DNA strands, 2-D gel electrophoresis	Initiation zone in 55 kb (2-D gel studies) Three fixed origins (ori-β, β' and γ) localized at 0.5-8 kb resolution (all the other studies)	(Heintz and Hamlin, 1982; Anachkova and Hamlin, 1989; Leu et al., 1989; Handeli et al., 1990; Burhans et al., 1990; Vassilev et al., 1990; Burhans et al., 1991; Gilbert et al., 1995; Kobayashi et al., 1998)
Ribosomal protein S14 (RSP14) gene	Hamster	PCR amplifcation of nascent DNA strands Analysis of Okazaki fragments	Origin mapped in ~2.5 kb region overlapping the gene	(Tascheva and Roufa, 1994)

Region studied	Organism	Method	Main conclusions	References
Rhodopsin gene	Hamster	Earliest replicated DNA	Origin contained in a 10 kb region overlapping the gene	(Gale et al., 1994)
APRT locus	Hamster	Imbalanced DNA synthesis	Origin mapped in a ~6 kb region	(Handeli et al., 1989)
CAD gene	Hamster	Analysis of Okazaki fragments	Origin localized in a 5 kb region in the CAD transcriptional unit	(Kelly et al., 1995)
APRT gene	Hamster	PCR quantitation of nascent DNA strands	Origin localized the promoter region encompassing a CpG island	(Delgado et al., 1998)
GADD gene	Hamster	PCR amplification of nascent DNA strands	Origin localized the promoter region encompassing a CpG island	(Delgado et al., 1998)
TK gene	Hamster	PCR amplification of nascent DNA strands	Origin localized the promoter region encompassing a CpG island	(Delgado et al., 1998)
GNAI3 and GNAT2 genes locus	Hamster	2D-gel electrophoresis Quantitation of nascent DNA strands	~1.7 kb origin in the intergenic region between GNAI3 and GNAT2	(Toledo et al., 1998)
Aldolase B locus	Rat	Hybridization of nascent DNA	Origin mapped in a 1 kb region containing the AldB promoter	(Zhao et al., 1994)
ADA gene (early S)	Mouse	Analysis of Okazaki fragments	11 kb origin region at 28.5 kb 5' of ADA gene	(Carrol et al., 1993)
ADA gene (late S)	Mouse	PCR of nascent DNA strands	1-2 kb origin region at 150 kb 3' of ADA gene	(Virta-Pearlman et al., 1993)
Ig heavy chain locus	Mouse	Semiquantitative PCR	Origin mapped in a ~600 bp region within the enhancer	(Ariizumi et al., 1993)
Ig heavy chain locus	Mouse (non B cells)	2-D gel electrophoresis	One/multiple origins within 90kb downstream of the 3' regulatory region	(Ermakova et al., 1999)

as "replication foci" or "replication factories" (Hozak et al., 1996). The foci formation is essential for replication and dependent on the presence of an intact nuclear lamina. In fact, when the *Xenopus* egg extract is depleted from the membrane fraction (Laskey and Madine, 1996), or the integrity of the nuclear lamina is perturbed (Goldberg et al., 1995; Ellis et al., 1997; Spann et al., 1997) replication cannot take place. Probably the lamina interacts with the nuclear matrix allowing the formation of chromatin domains permissive for origin establishment and assembly of replication factories (Zhang et al., 1996).

Chromatin structure

Chromatin organization plays a pivotal role in defining origins and in regulating their firing during S-phase. Early firing origins are located in the active transcribed regions of the chromatin whereas late firing origins are more often mapped at inactive chromatin domains (DePamphilis, 1996). Moreover, sequences acting at a distance determining the functional state of entire chromosomal regions, like insulators and Locus Control Regions, have a profound influence on the state of origins. In *Drosophila* the ACE sequence is required for high level of amplification initiating at the nearby major origin of replication oriβ. When ACE and oriβ were transferred on another chromosomal location, the efficiency of the amplification was very low, and was increased by the inclusion in the construct of an insulator element, protecting the origin from the chromatin mediated repression taking place at ectopic locations (Lu and Tower, 1997). As it has been already discussed, LCR controls the activity of the human β globin origin, differentially regulating transcription, replication and time of origin firing in cells of different lineages (Aladjem et al., 1995).
Thus, it appears that the variable activity of replication origins observed at different chromosomal positions is due largely to difference in chromatin organization throughout the genome. This difference may be mediated by the changes in protein composition of the complexes interacting with DNA. For example, histone H1 can reduce the frequency of initiation in *Xenopus* egg extracts, by limiting the assembly of prereplication complexes at sperm chromatin (Lu et al., 1998), and such changes in chromosome structure can determine which initiation sites will be used. Incubation of condensed chromosomes from metaphase arrested hamster cells in *Xenopus* egg extract elicits the assembly of nuclei and initiation at a novel site (oriδ), while oriβ, which is normally used in the physiological S phase, is repressed (Lawlis et al., 1996). Thus, changes in chromatin structure during the metazoan

development and during the cell cycle modify the pattern of active origins in a genome, presumably by affecting their accessibility to initiation factors.

Transcription

Transcription is not required for replication since the first S phase in mammalian embryos and the first rounds of replication in embryos of flies, frogs and fishes occur in absence of transcription. Nevertheless, there is a relation between transcriptional activity and replication. In fact, 45% of the origins mapped up to now are located in transcriptional promoters or enhancers (see table 1), suggesting that transcription factors may play a role in metazoan origins as they do in many viral and some yeast origins, but so far there are no clear evidences supporting this hypothesis (DePamphilis, 1999). It is also known that transcription at a locus is related to replication timing of that locus, since all the origins located near actively transcribed genes are early replicating (Biamonti et al, 1992, Kitsberg et al., 1994, Michaelson et al., 1997). From these evidences is possible to conclude that the positive correlation between transcription and replication reflects an open state of the chromatin, more accessible to the assembly of both replication and transcription complexes, and not a direct link between the two phenomena. In fact, in the case of β-globin, in non erythroid cells, replication occurs also in absence of transcription, even if the time of firing is shifted from early to late S (Hatton et al., 1988; Epner et al., 1988). An alternative way in which transcription might affect replication is in preventing association of the replication complex at origins located in a transcribed region, thus confining active origins in the intergenic space between transcription units (Pan et al., 1995; Haase et al., 1994)
Anyhow, at the moment, a role of transcription in determining metazoan replication origins remains speculative.

DNA methylation

In *E. coli*, DNA methylation plays a critical role in regulating DNA replication and preventing origin refiring (Kornberg and Baker, 1992).
The role possibly played by methylation in origin determination and/or regulation in metazoans is controversial. Previous studies indicated the presence of large, densely methylated islands in which all cytosines were methylated, regardless of their

dinucleotide composition (Tasheva and Roufa, 1994b). Although these results were subsequently shown to be artefactual, it is true that some origins in mammals are highly methylated (Rein et al., 1997 a and b). In fact both oriβ at the DHFR locus and the origin mapped inside the rhodopsin gene in hamster lay adjacent to a 350 bp sequence in which the frequency of ^{m}CpG is 10 times above the average. It was also observed that all the CpG dinucleotides within 2 Kb regions containing these two OBRs remain methylated throughout the cell cycle. Moreover, there are evidences that in cells defective in methylation, site specific initiation at ori β is lost (Rein et al; 1999). Recently it was shown by competitive PCR that nascent DNA is enriched in sequences coming from the CpG islands located at the promoter of three hamster housekeeping genes (Delgado et al., 1998). CpG island-like fragments are also enriched in a population of short nascent strands from human erythroleukaemic cells, suggesting that they constitute a significant fraction of the endogenous, early replicating origins (Delgado et al., 1998). Since CpG islands are G+C rich regions free of methylation, these results seems not in agreement with the observation of the methylation state at the DHFR and rhodopsin origins. Most probably, several origin families exist, with features that might be partly common and partly peculiar for each of them. In this contest, the absence of methylation might be restricted to origins located at the promoters of housekeeping genes, while other origins might require methylation to be activated.

Linking origin selection with cell cycle progression and metazoan development

The complexity of the origin determination in metazoans is probably due to the fact that origins are not fixed structures labelled throughout the cell cycle by an origin marking complex analogous to ORC in yeast. Footprinting data demonstrated that origin interacting proteins in metazoan are not present in M phase (Abdurashidova et al., 1998) when a thorough chromatin reorganization is taking place, erasing all the pre-existing chromatin structures. Origins needs to be reestablished at each cell cycle and the multiplicity of parameters that enters in specifying them, allows these structure to be extremely plastic. This plasticity is essential to respond to cell cycle regulation and to modify origin activity during the establishment of the different developmental programs through which a multicellular organisms is formed (Francon et al., 1999). Example of developmentally regulated origins are β-globin ori and the ori located at the IgH heavy chain

locus. In erythroyd cells, where the globin gene cluster is expressed, the β-globin origin fires early in S-phase, whereas in non erythroyd cell lines, the time of firing is shifted towards late S-phase (Epner at al., 1988; Hatton et al., 1988). In B-cells expressing the IgH gene, replication initiates from a multitude of origins scattered throughout the gene cluster, whereas in non-B cells the entire locus was shown to be replicated by a single replication fork and forms a transition region between early and late replication domains (Ermakova et al., 1999).
Recently, some insights were gained about the mechanisms through which the origins are establish in G1. CHO cell nuclei isolated at various time points in G1 were injected in *Xenopus* egg and initiation of replication at DHFR locus was monitored. Specific replication initiation at ori β was achieved only after a certain point in G1 defined as Origin Decision Point (ODP). Pre ODP nuclei replicated DNA efficiently but without any detectable preference for the DHFR origin (Wu and Gilbert, 1996). This means that pre ODP nuclei have already acquired the potential to replicate (replication licensing) but have not yet specified where replication should initiate. It was recently shown that passage through ODP takes place before the restriction point, after which origins are committed to fire even in absence of mitotic stimuli, it is independent from the presence of serum mitogens but it is abolished by protein kinase inhibitors. This suggests that origin specification is governed by a mitogen independent protein kinase pathway that drives the cells to the selection of which of many potential chromosomal sites may function as an origin in the upcoming S-phase (Wu and Gilbert, 1997).
It is possible to draw a parallel between what happens during the cell cycle in somatic cells and what happens during the early stage in the development of several animal species. In *Xenopus*, the early embryonic cell cycles rapidly oscillate between S and M phase and replication starts at random sites spaced 9-12 Kb. At a defined point in the blastula stage the initiation of replication becomes focused to specific chromosomal sites (Hyrien et al., 1995). One possible explanation for this developmental switch is that high concentrations of replication factors in the early embryo allow relaxed origin specification and these factors become titrated as DNA is replicated. An alternative explanation is that, in analogy to what happens at the ODP, changes in nuclear structure and chromatin organization which take place at the midblastula transition, focus replication to specific sites.

Considering all the data, the most inclusive working hypothesis is that metazoan replication initiation complex is focused to specific sequences by features of chromosome structure. The ODP and the

blastula stage of *Xenopus* development could represent the establishment of a nuclear structure that favours specific interaction between DNA and initiation complexes at particular chromosomal sites. If this model is correct, replicators are established in metazoans after ODP. At this moment specific interaction between replicators and initiatior proteins may occur defining specific origins of bidirectional replication in the genome.

Initiator proteins interacting with DNA replication origins at eukaryotic chromosomes

If replicator sequences exist in metazoan chromosomes, they have to be bound to the initiator protein complex that mediates origin activation. The main proteins involved in replication initiation in yeast, Orc and Mcm subunits and Cdc6 protein, are all highly evolutionary conserved in metazoans (Gavin et al., 1995, Kearsey and Labib, 1998, Quintana et al., 1997, 1998; Saha et al., 1998; Tugal et al.,1998) and an increasing number of evidences are indicating that their function in origin activation is also conserved throughout evolution. In the following paragraphs I will focus my attention on the metazoan homologues of the yeast prereplicative complex components, pointing on their interaction with the cell cycle regulation machinery.

Role of the nuclear membrane in concentrating replication proteins

An intact nucleus is generally required to observe initiation of DNA replication in all the *in vitro* systems described so far: *Xenopus* egg extracts (Wu et al., 1997), human cells S phase nuclear extract (Krude et al., 1997) and yeast S phase nuclear extract (Pasero et al., 1997), demonstrating that nuclear membrane and internal nuclear structure play one or more critical roles in regulating eukaryotic DNA replication.
The main function of the nuclear membrane is to regulate the access of replication factors to DNA substrates as was observed in the classical cell fusion experiments of Rao and Johnson (1970). The authors demonstrated the existence of a mechanism that prevents replication after S phase. They showed that when a G1 phase cell is fused with an S phase cell, the G1 nucleus replicates prematurely. This implies that the cytoplasm of the S phase cell contains an activator of replication and that the G1 nucleus is capable of responding to it. On the contrary, when a G2 phase cell

and an S phase cell are fused, the G2 nucleus does not replicate. Apparently G2 nuclei cannot respond to S phase replication signals because they lack an essential replication component that "licence" them for replication, which is excluded from the nucleus by the nuclear membrane. Further experiments performed in *Xenopus* extracts demonstrated that, if nuclear membrane is permeabilized, G2 nuclei can initiate replication (Leno et al., 1992). These data lead to a model in which the nuclear membrane promotes replication concentrating essential replication factors in the nucleus. After origin activation, these factors are somehow inactivated and origins cannot fire until after mitosis, when nuclear membrane breakdown allows new, active licensing factor to enter the nucleus and associate with chromatin (Blow and Laskey, 1988). An alternative model is that nuclear membrane acts concentrating inhibitory factors which prevent origin refiring during S, G2 and M. When, at the exit from M, nuclear membrane breaks down, the inhibitory factors are diluted out, leading to origin activation (Handeli and Weintraub, 1992; Mahbubani et al., 1997, McGarry et al., 1998).
Both these interpretations are correct. Nuclear membrane allows concentration of the replication factors that "licence" origins to fire, and, when the origins are fired, the accumulation of molecules (i.e. S phase cyclin-cdks) that activate origin firing but block the rereplication impeding the assembly of a functional initiation complex at the origins. In agreement with the hypothesis that the nuclear membrane acts concentrating the replicator proteins, it has been recently demonstrated that a single round of ORC dependent DNA replication can be achieved in *Xenopus* egg extracts in the absence of nuclear structure in a concentrated nuclear extract (Walter et al., 1998). This result suggests that the nuclear organization plays a role which is facilitative rather than obligatory in establishing replication forks, since once per cell cycle replication can be achieved in the absence of the nucleus.

Orc proteins in metazoans

ORC complexes have been purified from *Xenopus* and *Drosophila* (Gossen et al., 1995; Rowles et al., 1996). Although the properties of these complexes are not yet well defined, they appear to be essential for replication. Immunodepletion of Orc proteins blocks DNA replication in *Xenopus* egg extract (Carpenter et al., 1996; Rowles et al., 1996) and mutations in Drosophila Orc 2 gene causes a variety of DNA replication defects *in vivo*. In metazoans the proof that ORC interacts with origins of DNA replication is still lacking. The major difference between ORC complex in yeast and in

metazoans may rely in the dinamycs of its association with origins. In yeast, ORC complex is constitutively expressed and associated to the chromatin throghout the cell cycle. If also in metazoans the ORC complex marks replication origins costituting the landing pad over which the other components of the pre-RC are further assembled, it should not be bound to the origins throughout the cell cycle. Instead it should bind to the OBRs, defined as specific chromatin regions established after the Origin Decision Point, at the beginning of each G1. One of the strongest data in support of this hypothesis was obtained via *in vivo* footprinting at the lamin B2 origin locus, in which the observed cell cycle dependent footprint completely disappears in M phase (Abdurashidova et al., 1998). Whatever protein(s) interact(s) at this locus, it is not present in M phase and interacts again specifically during the following G1.

Not only the interaction between ORC complex and DNA is probably modulated during the cell-cycle, but also Orc genes expression seems to be subjected to cell-cycle mediated regulation. The expression of human Orc1 gene is proliferation dependent, negatively regulated in quiescent cells by the Rb-E2F complex and activated upon growth stimulation by G1 cyclin dependent kinase activity (Ohtani et al., 1996). E2F and pRB are major players in controlling G1 progression and S-phase entry. E2F binding sites are present at the promoters of many genes whose products are essential for commitment to S-phase (Johnson et al., 1993, Qin et al., 1994). At the beginning of G1 the transcription of these genes is repressed because E2F action is blocked by the interaction with pRB repressor (Weintraub et al., 1992). Late in G1, phosphorylation of pRB by cyclin dependent kinases frees E2F and triggers expression of E2F dependent genes that drive the cell in S-phase (Ewen et al., 1993; Kato et al., 1993; Ohtani et al., 1995, Botz et al., 1996, Geng et al., 1996). It is also becoming clear that ORC complex is subjected to developmental mediated regulation. It has been demonstrated in fact that the levels of Orc1 protein changes dramatically throughout *Drosophila* development. Orc1 is present exclusively in actively replicating cells and, at the stage in which chorion gene amplification occur in follicle cells of the ovary, its expression is restricted to foci of gene amplification (Asano and Wharton, 1999). The same relocalization at the site of gene amplification in response to developmental signals is shown by Orc2 protein, which expression is also under E2F control. Mutation in E2F or in its molecular partner dDP affects level of amplification and Orc2 protein localization in corresponding ways. These results indicate that E2F influences Orc2 protein localization either via E2F transcriptional targets or possibly by a more direct mechanism (Royzman et al., 1999).

Besides its function in replication, the ORC complex may also play different roles in regulating the state of the chromatin.
In yeast ORC complex is involved in the establishment of the silenced state of chromatin at the silent mating type loci, *HML* and *HMR* (Fox et al., 1995). In *Drosophila* the ORC complex interacts with HP1, a protein involved in the formation of heterochromatin, a metazoan nucleoprotein structure which shares many properties with the silenced mating type locus, allowing the proper chromosomal condensation and segregation (Pak et al., 1997).
In yeast, several temperature sensitive *orc5* mutants arrest the cells in early M phase, before the metaphase and anaphase transition and prior to the assembly of prereplication complex early in G1 (Dillin et al., 1998). The association of ORC complex with chromatin outside G1 and S phase may be important because information is contained in the pattern of ORC complex binding and/or ORC performs positive activities separate from its direct replication functions. From these data a role of the ORC complex is emerging in influencing the overall chromatin organization, and its involvement in initiation of replication and chromatin silencing might be viewed as an example of evolutionary variation of a general housekeeeping function.

Cdc6 protein in metazoans

The *Xenopus* counterpart of the yeast Cdc6 protein is essential for DNA replication, and interacts with ORC and MCM complexes (Coleman et al., 1996). It is possible to hypothesize that Cdc6, once localized at the origins by the interaction with the ORC complex, is acting also in metazoans as a clamp loader charging Mcm proteins onto the chromatin (Perkins and Diffley, 1998). Recently, the human homolog of the yeast Cdc6 protein was isolated (Williams et al., 1997) and the *in vivo* interaction between Orc1 protein and Cdc6 in human cells was demonstrated (Saha et al., 1998). Both in yeast and human (Elsasser et al., 1996; Saha et al., 1998) it was also shown that Cdc6 protein interacts with Cdk2, the kinase component that in association with the S phase promoting cyclins drives progression through S phase. The association of human Cdc6 with cyclin-Cdk2 is consistent with a role of the former as an adaptor protein and provides an attractive mechanism by which cyclins-Cdks can be recruited to the human prereplication complex in G1, providing a direct link between origin activation and cell cycle progression (Cardoso et al., 1993).
It is possible that cyclin-Cdk2 complexes phosphorylate Cdc6 protein and that this process triggers origin activation and inhibition of the Cdc6 association with chromatin, blocking the

origin refiring in S, G2 and M phases (Dutta and Bell, 1997). In late M and early G1 a mitotic inhibitor of cyclin-Cdck2 may associate with the complex, allowing reassociation of Cdc6 protein at the origins. Experiments in *S. cerevisiae* indicates that Sic1 protein might be the inhibitor of cyclin-cdk-Cdc6 interaction in yeast, whereas the identity of such inhibitor in metazoans is still unknown (Elsasser et al., 1996)
The levels of Cdc6 protein are differently regulated in yeast and human cells. In yeast the protein is degraded via ubiquitination as cells enter in S phase, and *de novo* synthesized as the cells enter the following G1 (Cocker et al., 1996; Piatti et al., 1995; Piatti et al. 1996; Santocanale et al., 1996). In human cells the time regulated transcription of CDC6 gene is dependent on E2F (Hateboer et al., 1998) and differently than in yeast, the Cdc6 protein is not degraded at the onset of S phase but it is translocated out of the nucleus (Saha et al., 1998). In this way human cells may regulate the functional pool of Cdc6 protein by selectively limiting its concentration in the nucleus when replication is initiated. The next round of replication then becomes dependent from the re-entry of the protein into the nucleus sometimes before the next S phase.

MCM complex in metazoans

Rao and Johnson demonstrated that, while G1 cells can replicate if fused with S phase cells, G2 cells cannot, because they lack a component that licence them for replication (Rao and Johnson, 1970). To account for this results Blow and Laskey proposed that chromosomes are licensed to undergo DNA replication once during each cell cycle by a regulator called Replication Licensing Factor (Blow and Laskey, 1988) which is inactivated by origin firing, ensuring that a single round of replication occurs. Following the dissolution of the nuclear envelope in mitosis, chromosomes are predicted to once again become licensed for replication via the binding of active RLF to chromatin.
Further studies demonstrate that the replication licensing depends from two biochemically distinct activities defined as RFL-M and RFL-B (Chong et al., 1995). Whereas the identity of RFL-B is not yet determined (Tada et al., 1999), analysis of the RFL-M demonstrated that it consists of complexes of all 6 *Xenopus* Mcm proteins which bind to chromatin in G1 and are removed as replication occurs (Kubota et al., 1995; Madine et al., 1995; Kubota et al., 1997, Thommes et al., 1997). RFL-M activity is depending from the presence at the chromatin of the *Xenopus* Origin Recognition Complex (Rowles et al., 1996) and *Xenopus* Cdc6 (Coleman et al., 1996). The dynamics of association of the initiator proteins at the

origins appears therefore to be conserved from yeast to Xenopus. The MCM complex may be loaded onto the chromatin by Cdc6 protein and act as a helicase promoting the unwinding of DNA at the origin and the formation of the replication fork (Ishimi et al., 1997). On the contrary, in human cells the picture seems to be more complex; in fact chromatin immunoprecipitation experiments suggested that MCM interaction keeps chromatin in an open conformation, structurally different from the bulk chromatin, but failed to detect the colocalization of Orc2 with the MCM complex (Ritzi et al., 1998). These data confirm the previous observation that, inside the MCM complex, Mcm3 associates preferentially with Mcm5, Mcm4 associates with Mcm6 and 7, while Mcm2 is loosely bound to the complex and might play a regulative role modulating the interaction of the complex with the chromatin (Burkhart et al., 1995; Ishimi et al., 1998). These results suggest that the interactions between chromatin bound Mcm proteins might be more complex and more dynamic than previously thought, and the composition and function the MCM complex may vary before and during S-phase (Coue et al., 1998).

Recent data point to the MCM complex as a target of the cell cycle regulation machinery that blocks reassociation of MCM complex in S, G2 and M phases, to avoid origin refiring. A yeast two hybrid screening revealed an interaction between Mcm7 protein and pRB, confirmed in *vitro* in the *Xenopus* system, in which the introduction of the protein blocks replication. A large body of evidence indicates that pRb indirectly inhibits the rate of cell proliferation via its interaction with E2F and several others transcription factors (Udvadia et al, 1995). It is proposed that pRb can directly downregulate DNA replication via interaction of its dephosphorylated form with Mcm7 protein in early G1. When pRb is phosphorylated by cyclin-Cdk kinases as cells approach the G1/S transition, it releases Mcm7 protein and enables it to interact with chromatin triggering replication initiation (Sterner et al., 1998).
Another regulative pathway controlling MCM complex activity involves geminin, a protein targeted to ubiquitin mediated proteolysis at the exit of mitosis by the Anaphase Promoting Factor (APC), the complex responsible for the degradation of the mitotic cyclins in human cells which drives them into a new G1 (Hochstrasser, 1996). Geminin is absent during G1 and starts to accumulate in S, G2 and M phases, interacting directly or indirectly with the Mcm complexes and blocking its association with chromatin and therefore origin refiring, up to the following G1 phase (McGarry et al., 1998). Geminin is therefore acting as a G2 inhibitor of replication initiation, in accordance with the ihibitory model, proposed in alternative to licensing factor hypothesis to

explain the failure of G2 nuclei in replicating DNA when fused with S cells. It is worth to note that in *Xenopus* extracts in which geminin has been immunodepleted, origin refiring does not take place. This is possibly due to the fact that cells may have evolved several mechanisms which operate to ensure that reinitiation does not occur (Handeli and Weintraub, 1992; Mahbubani et al., 1997; Donaldson and Blow, 1999).

The data reported above strongly indicate that regulation of origin activity is a crucial step and a convergent target in the control of cell cycle progression. A common frame is emerging in which cycle control and origin activation pathways are co-ordinated to ensure that the cell accomplishes correctly its major task: the faithful transmission of the genetic information to the progeny.

THE HUMAN LAMIN B2 ORIGIN LOCUS

The origin of DNA replication for which more details are currently available concerning the cis-acting sequences involved in origin activity is the one located immediately downstream of the lamin B2 gene. The genomic DNA segment encompassing the origin was originally identified in a library of nascent DNA sequences from early firing origins obtained from aphidicolin-synchronized HL-60 cells immediately after the release in S phase (Tribioli et al., 1987), and was localized by *in situ* hybridization in the short arm of chromosome 19, at the G-negative band of the sub-telomeric region p13.3 (Biamonti et al., 1992a). Later on, an origin activity was mapped by nascent strand abundance analysis via competitive PCR in an approximately 500 bp region, corresponding to the 3' end of the lamin B2 and the 5' portion of the intergenic spacer separating the lamin B2 gene from the downstream located ppv 1 gene (Giacca et al., 1994, see fig.1.11, panel A). Transcriptional studies of the region demonstrated that both lamin B2 and ppv1 transcripts are ubiquitously expressed and subjected to proliferation dependent regulation, with high steady state level of expression in proliferating cells and low levels in quiescent cells (Biamonti et al., 1992a, Biamonti et al., 1992b). The origin is firing in the first two minutes of S-phase in HL-60 cells, and it is active in all the cell types tested so far, including cells of myeloid, epithelial, neuronal and fibroblast origin and primary peripheral blood lymphocytes (Kumar et al., 1996). This behaviour is consistent with the fact that lamin B2 lies in a highly transcribed region of the genome showing an open chromatin structure (Dimitrova et al., 1996). Both ppv1 and lamin B2 genes are ubiquitously expressed

so that the chromatin region is probably derepressed in all the tissues, allowing efficient replication initiation from the ori. High resolution analysis of protein-DNA interactions in a 600-bp area encompassing the origin was carried out by the *in vivo* footprinting technique based on the ligation mediated polymerase chain reaction (LM-PCR). In growing HL-60 cells, sequences homologous to binding sites for known transcription factors (members of the helix-loop-helix family, NRF 1, Sp1 and UBF) were detected in the region corresponding to the promoter of the downstream gene. Upon conversion to a non proliferative state, a reduction in the intensity of these footprints was observed that paralleled the diminished transcriptional activity of the genomic area. In addition to these protections, in the region showing the peak of abundance in competitive PCR analysis of nascent DNA, a prominent footprint was detected. This footprint is asymmetric involving 70 nt on the lower strand and a smaller area on the upper strand, but, what is most intriguing, it is proliferation dependent, being absent from non-proliferating HL-60 cells (Dimitrova et al., 1996). This result indicates that the activity generating the proliferation dependent footprint might be involved in the process of origin activation.
The same approach was used to study dynamics of DNA-protein interaction at the origin area along the cell cycle. In G0 cells no protection is present; as the cells progress in G1 an extended footprint appears, covering more than 100 bp. When the cells enters in S-phase the protection shrinks to 70 bp, reproducing the pattern previously obtained for asynchronous growing cells (fig.1.11, panel B). In mitosis the protection totally disappears, and reappears in its extended form as the cell move in the next G1 (Abdurashidova et al., 1998). This modulation of the protein DNA interaction pattern at the region strictly resembles the assembly of the yeast pre-replicative complex in G1, which is converted in the post-replicative one upon origin activation (Diffley et al., 1994). Nonetheless a striking difference is the complete absence of DNA-protein interactions in M phase, whereas in yeast the ORC mediated pre-replicative complex marks the origins throughout the cell cycle. These data indicate that the protein complex bound at lamin B2 origin is reset in M phase and needs to be reestablished at any new G1.
We developed an assay that allows us to detect initiation events in genomes of any degree of complexity and we apply this approach on the study of the lamin B2 origin region. The aim of this work is to localize the start sites of bidirectional DNA replication at the lamin B2 origin at nucleotide resolution, to obtain data comparable

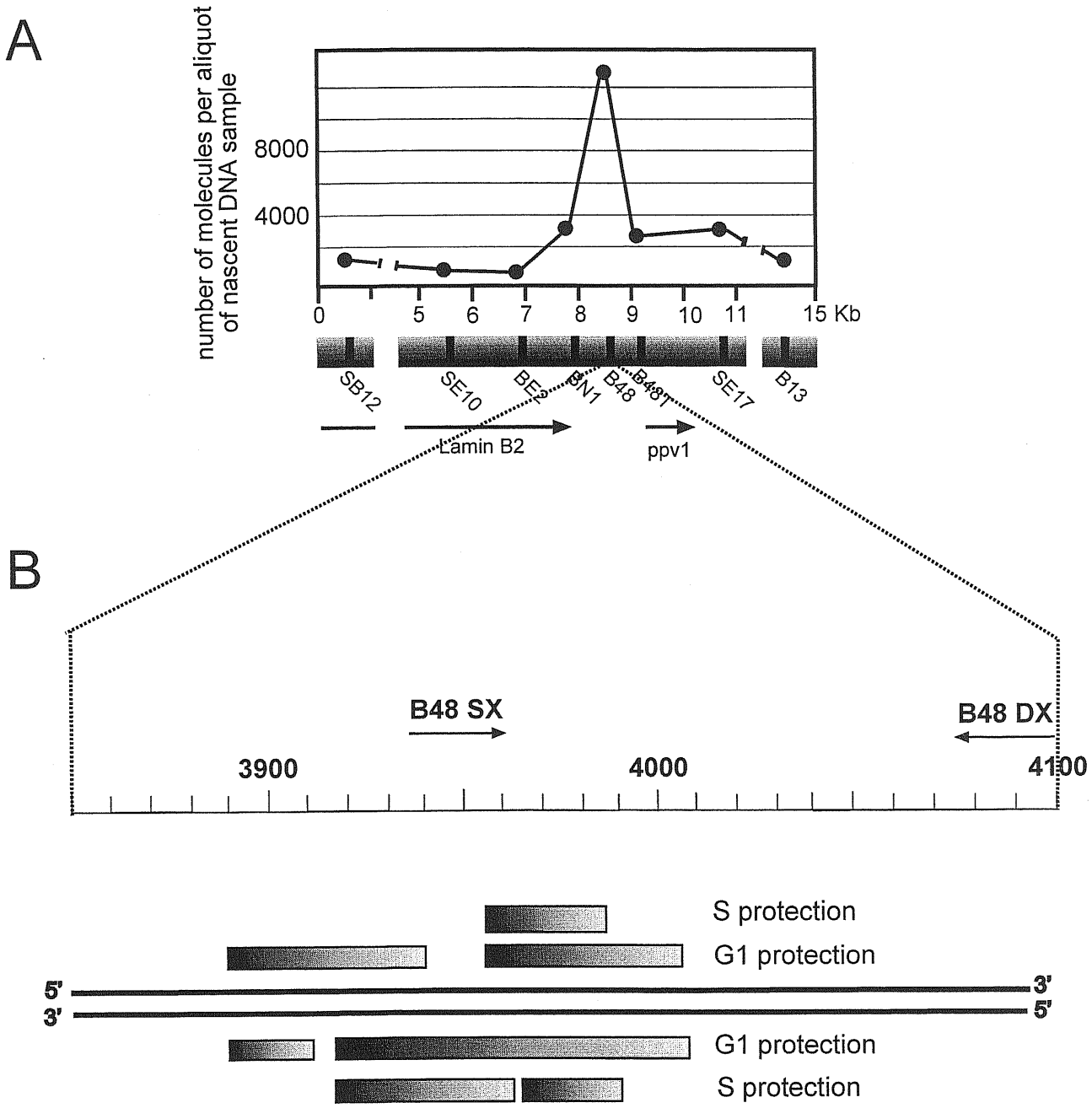

Figure 1.11: the lamin B2 origin region.

(A): Mapping of the lamin B2 origin by competitive PCR.
The origin is localized in a region of 500nt centered around the B48 marker that shows the higher abundance in nascent DNA samples in comparison with distal markers, as indicated in the graph. The positions and the relative distances of the PCR fragments analyzed in the mapping of the origin are indicated as black boxes on the grey bar below the graph.
The locations of the lamin B2 and ppv1 transcripts are also shown.

(B): The origin area is magnified and the cell cycle dynamics of DNA-protein interactions taking place at the lamin B2 ori are schematically represented.

to the nucleotide resolution mapping performed for eukaryotic virus origins and *S. cerevisiae* ARS1 (Hay and DePamphilis, 1982, Hendrikson et al; 1986; Bielinsky and Gerbi, 1998, 1999). In this way it may be possible to understand if replication origins analogous to those present in simple organism exist in metazoans, in spite of the differences that distinguish metazoan origins from origins at simple genomes. Metazoan origins in fact do not show a modular anatomy nor consensus sequences, are used at low frequency, since just a small population of cells is starting replication from a defined origin at any S-phase, and reestablish the binding with specifically interacting proteins at any new G1.
As a consequence, it may be possible to understand if the replicon model can be extended to higher eukaryotes. In this case the replicator sequence may be a structure defined by multiple parameters in which key roles are played by chromatin organization and nuclear architecture, and the initiator elements may be a multiprotein complex dynamically interacting with it in a cell cycle regulated fashion. If the complex responsible for the G1 extended protection in human cells is the metazoan counterpart of the pre-RC (i.e. the human initiator element), the start site of bidirectional replication might be located nearby or inside the protected area, in analogy to what was observed for ARS1 in yeast and for oris at chromosomes of eukaryotic DNA viruses.

MATERIALS AND METHODS

CELL CULTURE AND SYNCHRONIZATION

HeLa cells were cultured in Dulbecco modified Eagle's medium supplemented with 10% fetal calf serum (FCS), 2mM L-glutamine and 50 µg/ml gentamicin; cultures were kept in 5% CO_2 at 37°C.
For synchronization experiments HeLa cells at 50-60% confluence were piled up in M phase by incubation for 10 hours with 50 ng/ml nocodazole (Sigma) in complete medium. The cells were then detached by mechanical shake-off, washed three times with complete medium and replated. Three hours later the monolayer was washed to remove unattached cells, aphidicolin was added to the medium at the concentration of 5 µg/ml and the culture was incubated for 20 hours to accumulate cells at the G1/S border. The cells were washed twice with complete medium to remove aphidicolin and released into S phase in complete medium for 30 min.
When inhibition of Okazaki fragment synthesis was necessary, asynchronous cultures were incubated in presence of 2µM emetine (Sigma) for 1 hour (Burhans et al., 1991).
IMR90 cells were grown in Dulbecco modified Eagle's medium up to 60% confluence, and piled up in G0 by incubation in serum free medium for 3-4 days. Before re-entry into the cell cycle, G0 cultures were washed and then incubated in complete medium with 10% FCS for 5 hours to collect G1 cells, or in complete medium supplemented with 5 µg/ml aphidicolin for 24 hours to accumulate them at the G1/S border. Early S phase samples were obtained by washing the G1/S arrested cells and releasing them in complete medium for 30 min. Synchronization was followed by flow-cytometric analysis.

DNA EXTRACTION FROM HUMAN CELLS

Cells were detached by treatment with trypsin, washed at first in complete medium to inactivate the enzyme and then in cold PBS. Nuclei were isolated by resuspending the cells in RSB buffer (10mM Tris/HCl pH8, 10mM NaCl, 3mM $MgCl_2$) at a concentration of $2,5x10^7$ cells/ml and incubating the sample on ice for 5 min. An equal volume of 0.2% NP-40 in RSB buffer was added followed by an incubation on ice for additional 10 min. The nuclei so obtained were pelleted by centrifugation and treated with lysis buffer (200mM NaCl, 10mM Tris/HCl pH8, 25mM EDTA, 1% SDS and 600µg/ml proteinase K) overnight at 56°C.

The sample was extracted twice with phenol/chloroform/isoamyl alcohol 25/24/1 and once with chloroform/isoamyl alcohol. At the end an equal volume of isopropanol was added to the aqueous phase and the precipitate was collected by centrifugation for 30 min at 14000 rpm in Eppendorf centrifuge at 4 °C. The DNA pellet was rinsed twice with 70% ethanol and resuspended in TE buffer (10mM Tris-HCl pH8, 1mM EDTA pH8) in presence of 1unit/µl RNAse inhibitors (Boheringer Mannheim).
The genomic DNA extracted from synchronized cells was further digested with BamHI and EcoRI to obtain fragments of an average length of 10 kb.

YEAST CULTURE AND DNA EXTRACTION FROM YEAST CELLS

Yeast genomic DNA was isolated from exponentially growing *Saccharomyces cerevisiae* SP1 strain (kindly provided by S. Gerbi) at the cell density of $2x10^8$ cell/ml. Cells were harvested by centrifugation, washed in sorbitol buffer (1M sorbitol, 100 mM EDTA) and resuspended in spheroplast buffer (1M sorbitol, 100mM EDTA, 100mM β-mercaptoethanol, 3 unit/µl zymolase). Spheroplasts were then resuspended in lysis buffer (50mM Tris-HCl pH8, 50 mM EDTA, 1% SDS, 1mg/ml of proteinase K) and incubated at 30 °C. After lysis and debris precipitation by 3M potassium acetate, nucleic acids were extracted twice with phenol:chloroform:isoamyl alcohol (25:24:1), precipitated with isopropanol and resuspended in TE buffer.

GENOMIC DNA FRACTIONATION BY SUCROSE GRADIENT CENTRIFUGATION

150 µg of human DNA and 12 µg of yeast DNA were denatured 5 min at 95 °C and fractionated on a 5 to 30% (w/v) linear sucrose gradient (35 ml) by centrifugation for 17 hours at 20 °C, in a Beckmann SW28 rotor at 26000 rpm. In a parallel gradient, double stranded DNA size markers were fractionated; 34 one ml fractions were collected for both gradients and aliquots of the marker fractions were analyzed on agarose gel. Using the separation pattern of the marker DNA as sedimentation velocity reference, fractions containing ssDNA ranging in size between 800 and 1200 bp were pooled and the DNA was recovered by ethanol precipitation and resuspended in 20 µl of TE buffer.

Supplement to page 56

Yeast and human DNA is dissolved and denatured in TE buffer (10mM Tris-HCl pH 8, 1mM EDTA) before being loaded on sucrose gradient.
The 5% and 30% (w/v) sucrose solutions utilized in the preparation of the gradient are obtained dissolving sucrose in TNE buffer (10mM Tris-HCl pH 8, 300mM NaCl, 1mM EDTA).

PURIFICATION OF RNA-PRIMED DNA BY TREATMENT WITH λ-EXONUCLEASE

Size-fractionated DNA samples containing between 0.2 and 2 μg, and restricted denatured genomic DNA samples from synchronized cells (20 μg) were treated with polynucleotide kinase (New England Biolabs), to phosphorylate 5' hydroxyl ends exposed by aspecific DNA shearing and make all the parental contaminating molecules available for the λ-exonuclease digestion. The reaction was carried out in a total volume of 50 μl with 2.5 μl of enzyme (10 units/μl) in presence of 1mM ATP, in the condition described by Gerbi and collaborators (Gerbi and Bielinsky, 1997). The phosphorylated DNA was resuspended in water and subjected to λ-exonuclease treatment with 2 μl of enzyme (Gibco BRL) in 67 mM glycine-KOH pH 8.8, 2.5 mM $MgCl_2$, 50 μg/ml BSA, in a final volume of 20 μl at 37 °C overnight. The enzyme was heat inactivated 10 min at 75 °C and the samples were extracted once with phenol/chloroform/isoamyl alcohol and once with chloroform/isoamyl alcohol. Ethanol precipitation was carried out in dry ice for 20 min, in presence of 0.3M sodium acetate pH 7.2 and 2 μl of glycogen (20 μg/ml) as carrier and the DNA pellet was resuspended in TE. The RNA was degraded by incubation with RNAse A (Boheringer Mannheim, final concentration: 100 μg/ml), RNAse T1 (Boheringer Mannheim, final concentration: 1 units/μl), and RNAse T2 (Gibco BRL, final concentration: 0.6 units/μl) overnight at 37 °C, followed by protein digestion in the same buffer supplemented with 0.1% SDS and 600 μg/ml proteinase K. The DNA was then purified by phenol/chloroform extraction and ethanol precipitation as described above. In alternative, RNA was degraded by heating the samples 5 min at 95 °C in 0.1M NaOH. The solution was neutralized with HCl and the DNA was recovered by ethanol precipitation. The samples were resuspended in water and subjected to a second kinase reaction, to phosphorylate the 5' hydroxyl DNA ends exposed on nascent DNA after RNA primer digestion. The reaction was carried out in the conditions defined above. The DNA, purified and recovered by precipitation as previously described, was finally resuspended in 20 μl of TE buffer. Aliquots of the samples before and after the λ-exonuclease treatment were analyzed by competitive PCR and LM-PCR.

COMPETITIVE PCR ANALYSIS OF NASCENT DNA

Primer choice and competitor construction for human cells

The quantitation of the abundance of two different sequences in the lamin B2 gene region was performed by competitive PCR as

detailed by Diviacco et. al., 1992. The sequences chosen for amplification are defined as B48 and B13 (see results, figure 3.1); the former corresponds to the lamin B2 ori (within 474 nt.), whereas B13 is displaced by 5 kb. A unique competitor molecule was used for both markers. The core of the molecule is a 110 bp region derived from the β-globin gene harbouring a 20 bp insertion. The core competitor was built directly from the amplification products obtained by the overlap extension method (Higuchi *et al.* 1989; Ho *et al.* 1989). A set of four primers were synthesized for a region of the β-globin globin gene (fig. 2.1 a). Two external primers (PCO3 and PCO4) were synthesized together with two internal primers (PCO/+1 and PCO/+2) consisting of a common 5' tail of 20 nt linked to the specific sequences complementary to genomic targets on the 3'-end. Pairs of internal and external primers were used in two separate PCR reactions for the construction of two intermediate products (fig. 2.1 b and b'). These intermediate products were eluted from acrylamide gels, mixed, denatured and annealed (fig. 2.1 c). Subsequently, after one round of extension, the hybrid product was amplified using the external primers to obtain the core competitor molecule (fig. 2.1 d), which has the same sequence as the β-globin genomic target, except for the addition of 20 nt. in the middle. The forward and reverse primers spanning the lamin B2 origin area (SB12, SE10, B13, B48 see Giacca et al., 1994) arranged in a head to tail fashion were joined to the core molecule by PCR amplifications, using chimerical primers (fig. 2.1 e; Giacca et al., 1997). This approach allowed us to use a single competitor for the quantification of the relative abundance of different PCR markers. The competitor was quantified in competitive PCR experiments against a known amount of plasmid molecules harbouring the lamin B2 origin area. The PCR cycle profile was as follows: denaturation at 94° C, annealing at 56° C and extension at 72° C; time for each step was 30 sec; 35 cycles were performed with 1 unit of Taq polymerase (Boehringer Mannheim) in the conditions recommended by the manufacturer. Competitive PCR experiments were carried out challenging the same amount of size-selected DNA from asynchronous cells (before and after the λ-exonuclease treatment) and total DNA from synchronized cells (after the λ-exonuclease treatment) with 10-fold serial dilutions of the competitor.

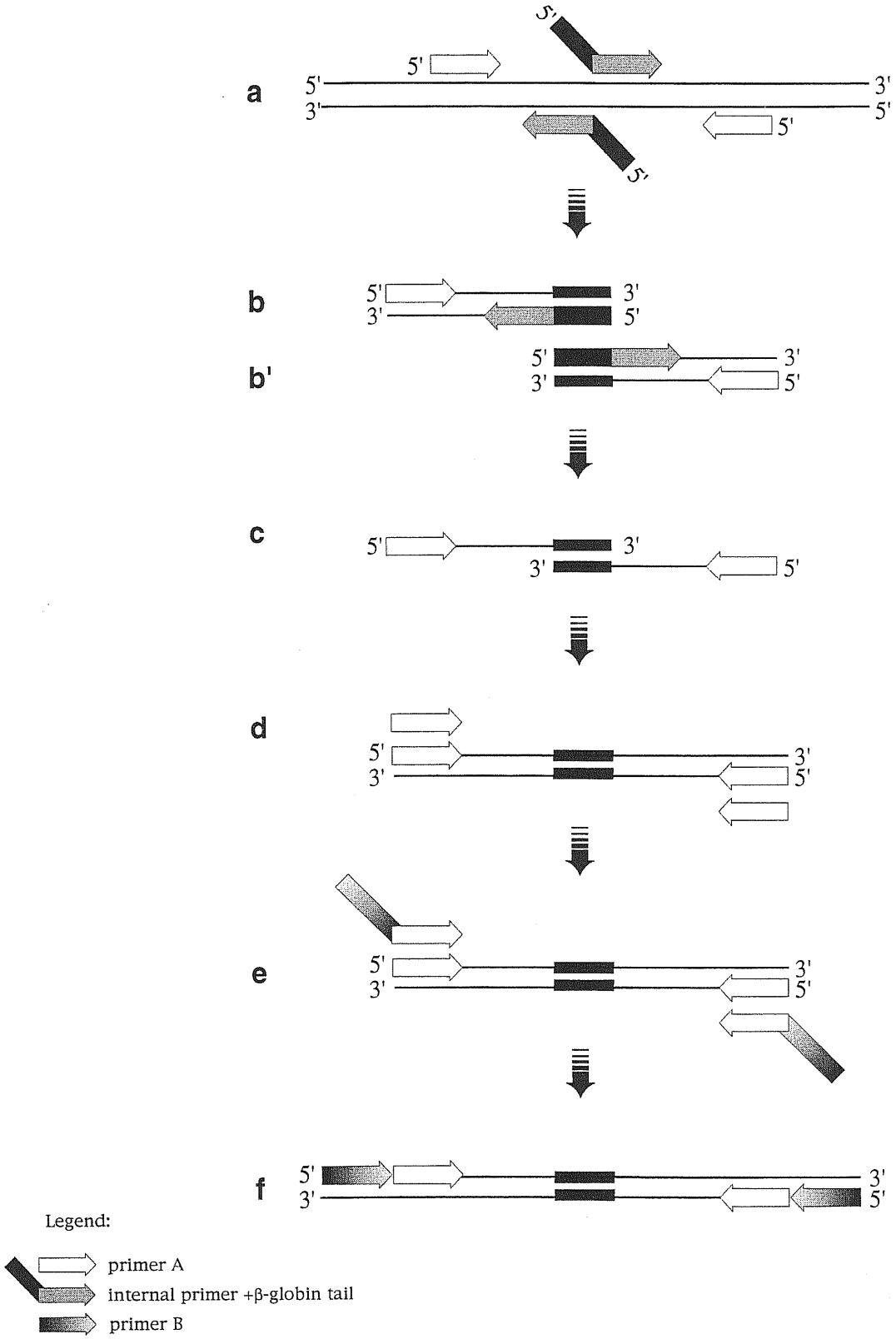

Figure 2.1: construction of a competitor molecule harbouring recognition sites for primer pairs amplifying two different genomic regions (A and B); see text for description.

Primer choice and competitor construction for yeast

Two target genomic DNA sequences were chosen on chromosome IV. The first in close proximity of ARS1 origin (Pink-region) and the second one (Blue-region) approximately 4 kb upstream of ARS1 origin. For these genomic DNA targets a unique competitor DNA molecule was constructed. The core competitor was built as described in figure 2.1 from a to d, and it is constituted by the Blue region carrying in the middle a 20 bp insertion from the human β-globin gene. The recognition sites for the pink primers are added to the competitor molecule by PCR with chimeric primers as detailed in figure 2.1 e and f.
The PCR cycle profiles were as follows: denaturation at 95 °C, 30 seconds, annealing at 63°C (for the blue region), or 58 °C (for the pink region), 20 seconds, and extension at 72 °C, 40 seconds; 35 cycles were performed. For each primer set a fixed amount of nascent DNA was coamplified with four-fold dilutions of the competitor.
The sequence of the primers used in competitive PCR analysis of human and yeast DNA, together with the internal primers carrying the 20 bp sequence of the human β-globin gene at the 5' end used in the construction of the competitor molecules are shown in table 1.

Table 1

Region	Template size, bp	primer	Sequence	Genomic Position and Source
human				
B48	164	B48dx	5'-GACTGGAAACTTTTTTCTAC-3'	3937-3957
		B48sx	5'-TAGCTACACTAGCCAGTGACCTTTTTCC-3'	4077-4104
				humlambbb
B13	168	B13dx	5'-GCCAGCTGGGTGGTGATAGA-3'	198-217
		B13sx	5'-CCTCAGAACCCAGCTGTGGA-3'	54-73
				unpublished
β-globin	110	PCO3	5'-CAACTTCATCCACGTTCACC-3'	
		PCO4	5'-ACACAACTGTGTTCACTAGC-3'	
		PCO/+1	5'-tail1-GGAGAAGTCTGCCGTTACTG-3'	
		PCO/+2	5'-tail2-TCAGGAGTCAGGTGCACCAT-3'	
				humhbb
yeast				
Pink	226	Pink F	5'-TTTGGATTGGAACTCGATTTCTG-3'	462300-323
		Pink R	5'-ACTCAGTAATAACCTATTTCTTAGC-3'	462501-526
Blue	299	Blue F	5'-CCCTGATTGACAAGCAACTC-3'	458356-376
		Blue R	5'-GGAGGAATTCGAGAATGTTTATCG-3'	468631-655
Blue/tail		Blue/tail1 F	5'-tail1-CACTTCTTAAATGCAGTCAG-3'	458446-466
		Blue/tail2 R	5'-tail2-CCGATAATAGTTATATTCC-3'	458529-548
				yeast genome sequence

Note: tail 1: 5'-ACCTGCAGGGATCCGTCGAC-3'; tail 2: 5'-GTCGACGGATCCCTGCAGGT-3'.

LM-PCR ANALYSIS OF THE NEWLY SYNTHESISED DNA

≈5 x 10^5 molecules of nascent DNA were subjected to LM-PCR according to the protocol of Quivy and Backer (1993). 1 µg of DMS modified genomic DNA was used in parallel as positive control for the LM-PCR reaction and to obtain a sequencing ladder of the lamin B2 region, allowing us to localize the position of the nascent DNA specific bands on the sequence. The primer sets used in this study are shown in table 2. The DNA samples are denatured 5 min. at 95 °C and annealed with 0.1 pmol of the primer # 1 of each set for 30 min. at 60 °C in 10 µl of 1X Vent polymerase buffer (New England Biolabs). When the annealing is completed, the primer # 1 is allowed to extend along its template by the addiction of dNTPs (fin. con. 0.5 mM) and Vent exo- (1 unit) in a final volume of 20 µl. The extension of the first primers is performed at the temperature indicated in table 2 with the following scheme: 5 min. at 60 °C, 10 sec. at 65 °C, 10 sec. at 70 °C and 10 min. at 76 °C.
The product of the elongation is ligated to the double stranded asymmetric linker prepared as detailed in Ausubel et al., 1994. 20 µl of extension reaction are mixed with 5 µl of ligase buffer (New England Biolabs), 5 µl of annealed asymmetric linker (200pmol/µl), and 19 µl of 40% PEG 8000, to increase the viscosity of the reaction medium in order to maintain the molecules in an extended position to facilitate ligation. 1 µl of DNA ligase (400 units, NEB) is added to mixture and the ligation is performed overnight at 16 °C. 150 µl of TE are added to the ligated product which is purified by phenol chloroform extraction and precipitated with ethanol in presence of 3M ammonium acetate. The DNA pellet is washed in 70% ethanol, dried in speedvac and resuspended in 20 µl of water. The ligated product is subjected to PCR amplification in 1X Vent buffer, 4 mM $MgSO_4$, 0.2 mM dNTPs, 10 pmol of primer # 2, 10 pmol of long linker primer for LM-PCR, 1 unit of Vent in a total volume of 50 µl. The PCR conditions are as follows: pre-PCR denaturation: 3 min. 95 °C, denaturation: 1 min. at 95 C°, annealing: 2 min. at the temperature indicated for each primer # 2 in table 2, extension: 3 min. at 76 °C plus 3 seconds added at any further cycle. 18 cycles of amplification are performed. 20 µl of the amplification reaction are subjected to radioactive extension in presence of 0.8 pmol of $\gamma^{33}P$ labelled primer # 3 in 1X Vent buffer, 2.6 mM $MgSO_4$, 0.2 mM dNTPs, 0.5 unit of Vent, in a final volume of 30 µl. The condition of the radioactive extension are as follows: pre-PCR denaturation: 3 min. at 95 °C; denaturation: 1 min. at 95 °C; annealing: 2 min at the temperature indicated in table 2 for each primer # 3; extension: 7

min. at 95 °C. 5 cycles of extension are performed. The extension reaction is blocked adding EDTA (fin. conc. 40 mM). The DNA is purified by phenol-chloroform extraction and precipitated with ethanol in presence of 0.3 M sodium acetate, washed with 70% ethanol, dried briefly and resuspended in loading buffer for sequencing gels (95% formamide, 100 mM EDTA, 0.5% bromophenol blue, 2.5% xylene cyanol). The samples are denatured 3 min. at 95 °C and loaded on a 8% acrylamide sequencing gel.

IN VITRO MODIFICATION OF GENOMIC DNA

The control sequencing ladder was obtained subjecting to LM-PCR reaction genomic DNA treated *in vitro* with Dimethyl Sulfate (DMS) which methylates guanosines. In basic condition at high temperature methylated guanosine are subjected to β-elimination which leads to DNA breaks. At limiting DMS concentration modification occurs once per DNA molecule, giving a population of fragments each generated by the breakage on a different guanosine (Saluz and Jost, 1987).
The DMS treatment is performed as follows: 175 µg of total genomic DNA are treated with DMS (fin. con. 0.25%) for 2 min. at room temperature. The reaction is blocked by the addition of β-mercaptoethanol (fin. conc. 200 mM) and DNA is immediately precipitated with ethanol in presence of 0.3 M sodium acetate. DNA is recovered by centrifugation, resuspended in 200 µl of 1M piperidine solution in water, heated for 30 min. at 95 °C, frozen in dry ice and dry in speedvac overnight to eliminate any trace of the organic solution. The DNA pellet is resuspended in 360 µl of TE buffer and precipitated twice, the first time with 2.5 volumes of ethanol and 0.3 M sodium acetate, the second time with the same volume of isopropanol and 3 M ammonium acetate. The DNA pellet is resuspended in 50 µl of water, frozen in dry ice and dried in speedvac from three hours to overnight. The recovered DNA is finally resuspended in 50 µl of TE buffer

table 2

Set	Primer	Length, bp	Sequence	Genomic Position	Annealing Temperature, °C
	human				
A	1	25	5'-TTACCTACACGAGCTACCCGTGGTT-3'	4385-4409	60
	2	22	5'-TGGTTGCGACTCCGCGGGAAGA-3'	4368-4389	68
	3	27	5'-TCCGCGGGAAGAGGGAGGCCCTGAGTT-3'	4353-4379	70
B	1	28	5'-GGCTAGTGTAGCTAGTGTAAACAGGACC-3'	4091-4118	60
	2	25	5'-GTAAACAGGACCCAGGCGATGCATG-3'	4107-4131	67
	3	27	5'-CAGGACGGAGGCGATGCATGGGACCCT-3'	4112-4138	70
C	1	25	5'-TCGCATCACGTGACGAAGAGTCAGC-3'	4179-4203	60
	2	25	5'-GAGTCAGCTTGTGCAACAGCGTCGG-3'	4162-4189	66
	3	27	5'-GCTTGTGCAACAGCGTCGGAGGCTCAC-3'	4154-4180	68
D	1	25	5'-GTCACAGCACAACCTGCAAAAACGG-3'	3795-3821	60
	2	25	5'-CAAAAACGGAGCTGGGCTGCAGCTG-3'	3813-3837	70
	3	26	5'-GGGCTGCAGCTGGGGCTGGCATGGAC-3'	3826-3851	72
E	1	25	5'-GGGGTGGAGGGATCTTTCTTAGACA-3'	4049-4074	60
	2	27	5'-GACATCCGCTTCATTAGGGCAGAGGCC-3'	4026-4052	68
	3	27	5'-TCATTAGGGCAGAGGCCCGGCTCGAGC-3'	4016-4042	70
F	1	30	5'-GAGTCCCTCAGATCTTTAACAAAGAACTGC-3'	3712-3742	60
	2	25	5'-AACTGCCGCGTGCAGGCTTCAGACC-3'	3693-3718	68
	3	30	5'-CGCGTGCAGGCTTCAGACCAACCCCAGCCA-3'	3682-3712	72
G	1	21	5'-TGCACAGCGCCAGGTTAACGC-3'	3573-3594	60
	2	25	5'-CAGGTTAACGCTGAACGGTGCCCCG-3'	3584-3609	68
	3	26	5'-CTGAAGCCTGCCCCGGTGAGCCCAAG-3'	3593-3619	70
H	1	21	5'-GGGCTGGCATGGACTTTCATT-3'	3837-3858	60
	2	29	5'-GGCATGGACTTTCATTTCAGAGATTCGGT-3'	3842-3871	68
	3	33	5'-CGGTTTTTAAGAAGATGCATGCCTAGCGTGTTC-3'	3867-3900	70
	yeast				
Fy	1	25	5'-AGGTGAACTTTTGGATTGGAACTCG-3'	442291-315	60
	2	27	5'-TCTGACTGGGTTGGAAGGCAAGAGAGC-3'	462319-345	68
	3	30	5'-CTGACTGGGTTGGAAGGCAAGAGAGCCCCG-3'	462320-349	72
Ry	1	25	5'-GTATTTATATACTAAGCTCGGGGCG-3'	462820-826	60
	2	27	5'-TAAGCTGCCGGCGGTTGTTTGCAAGAC-3'	462788-814	68
	3	30	5'-AAGCTGCCGGCGGTTGTTTGCAAGACCGAG-3'	462784-813	72
SL*		11	5'-GAATTCAGATC-3'		
LL*		25	5'-GCGGTGACCCGGGAGATCTGAATTC-3'		68

The numbering of the primers for the human Lamin B2 region refers to the file humlambbb of GeneBank (accession number M94363).

The numbering of the primers for the ARS1 region refers to the sequence of chromosome IV of *S. cerevisiae* genome.

*SL and LL are the short and long linker primer, respectively

RESULTS

ISOLATION OF RNA-PRIMED NASCENT DNA

The portion of the lamin B2 gene area that we are exploring is shown in Figure 3.1, panel A. Our previous localization of the lamin B2 ori was performed using as starting material the nascent single strands of a size close to 1 kb isolated from asynchronously growing cells, after denaturation and sucrose gradient fractionation; we have demonstrated previously that no significantly greater enrichment for the recently activated ori sequences is obtained if other purification steps are added, such as pulse-labelling with BrdUrd and affinity column fractionation (Kumar et al., 1996). In such samples of nascent DNA from proliferating cells we have observed by competitive PCR a tenfold enrichment of the ori containing sequence (corresponding to the fragment labelled B48) with respect to control sequences (typically, fragment B13, which is 5 kb removed to the right of B48: see fig. 3.1, panels A and B). We explored the possibility that a further enrichment could be obtained by treating the so isolated nascent DNA with the λ-exonuclease (that is unable to degrade RNA-primed sequences) thus reducing the background caused by the contamination with randomly sheared DNA, and increasing the fraction of molecules bearing a 5' RNA primer. This is indeed the case: panel B shows the relative abundance measured by competitive PCR of the B48 and B13 markers in a sample of nascent DNA isolated from exponentially growing HeLa cell cultures, yielding a tenfold enrichment, and the results of the measurement performed on the same sample following hydrolysis by λ-exonuclease, yielding instead a greater than 50-fold enrichment. The same experiment performed in human quiescent normal lymphocytes does not show any enrichment of B48 sequences over B13 ones before λ-exonuclease treatment, nor the presence of any residual DNA after such treatment.

The effectiveness of the treatment with λ-exonuclease was further strengthened by the experiment shown in fig. 3.1, panel C: here we collected HeLa cells synchronized in G1, early S or late S phase and we extracted their DNA, denatured it and exposed it to λ-exonuclease treatment without any prior size fractionation. If the hydrolytic step were really effective and selective, it should remove all sheared DNA and leave only RNA-primed nascent one. In fact, from the G1 sample we could not recover any detectable amount of DNA, as determined by PCR amplification (data not shown); a measurable amount of DNA was obtained from the cells in early S; the lamin B2 ori is known to fire within the first minute

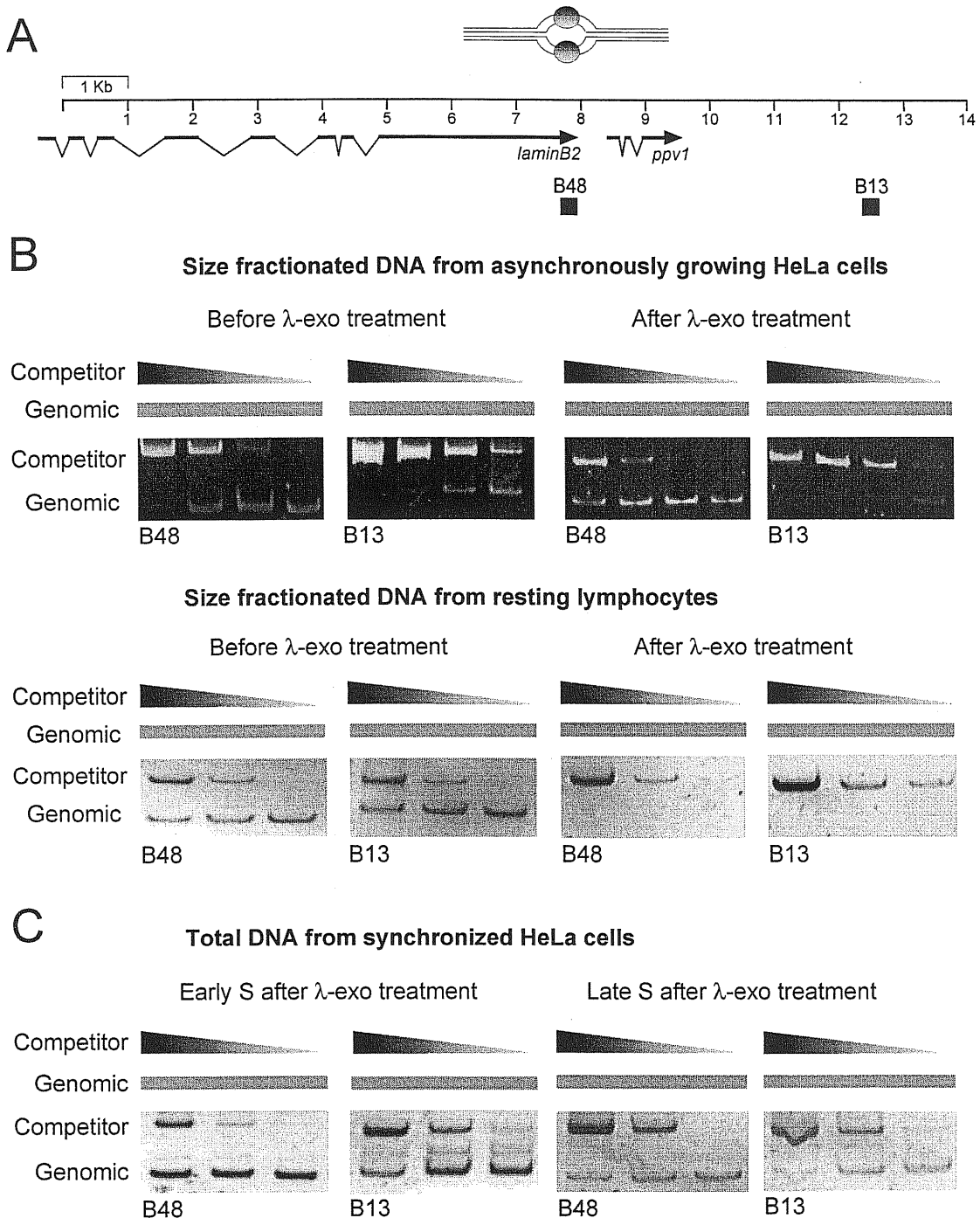

Figure 3.1: Isolation and analysis of the newly replicated DNA.

(A) Schematic representation of the human lamin B2 genomic region containing the origin of DNA replication analyzed in this study. The bubble structure marks the position of the ori. The location of lamin B2 and ppv1 transcripts are shown. The positions of the regions amplified by PCR are indicated by black boxes.

(B) Competitive PCR analysis of size-fractionated nascent DNA isolated from asynchronously growing HeLa cells or from human normal G0 lymphocytes. DNA samples were analyzed for their abundance in origin (B48) or control (B13) sequences before or after λ-exonuclease treatment. For each region a fixed amount of DNA was co-amplified with amounts of competitor molecules decreasing by tenfold dilutions, as detailed in Materials and Methods. The number of competitor molecules inserted in the PCR reactions ranges from 10^4 to 10 molecules for the first set of reactions in panel B, and from 10^3 to 10 molecules for all the other reaction sets. The positions of the competitor and genomic DNA amplification products are shown.

(C) Competitive PCR analysis of total DNA isolated from HeLa cells synchronized at the beginning and at the end of S phase and treated with λ-exonuclease.

of S (Tribioli et al., 1987), and, accordingly, we observe in the DNA sample an over tenfold enrichment in the B48 sequence over the B13 one (see above and panel A); instead in late S, when other oris fire and the lamin B2 replicon is presumably entirely duplicated, no enrichment of B48 over B13 is observed.
We can thus conclude that, in the preparations of size-fractionated and lambda exonuclease digested DNA, we are dealing with a population of molecules offering a good representation of the newly replicated sequences at the oris firing shortly before the arrest of the culture. The nascent DNA so prepared can thus be utilized in experiments aimed at mapping the RNA-DNA junctions resulting from the DNA replication process. To this aim, after removing the RNA primer and performing second strand synthesis with appropriate primers, it is possible to measure the size of these strands, as was performed for the similar experiment in yeast (Bielinsky et al., 1999). On the other hand, in view of the size and complexity of the human genome, in order to obtain a detectable signal after second strand synthesis one has to amplify the duplex fragments so obtained by ligation-mediated PCR (LM-PCR) (Quivy and Becker 1993). We designed therefore a novel approach in which the actively replicating origin DNA is selectively isolated by size fractionation or by cell synchronization and is further purified by the treatment with λ-exonuclease (fig 3.2a). This material is analyzed by LM-PCR as described in fig. 3.2b and in Material and Methods. The amplification step of LM-PCR increases the sensitivity of the method, making it a valuable tool to attempt the mapping of RNA-DNA junctions in complex genomes.

VALIDATION OF THE NOVEL APPROACH TO MAP 5'ENDS OF NASCENT DNA ON YEAST ARS 1 ORI

In order to validate our system of nascent DNA extraction and analysis, we applied the approach we intended to use in human cells to map the 5' ends of nascent DNA at ARS1 ori in yeast, as already performed by Bielinsky and Gerbi with a different technique (see introduction). Accordingly, we isolated from asynchronously growing yeast cells nascent DNA of approximately 1 kb in size by sucrose gradient fractionation of the total denatured DNA. At this point we would expect that the isolated nascent DNA should be enriched in ori sequences; the results of the quantification of the abundance of ori vs. control sequences

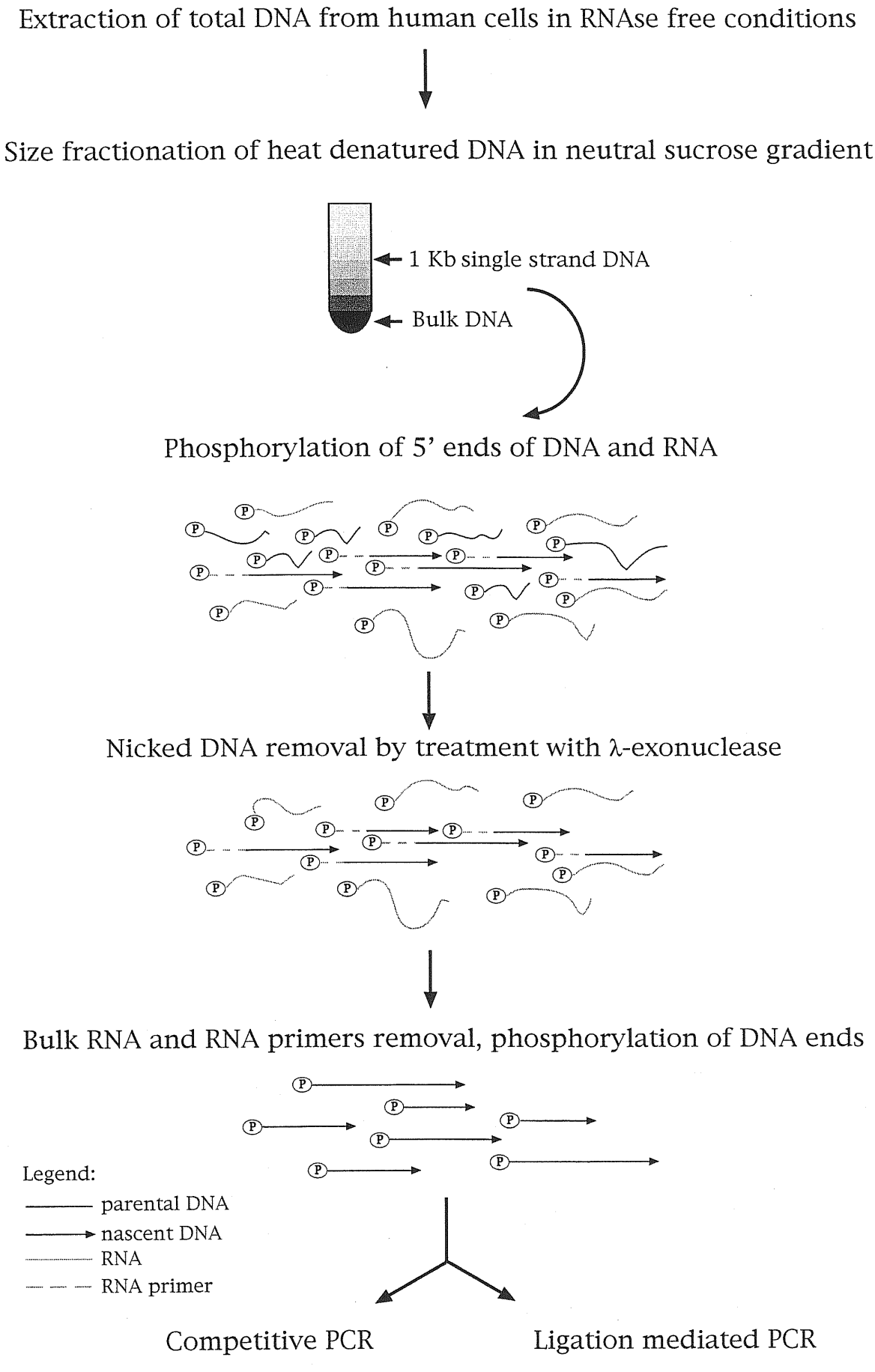

Figure 3.2a: nascent DNA isolation from human cells.

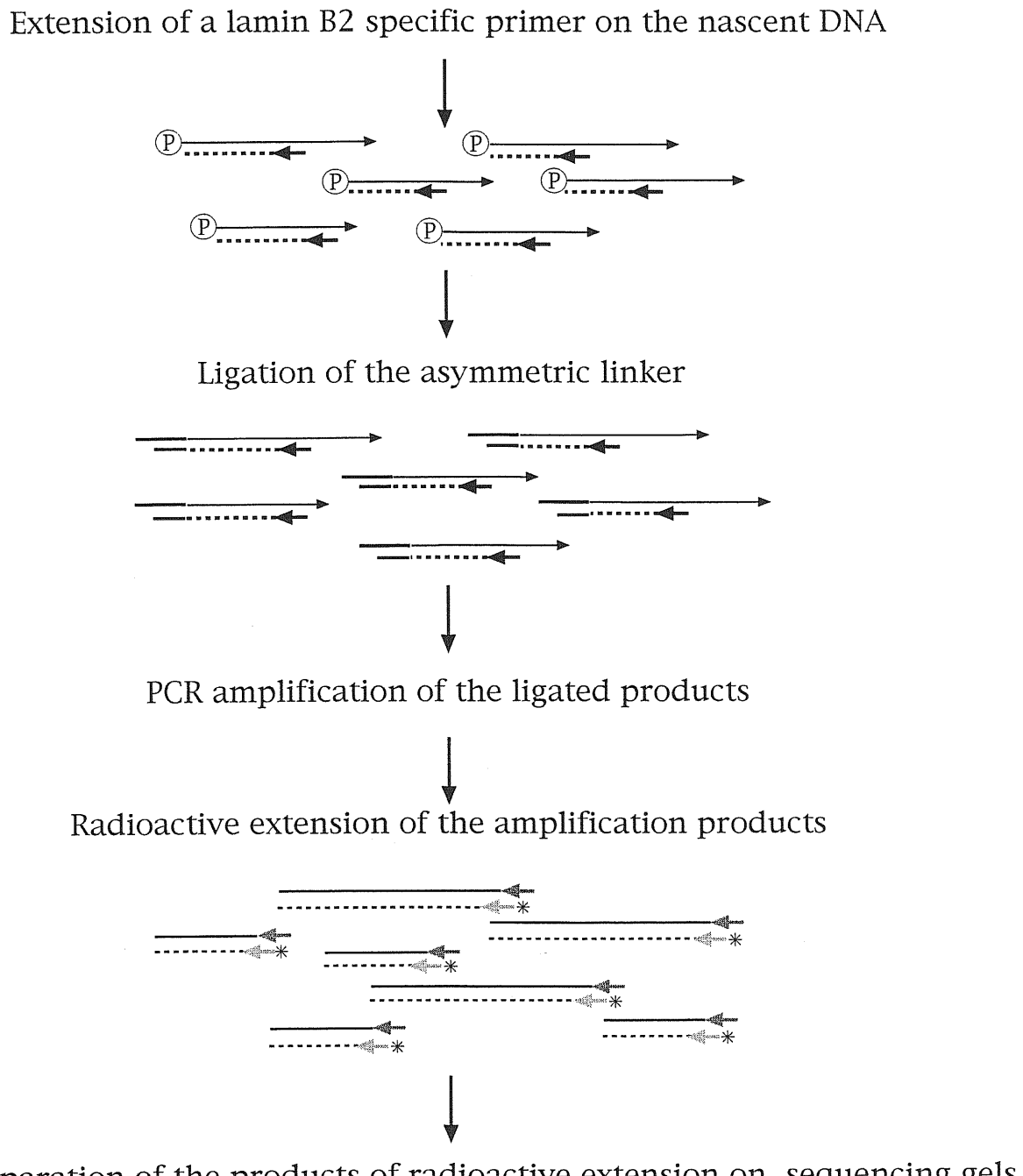

Figure 3.2b: LM-PCR analysis of nascent DNA (see materials and methods).

(fig. 3.3 panel A) demonstrate that, indeed, also in this case we obtain an over five-fold enrichment of ori sequences even before λ-exonuclease digestion (fig. 3.3 panel B). Therefore the size selection by sucrose gradient fractionation was proven to be a valuable tool to purify nascent DNA, applicable also in systems different than human cells.

Before moving to the LM-PCR analysis of the nascent DNA so isolated, we decided to ascertain that the ligation and amplification steps do not introduce biases or distortions with respect to the single step primer extension procedure in the detection of stop sites of second strand synthesis (Santocanale et al., 1997). We therefore compared the size distribution of non-amplified and amplified second strands. To this purpose, total yeast genomic DNA was isolated, randomly fragmented by dimethyl sulfate treatment and subjected to single-step primer extension and to LM-PCR using suitable primers distributed along the ARS 1 region (fig. 3.4 panel A). The resulting autoradiographs are shown in Figure 3.4, panel B; the patterns of the single-step extended fragments and of the LM-PCR amplified ones are exactly superimposable, allowing for the difference in size given by the linker which is ligated to the double strand molecule generated by the extension of the specific primer along the single strand *in vitro* modified DNA (see fig. 3.2b). This observation dispels the doubt that the ligation and amplification treatment may give an unfaithful representation of the relative abundance of the different fragments present in a given sample. Thus, we should be able to reproduce the identification of the start sites of DNA synthesis at the yeast ARS 1 ori performed by Bielinsky and Gerbi, utilizing the technique we intend to apply to our ori.

The λ–exonuclease treated material was analyzed by LM-PCR using two sets of primers located on either side of the ARS 1 ori (fig. 3.4, panel A). As the data show, we identify exactly the same start sites as detected previously by Bielinsky and Gerbi (fig. 3.4, panel C), but the much greater sensitivity offered by the LM-PCR amplification allowed us to use an amount of nascent DNA lower by five orders of magnitude. The procedure we have established and here described appears thus adequate to identify at the nucleotide level the start sites of bidirectional replication in genomes of much greater complexity, including the human one.

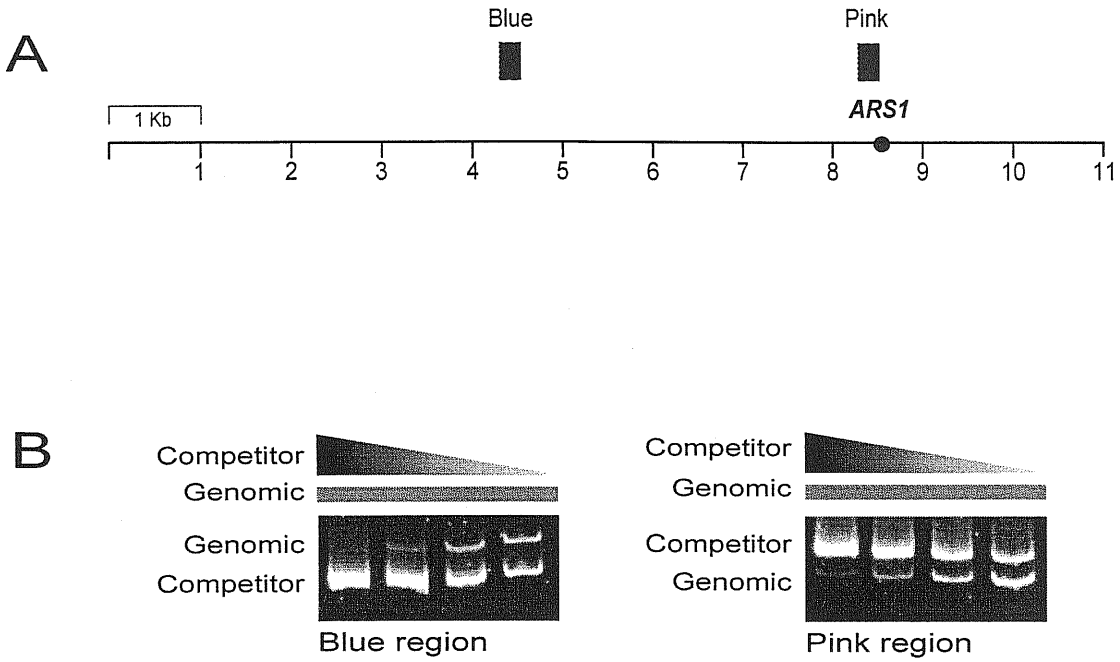

Figure 3.3: Validation of the assay on the yeast chromosomal ARS1 ori: evaluation of the enrichment of ori containing sequences in samples of size fractionated yeast DNA.

(A) Schematic representation of the yeast genomic region containing the ARS1 ori. The positions of the segments amplified by competitive PCR are indicated by the black boxes at the top.

(B) Quantitation of the abundance of selected genomic markers in samples of size-fractionated, newly replicated DNA isolated from exponentially growing yeast cells. The location of the primers, their nucleotide sequences and construction of the competitor are described in Materials and Methods. The competitor and genomic DNA amplification products are indicated. A fixed amount of size fractionated DNA was challenged against serial fourfold dilution of competitor. The number of competitor molecules inserted in the PCR reactions ranges from 10^4 to 1.5 x 10^2.

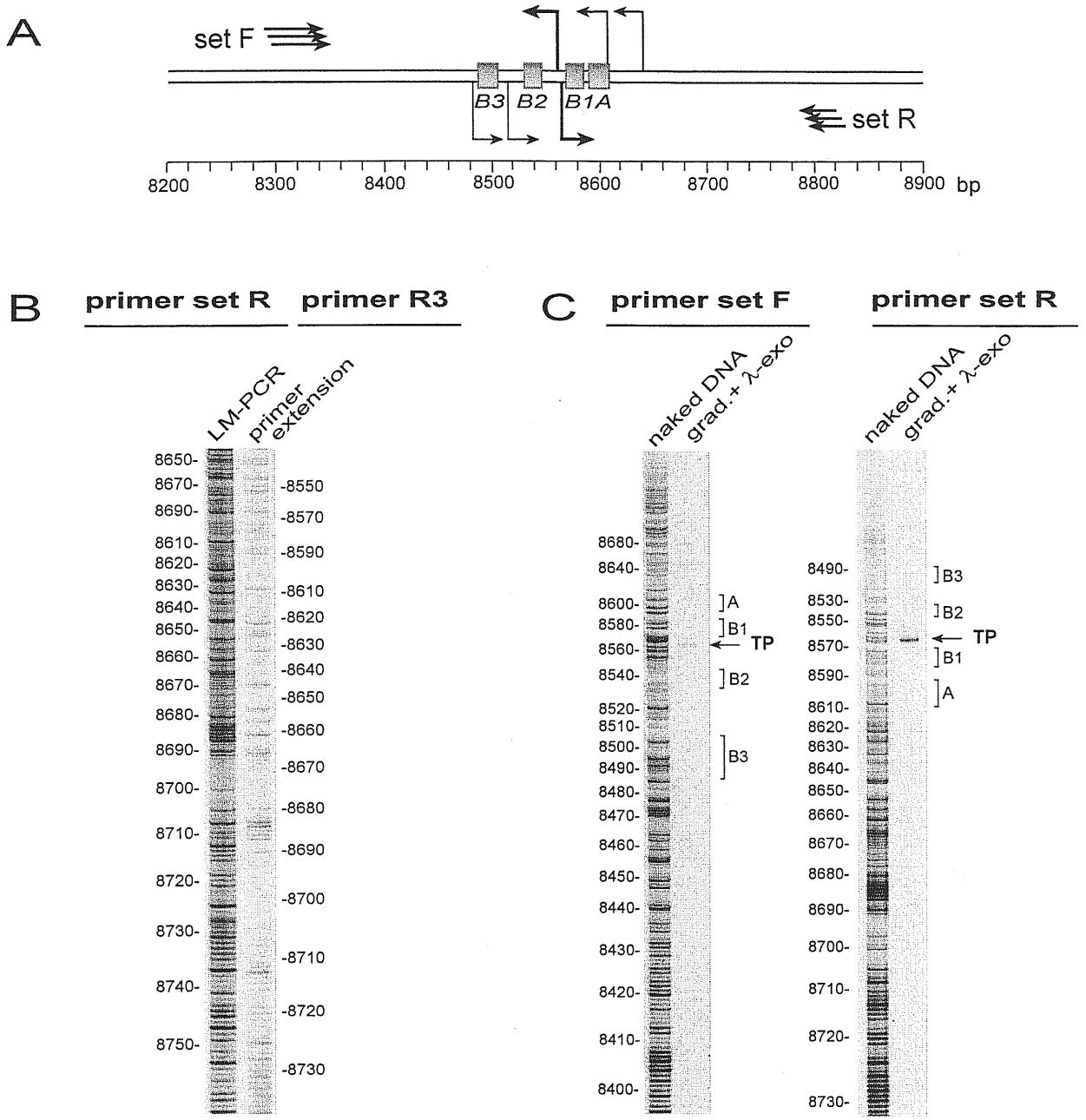

Figure 3.4: Validation of the assay on the yeast chromosomal ARS1 ori: assessment of LM-PCR as a tool to map start sites of nascent DNA synthesis.

(A) Yeast ARS1 region. The positions of the sequence elements A, B1, B2 and B3 are marked by boxes. The clusters of arrows indicate the localization and orientation of primer sets used for LM-PCR.

(B) Comparison of the size distribution of amplified and non-amplified molecules. Yeast genomic DNA was subjected to in vitro fragmentation by dimethyl sulfate (DMS) and analyzed by LM-PCR (primer set R) and by primer extension (primer R3).

(C) LM-PCR analysis of the size fractionated, λ-exonuclease digested nascent DNA isolated from asynchronously growing yeast cells. The primer sets F and R were used to reveal the position of RNA-DNA junctions in the nascent DNA complementary to the upper and lower parental strand. The positions of the transition points (TP) from continuous to discontinuous DNA synthesis are shown and coincide with the ones localized by the RIP mapping technique (Bielinsky and Gerbi, 1999; see introduction).

IDENTIFICATION OF PRECISE START SITES OF NASCENT DNA SYNTHESIS WITHIN THE ORIGIN AREA

In all the systems studied so far in sufficient detail the start sites are located within or very close to the ori region interacting with the initiation specific proteins; hence, in our search for the possible start sites, we moved initially from two primer sets located on either side of the protected region (see the map of the area in fig. 3.5, panel A) and firing into it towards each other, namely primer sets D and E. The results obtained with primer set D on nascent DNA from asynchronously growing HeLa cells are shown in fig. 3.5, panel B; even without λ-exonuclease treatment discrete stop sites of second strand synthesis are visible, the first one corresponding to nucleotide 3933, plus a second one at nucleotide 4076, and a still further one for which a precise size cannot be measured; these data are much more evident in the sample treated with the λ-exonuclease, in agreement with the greater purification of nascent chains given by this treatment. Thus, we have evidence for the presence of an RNA-DNA junction at nucleotide 3933, i. e. within the sequence protected *in vivo*, as well as of, at least, another one located 143 nucleotides upstream, outside the protected area.
The analysis with primer set E (fig. 3.5, panel B), gives comparable, specular results: two clear stop sites of second strand synthesis (i. e. RNA-DNA junctions) are visible, one at nucleotide 3930, within the *in vivo* protected area, and the other one at nucleotide 3787, i. e. 143 nucleotide upstream and well outside the protection.
In order to confirm that we are detecting indeed the RNA-DNA junctions corresponding to the initiation of new DNA strands, we explored their detectability in non-DNA-duplicating cells; to this purpose, in the first place, we performed the same experiment just described on size-fractionated DNA from quiescent human lymphocytes: the quiescent cell DNA, as expected, does not give evidence of any stop site for primer set D (fig 3.5, panel B). Furthermore, we collected cells synchronized in G1 phase and immediately after the G1/S border, extracted their DNA and analyzed it with primer set E; the results shown in fig. 3.5, panel B show clearly that, whereas no stop sites are visible in G1 cells, those that have just entered S (the lamin B2 ori is one of the first to fire in S (Tribioli et al., 1987)) show precise stop sites in exactly the same positions as the asynchronously growing cells.

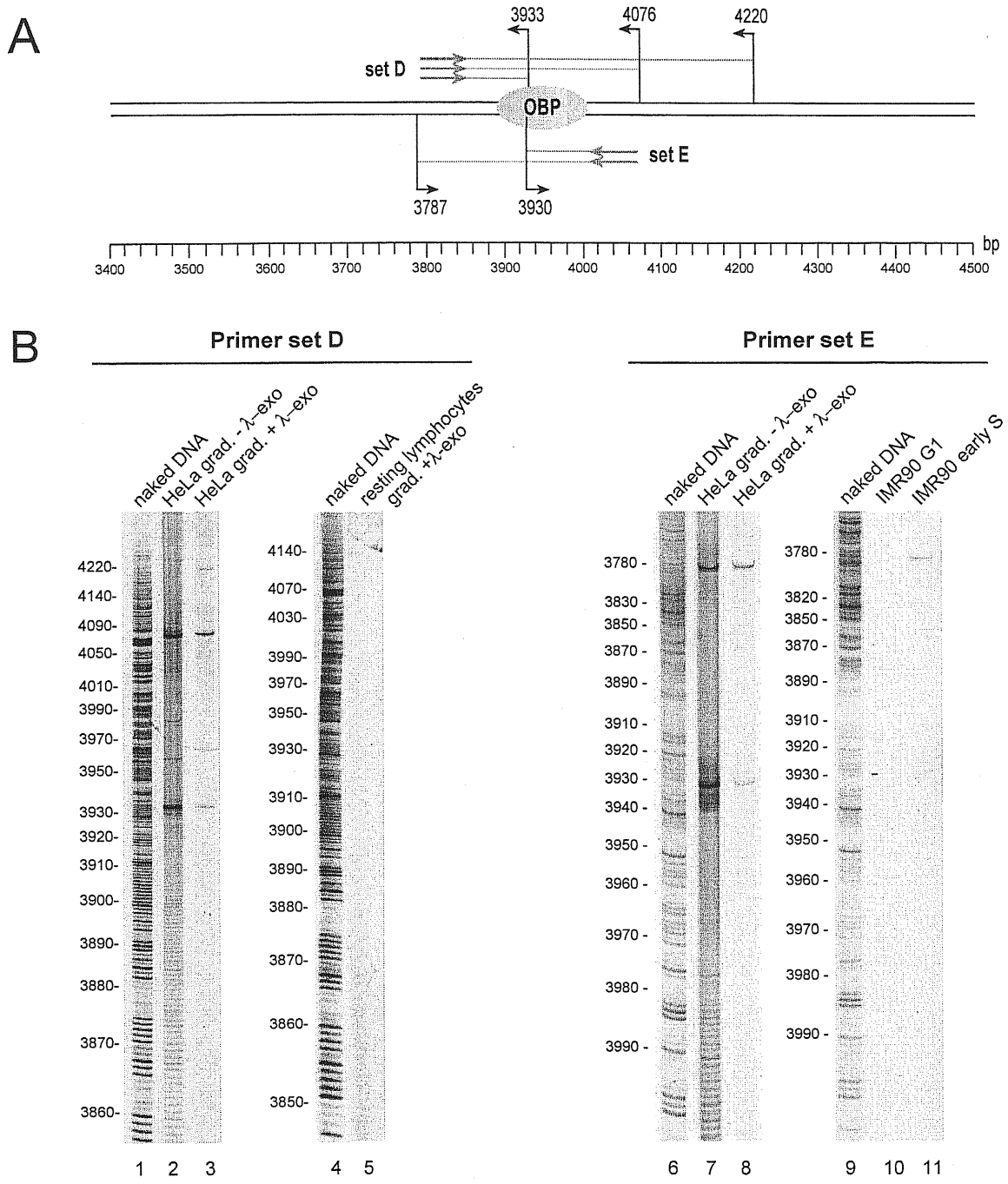

Figure 3.5: Identification of precise start sites of nascent DNA synthesis within the protected area.

(A) Schematic representation of the DNA-protein interactions (origin binding proteins, OBP) at the lamin B2 ori. The cluster of arrows indicates the localization and orientation of the primer sets D and E used to identify the 5' ends of the nascent DNA, complementary to the upper and lower parental strands, respectively. The thin lines connected to the primers indicate the lengths of their relative extended products. Start sites of the nascent DNA synthesis are shown by arrows, pointing into the direction of DNA synthesis.

(B) LM-PCR analysis of size-fractionated DNA not subjected (lanes 2, 5 and 7) or subjected (lanes 3 and 8) to λ-exonuclease treatment, and of total DNA isolated from cells synchronized at the indicated phase of cell cycle (lanes 10 and 11). Primer set D was used to visualize the position of RNA-DNA junctions on the newly synthesized DNA complementary to the upper strand from nucleotides 3850 to 4200. Primer set E was used to visualize the position of RNA-DNA junctions on the newly synthesized DNA complementary to the lower strand from nucleotides 4000 to 3740. Lanes 1, 4, 6 and 9 are Maxam and Gilbert G-sequencing reactions performed on naked DNA. The numbering on the sequences refers to the file humlambbb of GenBank (accession number M94363).

We can thus conclude at this point that the observed stop sites of second strand synthesis correspond indeed to the start sites of nascent DNA synthesis. Furthermore, it is tempting to surmise that nucleotides 3933 and 3930, located well within the protected area, might correspond to the starts of the leading strands moving leftward and rightward, respectively, whereas the RNA-DNA junctions observed approximately 140 nucleotides upstream of both and outside the protected area could derive from the addition of the first Okazaki fragment to the respective leading strand start. We therefore enquired whether primers displaced at different distances on either side of the proposed start sites give results consistent with this interpretation.

EXTENSION OF THE SEARCH FOR START SITES ON THE NASCENT DNA COMPLEMENTARY TO THE UPPER PARENTAL STRAND

In order to investigate more extensively the distribution of RNA-DNA junctions in the ori region, we first analyzed the nascent DNA complementary to the upper strand with primer sets H, D, G and B (see the map in fig. 3.6, panel A for their position). If we move closer to the putative start site of bidirectional DNA replication with primer set H, displaced by only 50 nucleotides from primer set D, we obtain the results shown in fig. 3.6, panel B: in the same experiment also primer set D is used in parallel, and, as already shown above, it indicates again nucleotide 3933 as likely start site of a leading strand and nucleotide 4076 for the possible start of the first lagging strand fragment. In perfect agreement with this result, primer set H shows stop sites at nucleotides exactly corresponding to the two just mentioned, plus a third one, in a position not clearly resolved (around nucleotide 4200) but compatible with that of another Okazaki fragment start. In any case this observation supports the conclusion that we are dealing with replication intermediates and not with some kind of fragmentation caused by the procedure.
If we now use primer set G, that is displaced by over 300 nucleotides on the left of nucleotide 3933, i. e. at a distance greater than the length of an Okazaki fragment (Burhans et al., 1990), we obtain the result shown in fig. 3.6, panel B: only one single stop site is visible, at a position not clearly measurable in view of the poor resolution of the high-molecular size region of the electropherogram, but compatible with nucleotide 3933. The long blank region before this smallest detectable fragment demonstrates that we are dealing here with continuous DNA synthesis, indicating nucleotide 3933 as the transition point

between continuous and discontinuous DNA replication. Thus, the data obtained with three different primer sets (D, G and H) all indicate nucleotide 3933 as the start site for the leftward leading strand.
A confirmation of this tentative conclusion is given by the results obtained with primer set B, located approximately 200 nucleotides on the right of nucleotide 3933. The results are reported in Figure 3.6, panel B; we observe here several stop sites, three of which can be clearly located at nucleotides 4220, 4285 and 4365, respectively, whereas two, still more distant ones, cannot be clearly located. The data can thus be interpreted as deriving from the linking to the initiated ori of four subsequent Okazaki fragments, with lengths of 143 (for the one identified with primer sets D and H above), 144, 65 and 80 nucleotides from the first to the fourth, respectively; conversely, this result is compatible with the presence of multiple alternate start sites of leading strand synthesis.
That, also for the results observed on the nascent DNA on the right side, we are indeed dealing with replication intermediates (and not with some preferential fragmentation of the DNA induced by the treatment) is shown by the data of Figure 3.6, panel C: in this case the experiment with primer set B was performed also in cells synchronized in G1, at the beginning of S, and in mitosis: whereas the cells that have just initiated DNA synthesis show the presence of fragments of the already observed sizes, the cells not synthesizing DNA are totally devoid of them.

EXTENSION OF THE SEARCH FOR START SITES ON NASCENT DNA COMPLEMENTARY TO THE LOWER PARENTAL STRAND

For the extension of the analysis of the nascent DNA complementary to the lower strand we utilized primer sets C, A, and F. Primer sets C and A are displaced by approximately 250 and 450 nucleotides, respectively, to the right of the start site mapped with primer set E at nucleotide 3930 (cf. two paragraphs above) whereas primer set F is located approximately 190 nucleotides to the left of the same nucleotide (see fig. 3.7, panel A for the position of the primer sets). With primer set C (see fig. 3.7, panel B) we obtained again a clear indication that the first stop site corresponds to nucleotide 3930, whereas a second stop, at a length too high to be exactly localized, is certainly compatible with corresponding to nucleotide 3787, indicated by primer set E as a possible start of the first retrograde lagging strand fragment. When

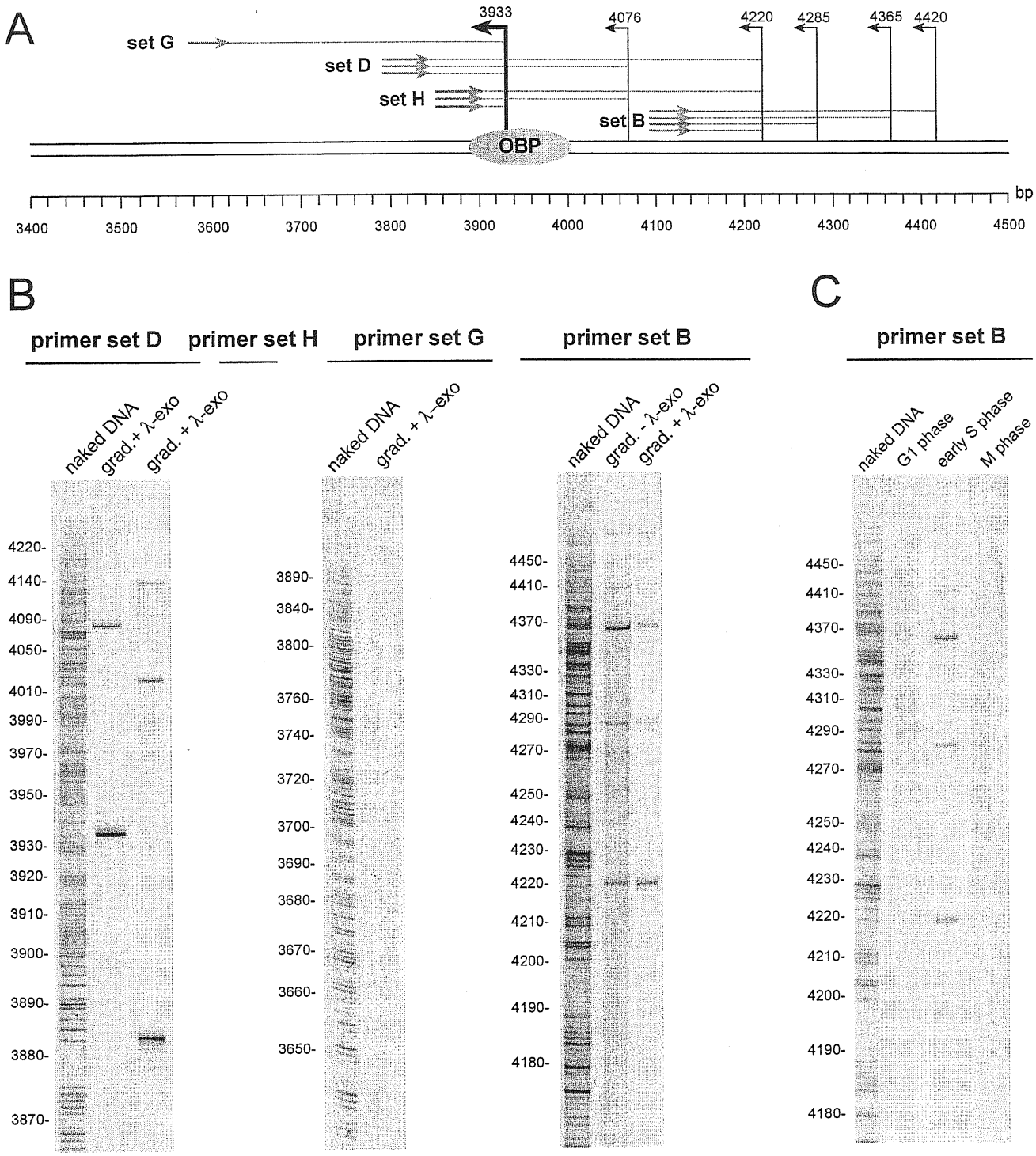

Figure 3.6: Distribution of the start sites of nascent DNA synthesis in a 1kb region centered around the ori. Lower nascent strand.

(A) Localization and orientation of the primer sets D, H, G and B used to identify the position of RNA-DNA junctions on the lower nascent strand from nucleotide 3400 to 4500. The position of the first detectable start site of leading strand synthesis is indicated by a thick arrow, whereas the start sites more upstream are indicated by thinner arrows, all pointing in the direction of synthesis.

(B) LM-PCR analysis of size-fractionated newly replicated HeLa cells DNA subjected to λ-exonuclease treatment.

(C) LM-PCR analysis of total genomic DNA isolated from cells synchronized at the indicated phase of the cell cycle.

we used primer set A (fig. 3.7, panel B), only one stop site at a very large size (compatible with nucleotide 3930) was observed; the long, over 450 nucleotide region free of stop sites indicates that we are dealing with an area of continuous leading strand synthesis, and that nucleotide 3930, corresponding to the first stop site detected by primer sets E, C and A, is the start site of the rightward moving leading strand and hence the transition point between continuous and discontinuous synthesis for the nascent DNA complementary to the parental lower strand.
When we move decidedly to the left of the proposed leading strand start, using primer set F, we obtain the results reported in fig. 3.7 panel B; these show clearly a number of start sites (resistant to λ-exonuclease) that can be located precisely at positions 3645, 3588, 3553, 3501, and, less precisely, around positions 3435 and 3360. As in the case of the data obtained with primer set B (see previous paragraph) for the rightward moving fork, these data (including those of fig. 3.6, panels B and C) can be interpreted as evidence that, in the advancement of the fork moving leftward, the first Okazaki fragments synthesized and ligated to the leading strand have sizes, in the order of movement, of 143, 142, 67, 35, 52, plus two more of approximately 65 nucleotides. Alternatively, again, those 5' ends of nascent DNA could arise from a multiplicity of start sites of leading strand synthesis.
Thus, the analysis of the replication events performed with 8 primer sets spanning a region of almost 1 kb allowed us to identify nucleotides 3933 and 3930 (well within the protected area of the ori region) as the transition points between continuous and discontinuous DNA synthesis for the nascent DNA complementary to the upper and lower parental strand, respectively. At this stage the alternative possibility that the stop sites of second strand synthesis could arise from a multiplicity of start sites of leading strand synthesis cannot be ruled out.

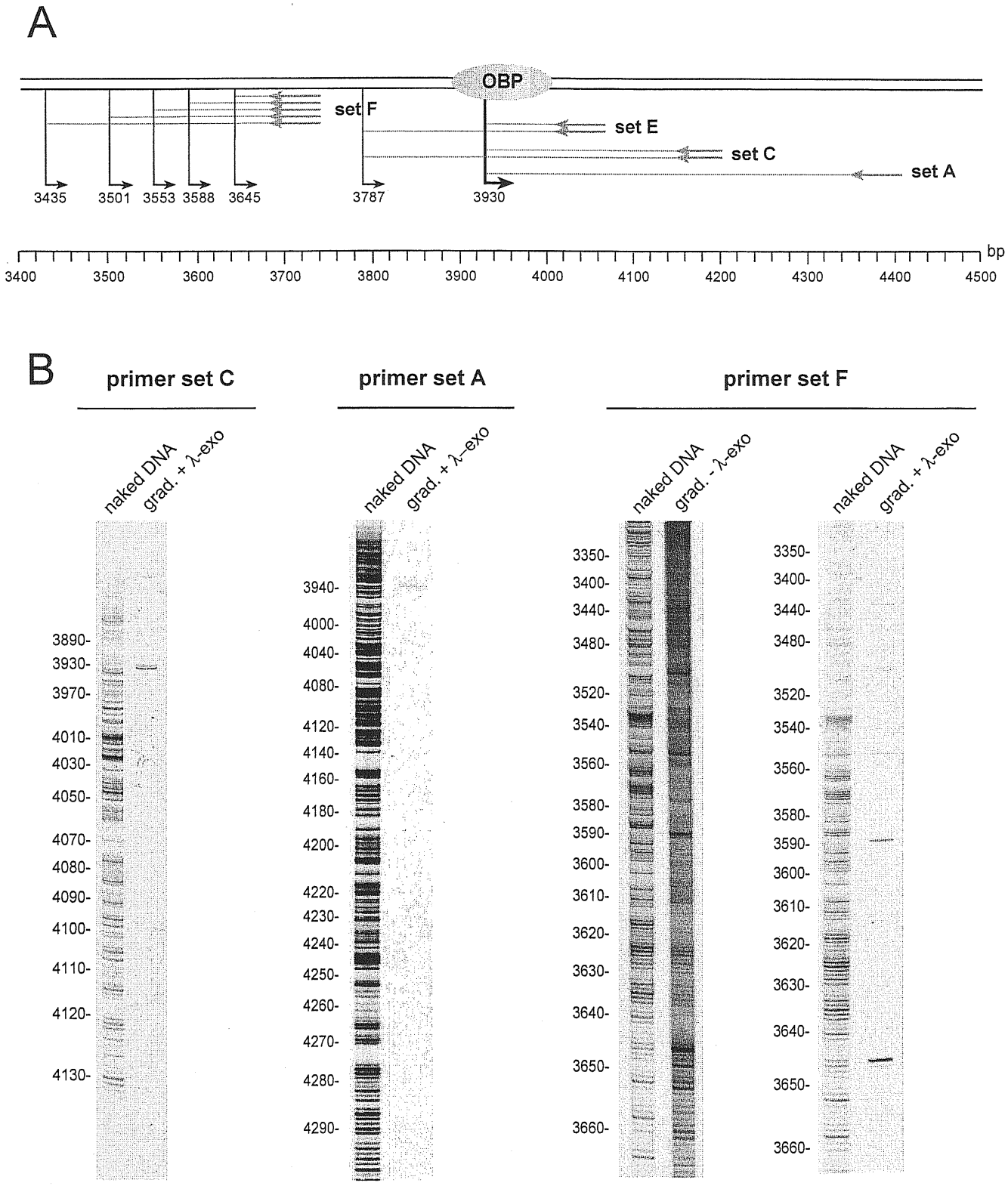

Figure 3.7: Distribution of start sites of nascent DNA synthesis in a 1kb region centered around the ori. Upper nascent strand.

(A) Localization and orientation of the primer sets A, C, E and F used to identify the position of RNA-DNA junctions on the upper nascent strand from nucleotide 4350 to 3400. The positions of the first detectable start site of leading strand synthesis and of the more upstream ones are indicated by a thick and thinner arrow(s) respectively.

(B) LM-PCR analysis of size-fractionated nascent DNA isolated from asynchronously growing HeLa cells and subjected or not subjected (as indicated) to λ-exonuclease digestion.

EFFECT OF THE INHIBITION OF OKAZAKI FRAGMENT SYNTHESIS ON START SITE DETECTION

If our first interpretation is correct, we should expect a fading or disappearance of the start sites corresponding to the Okazaki fragments by treating the cells in a way that inhibits their synthesis, as it occurs when they are incubated in presence of emetine (a protein synthesis inhibitor that has been shown reproducibly to affect eukaryotic DNA replication by blocking Okazaki fragment synthesis through a still unknown mechanism, Burhans et al., 1991). Thus, we first incubated HeLa cells with emetine for one hour, a condition in which Okazaki fragment synthesis is strongly inhibited. Nascent DNA was then isolated and probed for the presence of start sites by the procedure described above, using primer sets D and E, both of which have before given evidence (see fig. 3.5) of leading strand starts, accompanied by putative Okazaki fragment starts. As the results reported in fig. 3.8, panel B show, for both primers the emetine treatment does not affect the intensity of the signals at the positions corresponding to the starts of the leading strands, whereas it causes a fading, down to disappearance, of the signals interpreted as Okazaki fragment starts.
Furthermore, HeLa cells were synchronized at the G1/S border and released in S-phase for one hour in presence of emetine; their DNA was extracted and challenged with our procedure using again primer sets D and E (fig. 3,8, panel C). The results show clearly that, also in these cells, whereas the signals of the start sites of the leading strands are not affected by the treatment, the signals of the fragments of greater size, corresponding to the starts of Okazaki fragments, are significantly reduced on both sides of the ori.
These results demonstrate that at the lamin B2 origin the leading strand has a unique start site. The leading strand start site coincides with the transition point between continuos and discontinuous DNA synthesis, from which bidirectional DNA replication fires in all the cells that are initiating DNA synthesis from this ori at the moment of the analysis, since no polymorphism in the selection of the leading strand start sites is detectable. The other initiation events revealed at the area are due to Okazaki fragment synthesis.

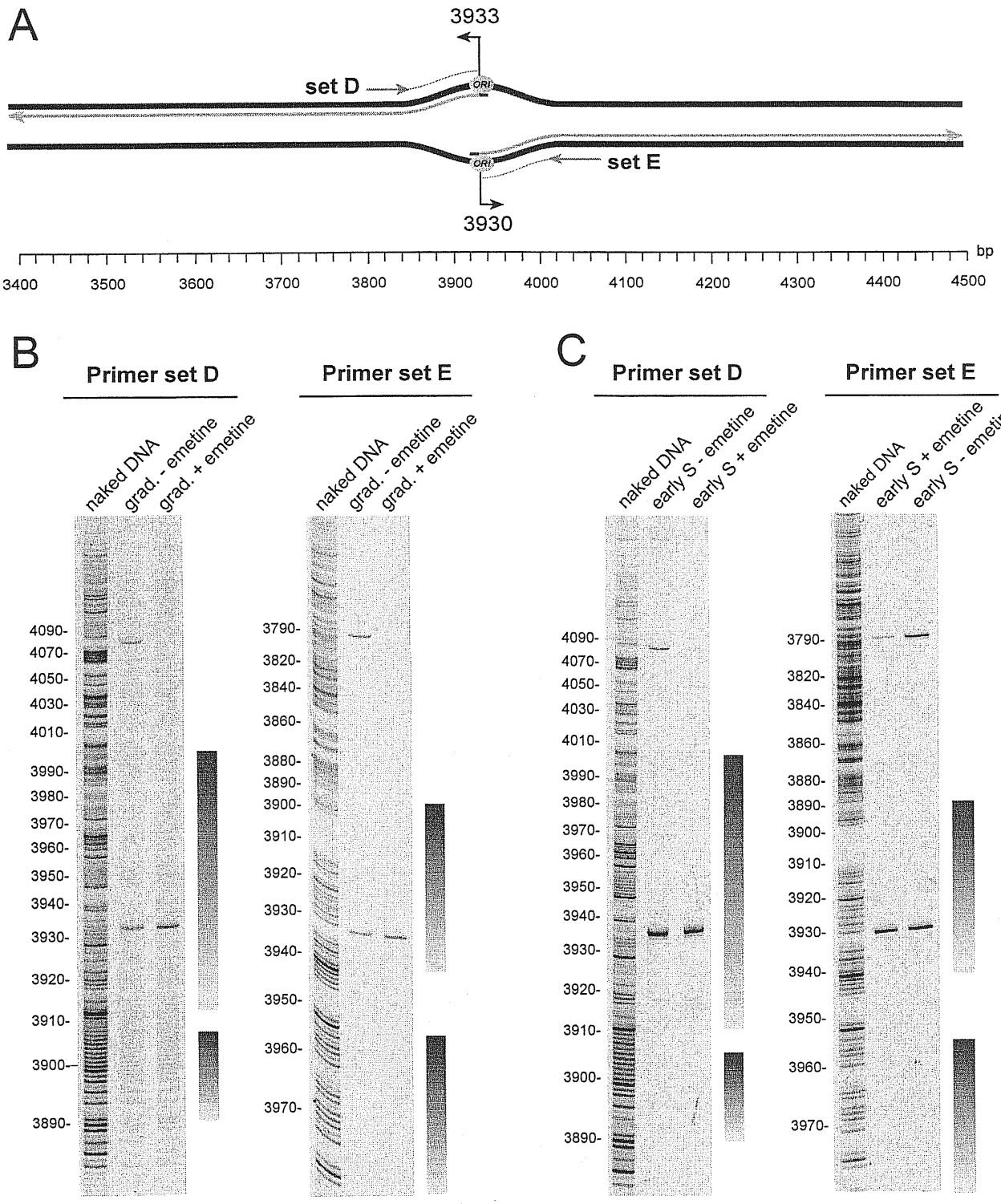

Figure 3.8: Lamin B2 ori has a single start site of leading strand synthesis.

(A) Map of the start sites of leading strand synthesis within the lamin B2 ori. The positions of the 5' ends of the nascent DNA synthesized in the presence of emetine on the two helices are indicated by black arrows pointing in the direction of DNA synthesis.

(B) LM-PCR analysis of size-fractionated, λ-exonuclease treated nascent DNA isolated from asynchronous cells growing in the presence or absence of emetine. The sequence corresponding to the OBP site is indicated by boxes.

(C) LM-PCR analysis of total, λ-exonuclease treated DNA extracted from cells synchronized at the beginning of S phase in the presence or absence of emetine.

DISCUSSION

When the replicon model was postulated (Jacob et al., 1963), it was tempting to consider it as a general paradigm to explain origin function in any organism, from *E. coli* to metazoans. It was later demonstrated that the model holds true for eukaryotic viruses and for *S. cerevisiae,* in which the start sites of bidirectional DNA replication were mapped at nucleotide level (Hay and DePamphilis, 1982; Hendrickson et al., 1986; Bielinsky and Gerbi, 1998, 1999). In these systems the sites of replication initiation are in fact localized in the region protected by the specific proteins acting as initiator factors: T Ag, EBNA1 and ORC complex respectively in the case of SV40, Epstein-Barr virus and budding yeast. So far, it was impossible to evaluate the validity of the model at metazoan origins of replication since none of the origins mapped at metazoan chromosomes were described at nucleotide level. The data presented in this thesis demonstrate that replication origins analogous to the ones described for simple organisms exist in metazoans, confirming the general validity of the replicon model, in spite of the major differences distinguishing yeast and viral origins from the those present in higher eukaryotes. On the contrary of what is observed in simple organisms, metazoan origins do not present apparently a modular anatomy, are used at low frequency in S-phase and need to be reestablished at any new G1, because the initiator proteins are removed from the chromatin during the condensation of the chromosomes in mitosis.

SETTING AND VALIDATION OF AN ASSAY TO MAP ORIGINS OF BIDIRECTIONAL REPLICATION IN GENOMES OF ANY DEGREE OF COMPLEXITY

In order to map the start of bidirectional replication at nucleotide level at a human chromosome, we attempted to set a method with an improved selectivity and sensitivity in comparison with the techniques previously designed to map the transition point between continuos and discontinuous DNA synthesis in simple genomes.
The assay that we established is based on the analysis of actively replicating sequences, isolated by nascent DNA purification via size fractionation of total DNA extracted from asynchronously growing cells. All the origins active at the moment of the analysis are represented in this preparation. In alternative, we performed the analysis on total DNA extracted from cells synchronized in early S-phase. The rationale of this approach is that just short DNA can productively react in the conditions of the assay. The only short DNA stretches present in the sample derived from early S cells

should come from early firing origins. Lamin B2 region, known to fire in the first 2 minutes of S-phase (Tribioli et al., 1982) is part of this origin sub-population. Size fractionated DNA or total DNA extracted from synchronized cells is further treated with λ-exonuclease, which selectively preserves the RNA-primed nascent strands digesting the aspecifically broken DNA. The sample is then analyzed by LM-PCR. The newly synthesized DNA is annealed to primers specific for the region of interest and extended up to the 5' end of the template which corresponds to the start site of DNA synthesis. The size fractionation or alternatively the synchronization, coupled with the λ-exonuclease treatment, selects for active origin sequences reducing the complexity of the sample under study. Nonetheless, the signal specific for the region of interest would be undetectable, if it would not be enhanced by an amplification step. The double stranded molecule obtained from the extension with the specific primer is therefore ligated to a double stranded linker. This step creates a population of amplifiable molecules, bearing at one end a recognition site for the selected genomic primer and at the other end the recognition site for the linker primer (fig. 3.2b). The amplification improves dramatically the sensitivity of the method allowing to localize 5' ends of nascent DNA in complex genomes. The concern in using amplification is that the PCR step might introduce artefacts due to the preferential amplification of some ligated molecules, resulting in a bias in the detection of nascent DNA 5' ends. The comparison between hot extension and LM-PCR performed on *in vitro* modified yeast DNA shown in fig. 3.4, panel B, demonstrates that this is not the case, since the bands are equally represented in both the samples.

Once the reliability of the LM-PCR was demonstrated, we tested the whole procedure attempting to localize with our approach the 5' ends of nascent DNA at the ARS1 ori, already mapped by Gerbi and collaborators using a different strategy (Bielinsky and Gerbi, 1998, 1999). Our data are in perfect agreement with the data previously obtained. Therefore these results validate our system of nascent DNA isolation and analysis. The advantage of this assay in comparison with the other replication start sites mapping approaches is that its increased selectivity and sensitivity allow to reduce by four order of magnitude the amount of material needed to perform the analysis, offering therefore a tool to map start sites of DNA synthesis in genomes of any degree of complexity.

AT THE LAMIN B2 ORIGIN, INITIATION EVENTS ARE TAKING PLACE IN THE REGION PROTECTED BY A HUMAN ORIGIN BINDING COMPLEX

The lamin B2 origin was mapped by nascent strand abundance analysis via competitive PCR in a region of roughly 500 bp where multiple specific DNA-protein interactions are taking place. The most prominent protein-DNA interaction observed is proliferation dependent and is modulated along the cell cycle strongly resembling the assembly of pre and post-replicative complex at yeast origins (Dimitrova et al., 1996; Abdurashidova et al., 1998). Even if these findings do not demonstrate a direct role of the proteins that generate the footprints in the control of DNA replication, they hint at them as at the mammalian counterparts of the initiator proteins in yeast. The search for start sites of DNA replication at the lamin B2 origin reveals that initiation events are occurring in the region interacting with the cell cycle modulated binding activity. This activity can therefore be considered as a human origin binding complex, since it is bound to a region in which initiation of DNA replication occurs. The detected stop sites of second strand synthesis represent indeed 5' end of nascent DNA because are resistant to the treatment with λ-exonuclease, and their appearance is proliferation dependent and cell cycle regulated, being absent in quiescent and in G1 cells, where DNA synthesis does not take place.

THE LAMIN B2 ORIGIN PRESENTS A UNIQUE START SITE OF LEADING STRAND SYNTHESIS

Figure 4.1 gives a summary of our results and shows the protein-DNA interactions occurring in the ori area during the cell cycle along with the nucleotide sequence at and around the initiation sites. The monitoring of the initiation events in a region of 1 kb encompassing the origin area points to nucleotide 3933, inside the region protected by proteins both in G1 and in S phase, as the transition point between continuos and discontinuous DNA synthesis for the nascent DNA complementary to the parental upper strand. On the other strand, the transition point between continuos and discontinuous DNA synthesis is located at nucleotide 3930. Since the further stop sites of second strand synthesis are disappearing when the cells are treated with emetine which blocks the lagging strand synthesis, they represent 5' ends of Okazaki fragments sequentially ligated to

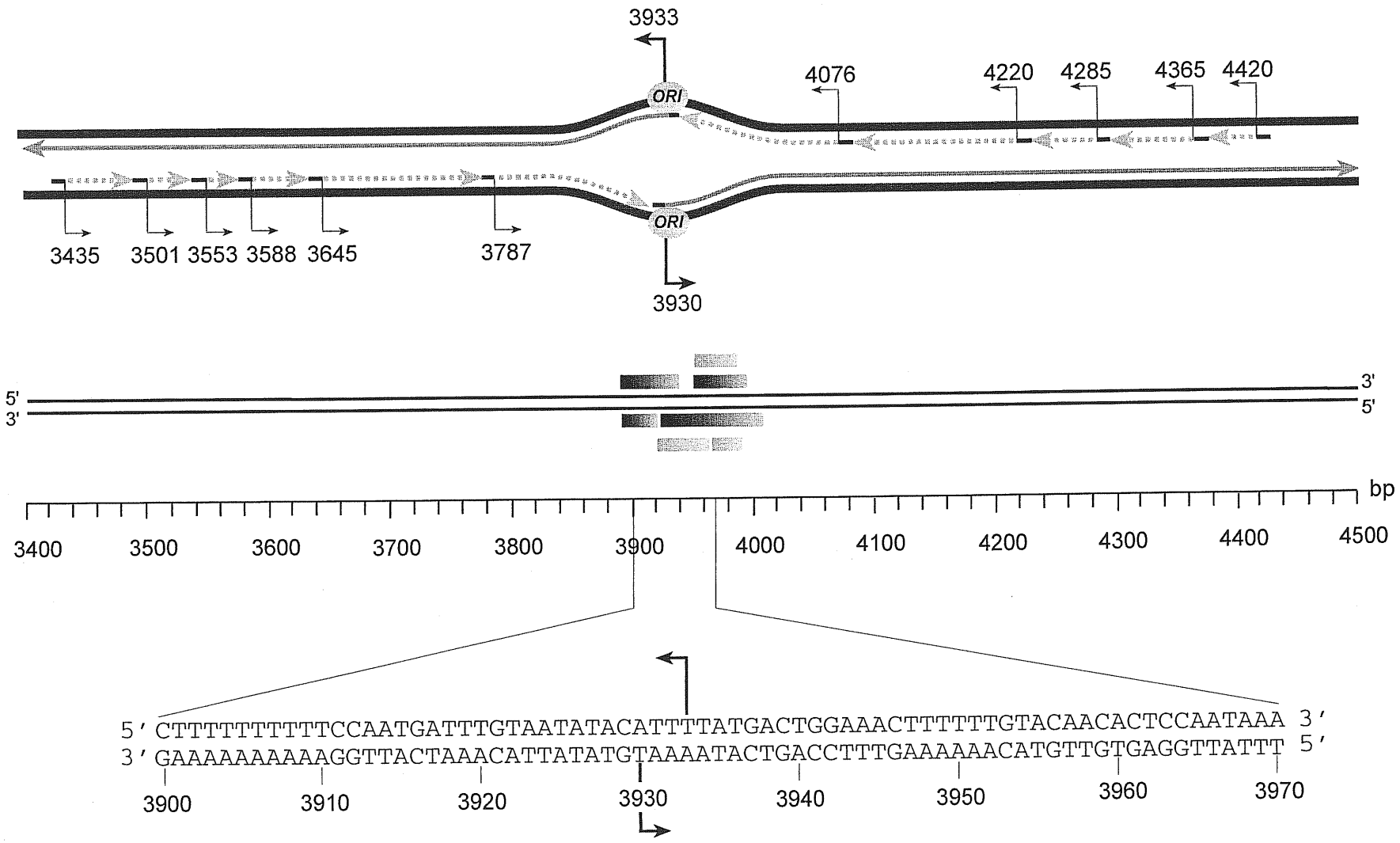

Figure 4.1: Map of the start sites of bi-directional DNA synthesis at the human lamin B2 ori.

The transition points from continuous to discontinuous DNA synthesis on the two strands are shown in the upper part of the figure by thick arrows; the 5' ends of Okazaki fragments are indicated by thin arrows. The middle part of the figure shows the position of pre- (dark boxes) and post- (pale boxes) replicative complexes observed in vivo in G1 and S phase of the cell cycle, respectively. In the lower part of the figure the sequence of the region containing the start sites of leading strand synthesis is also shown.

the leading strand and not alternative start site of leading strand synthesis. These data demonstrate that the lamin B2 origin presents a unique start site of leading strand synthesis at each parental strand. Therefore, when the origin is used, the leading strand synthesis fires from the same nucleotide without polymorphism in the selection of the start site among a population of cells. The two leading strand start sites are overlapping, defining an Origin of Bidirectional Replication of three nucleotide, which is located inside the region interacting in a cell cycle dependent fashion with the human origin binding proteins. The transition point occurs in an AT-rich region which could act as a DUE (DNA unwinding element). The significance of the sequence elements at which initiation occurs, will be understood only when a genetic dissection via mutation analysis of the origin locus will be available.

SIZE DISTRIBUTION OF OKAZAKI FRAGMENTS AT THE ORIGIN REGION

The finding that the lamin B2 has a unique start site of leading strand synthesis indicates that the origin determination is a tightly regulated process. The situation is similar at the chromosomal ARS1 in yeast (Bielinsky et al., 1999) and indicates that defined constraints of chromosomal architecture both yeast and human localize initiation complexes on a fixed position in an homomorphyc chromatin environment. In human cells this process should take place at any cell cycle when the cell exits from mitosis and enters a new G1, since the protein complex that marks the lamin B2 origin disappears in M phase and is reestablished in the following G1 phase. Probably the chromatin structure defined at the origin region by the binding of a pre-replicative complex which in turn recruits the replisome is responsible for the tight regulation of the length of the Okazaki fragments adjacent to the origin. In fact, from our data it is apparent that, on the retrograde arms, in both directions, the first two Okazaki fragments are of identical length (143 nt). These data hint at the possibility that during origin activation and initial movement, the two forks issuing from one origin may be strictly correlated with each other and possibly physically bound in a single, double replisome. The control over Okazaki fragment length is lost while the fork proceeds since the following Okazaki fragments appear to initiate more frequently, giving rise to shorter fragments, ranging in size between approximately 70 and as few as 35 nucleotides. An alternative interpretation of these data is that polymorphism in the selection of start sites of the Okazaki fragments takes place moving outward

from the origin region. If this is true, the 5' ends detected more frequently as the fork proceeds, constitute alternative start sites of Okazaki fragment synthesis.

MAPPING INITIATION EVENTS AT METAZOAN ORIGINS: INITIATION SITES AND INITIATION ZONES.

As detailed in the introduction, the origin determination in metazoans is a multifactorial process in which key roles are played by nuclear environment, chromatin structure, transcriptional state of the region and epigenetic modification of the origin sequence (DePamphilis et al., 1999). It is possible to hypothesize that origin determination is a combinatorial event in which several elements contribute to define a place to start bidirectional replication. It is therefore possible that several origin types exist, some, analogous to the lamin B2 origin, presenting a unique start site of leading strand synthesis, whereas others presenting a region in which many sites are available to initiate bidirectional replication. Our assay, based on the selective detection of 5' end of nascent DNA, allows to map initiation events both in case of localized origin and in case of initiation events delocalized in a broad initiation zone. As detailed in the introduction, mapping data obtained by 2-D gel techniques hint at the existence of broad initiation zones at some metazoan replication origins (Little et al., 1993, Dijkwel et al., 1995), whereas data obtained either by nascent strand length and abundance analysis and by fork polarity assays suggest the presence of localized initiation sites at these same regions (Burhans et al., 1990; Kobayashi et al., 1998). The assay that we set may offer a tool to reconcile the data from 2-D gels, which give a comprehensive picture of the replication state of a region monitoring active and passive replication structures, and the data from nascent strand abundance analysis, which detect the molecules actively replicating at the moment of the study. Our system could in fact reveal the major initiation events at the primary initiation sites defined by nascent strand abundance analysis, but is sensitive enough to monitor the nearby low frequency events, possibly responsible for the detection of bubbles arcs throughout broad initiation zones.

FUTURE PERSPECTIVE

Identification of start site of bidirectional DNA replication and origin binding proteins at different chromosomal location in metazoan genomes

The assay that we set up could allow us to map the start site of bidirectional replication at nucleotide level at any metazoan origin, provided that the candidate region has already been localized in an area of around 1kb by mapping techniques with a lower definition. Several human mammalian origins fulfil these requirements (Kitsberg et al., 1994; Berberich et al., 1995, Delgado et al., 1998, Kobayashi et al., 1998) and therefore may be defined at molecular level by means of our assay. This analysis may give information on the sequence at which replication starts at different origins, leading to the identification of common sequence elements possibly acting as DUE and/or recognition sites for origin binding proteins and may clarify if the homologues of the yeast components of pre and post-replicative complexes are associated with metazoan origins. Orc, Cdc6 and Mcm proteins appear in fact to be essential in all the metazoan system tested so far and extensive studies on the replication licensing phenomenon in the *Xenopus* system allow to define a stepwise assembly of the replication initiation complex at the chromatin similar to the assembly of pre-replication complex in yeast. Chromatin-bound ORC complex recruits Cdc6 protein which probably acts by loading the MCM complex to chromatin, that may in turn act as the replicative helicase, locally unwinding the DNA and allowing bidirectional DNA replication to start specifically from the origin region (Coleman et al., 1996; Romanowski et al., 1996; Rowles et al., 1996; Yshimi et al 1997; Perkins and Diffley, 1998). At the moment there are not sufficient data available to assess the validity at such a model at metazoan replication origins *in vivo*, but the high degree of phylogenetic conservation among the proteins crucial for replication initiation (Dutta and Bell, 1997; Donaldson and Blow, 1999) both in yeast and metazoans, leads to the hypothesis that the general principles which govern replication initiation are evolutionary conserved, pointing to the Orc, Mcm and Cdc6 proteins as the best candidates to be initiator proteins in metazoan.

The identification of the components of the origin binding complex will offer the opportunity to analyze the functional significance of each bound protein in defining the leading strand start site at the origin region, and would allow to discriminate if these complexes

are common or differ from origin to origin in dependence of different chromatin localization (active versus inactive chromatin domains), time of replication in S-phase, time of activation during development.

Cell-cycle control of origin activation and definition of the start site of leading strand synthesis at metazoan origins

The major players in controlling DNA replication initiation are also the target of cell cycle mediated regulation, which provides a link between replication initiation and control of cell cycle progression. Several evidences are indicating that cell cycle regulators are acting on different putative components of metazoan preinitiation complex, to licence origins to fire and ensure once per cell cycle replication. The major regulators of cell cycle progression (i.e. G1 cyclin-dependent kinases and E2F/pRb complex) control the activity of the putative metazoan origin binding proteins either by regulating their level of transcription or by direct interaction with the initiation complex bound at the origins (Cardoso et al., 1993; Elsasser et al., 1996; Saha et al., 1998). As in the case of yeast, origins are licensed to fire in G1 when the activity of CDKs is low. CDKs activation during cell cycle progression leads to the phosphorylation of key components of the initiation complex such as Cdc6 and Mcm proteins, which are displaced from the chromatin leaving an inactive initiation complex which cannot promote reinitiation of DNA replication during the same cell cycle (Elsasser et al., 1996, Jallepalli and Kelly, 1996; Kimura et al., 1995; Todorov et al., 1995; Hendrickson et al., 1996, Lei et al., 1996). There is also increasing evidence that the expression and the localization of candidate initiator proteins in metazoans are developmentally regulated. In *Drosophila* embryo, the level of Orc 1 subunit changes dramatically during development, accumulating preferentially in the active proliferating cells (Asano and Warton, 1999) and both Orc1 and Orc2 are redistributed selectively at the amplification loci when chorion gene amplification occurs at the follicle cells (Asano and Warton 1999; Royzman et al., 1999).
The availabily of the information about the precise localization of the start sites of bidirectional replication would provide a tool to look at the metazoan origins *in vivo* and to analyze the functional significance of cell cycle and developmental stimuli acting on the initiator proteins in regulating origin activity and leading strand start site selection. The initiation of DNA replication may be analyzed considering the changing in the pattern of initiation

events taking place at the origin in response to growth or differentiation signals.

Developmental control of origin activation and definition of the start site of leading strand synthesis at metazoan origins

Metazoan origins are dynamic structures subjected to developmental control (Francon et al., 1999).
The plasticity in origin determination is modulated by the multiple parameters defining a origin which allow its activity to be modified during the establishment of the different developmental programmes through which a multicellular organism is formed. Since interaction with origin specific proteins is reset at any M-phase (Abdurashidova et al., 1998) it is likely that a time window exists in G1 in which origin definition takes place and developmental stimuli may act in reorganising the chromatin structure changing the pattern of sites available to initiation complex binding. This hypothesis is confirmed by the fact that in early G1 nuclei, forced to replicate in *Xenopus* egg extract, DNA synthesis starts randomly, while after a certain point in G1 origin definition takes place and replication initiation is localized at particular sites in the genome (Wu and Gilbert, 1996, 1997). It is possible that a reorganization of the chromatin structure is established at the ODP, which focuses the assembly of the replication initiation complex to specific sequences. A similar phenomenon seems to take place in the *Xenopus* development, when after the midblastula stage, replication origin previously randomly distributed along the genome become localized at defined loci (Hyrien et al., 1995). Considering these findings, it is not surprising to observe chromosomal aberrations possibly due to amplification of some regions of the genome when there is a perturbation in the activity of key regulators of the state of the chromatin like SWI/SNF complex and histone deacetylases (Avithal et al., 1999; Kim et al., 1999; Koipally et al., 1999). The assay that we set offers the opportunity to map the start site of DNA replication in different conditions, and to monitor the changes in start site selection in response to developmental signals. Start site identification may be coupled with structural studies to identify the chromatin features that are critical in driving the assembly of the initiator complex on a peculiar sequence, and to understand how chromatin remodelling acts on the definition of a place at which to start replication in the metazoan genomes.

CONCLUSION

We designed an assay to map the start sites of bidirectional DNA replication in genomes of any degree of complexity; the assay is based on LM-PCR analysis of nascent strands from actively replicating origins. The increase in selectivity, given by the nascent DNA isolation procedure and in sensitivity, given by the specific amplification of the region under study, offers a tool to identify 5' ends of nascent DNA in genomes of any degree of complexity. The assay has been validated on the yeast ARS1 chromosomal origin, giving results perfectly compatible to the ones obtained with a different mapping strategy at nucleotide definition (Bielinsky et al., 1999).

The mapping results of the start sites of leading strand synthesis at the lamin B2 origin region are comparable to the data obtained at the same resolution for yeast and eukaryotic viruses, and reveal that the lamin B2 ori is located in a specific genomic region in which leading strand synthesis starts homogeneously from two unique nucleotides that define the transition point between continuos and discontinuous DNA synthesis at both parental strands. The transition points are located inside an area interacting specifically with proteins in a cell cycle dependent fashion, resembling the situation in yeast, in which the replication start sites are located near the region interacting with the putative initiator element: the ORC complex. These findings demonstrate for the first time that the replicon model is valid also in metazoans and that origins of replication analogous to the ones described at simpler genomes exist in higher eukaryotes, regardless the major differences distinguishing metazoan origins from origins in simple genomes. Metazoan origins in fact do not share apparently any sequence homology, are used at low frequency at any S-phase and are not marked by binding with specific proteins throughout the cell cycle (Abdurashidova et al., 1998). This phenomenon would allow origins to change during the development and to be redefined in response to the various differentiation programmes that enter in the formation of a metazoan organism.

ACKNOWLEDGEMENTS

I'm grateful to Prof. A. Falaschi for having offered me the opportunity to work for a Ph.D. in his lab.
The data presented in this thesis are the result of a team work performed by Prof. G. Abdurashidova and me. I'm thankful to Prof. Abdurashidova for the continuos help and support during the experimental work and for having provided me with her extraordinary expertise.
My sincere thanks go to Dott. Raffaella Klima due to her fundamental contribution in obtaining the data on *S. cerevisiae* ARS1.
Finally I would like to thank all my colleagues for helpful discussions, especially Ms. Vesna Todorovic, who contributed in compiling part of this thesis, and all my friends for their patience and moral support during the last years.

REFERENCES

Abdurashidova G, Riva S, Biamonti G, Giacca M, Falaschi A (1998). *Cell cycle modulation of protein-DNA interactions at a human replication origin.* EMBO J 17, 2961-9.

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K. editors (1994). *Current Protocols in Molecular Biology.* Wiley, New York.

Aladjem MI, Groudine M, Brody LL, Dieken ES, Fournier RE, Wahl GM, Epner EM (1995). *Participation of the human beta-globin locus control region in initiation of DNA replication.* Science 270, 815-9.

Aladjem MI, Rodewald LW, Kolman JL, Wahl GM (1998). *Genetic dissection of a mammalian replicator in the human beta-globin locus.* Science 281, 1005-9.

Anachkova B and Hamlin JL (1989). *Replication in the amplified dihydrofolate reductase domain in CHO cells may initiate at two distinct sites, one of which is a repetitive sequence element.* Mol Cell Biol 9, 532-40.

Aparicio OM, Weinstein DM, Bell SP (1997). *Components and dynamics of DNA replication complexes in S. cerevisiae: redistribution of MCM proteins and Cdc45p during S phase.* Cell 91, 59-69.

Araujo FD, Knox JD, Ramchandani S, Pelletier R, Bigey P, Price G, Szyf M, Zannis-Hadjopoulos M (1999). *Identification of initiation sites for DNA replication in the human dnmt1 (DNA-methyltransferase) locus.* J Biol Chem 274, 9335-41.

Ariizumi K, Wang Z, Tucker PW (1993). *Immunoglobulin heavy chain enhancer is located near or in an initiation zone of chromosomal DNA replication.* Proc Natl Acad Sci U S 90, 3695-9.

Asano M, Wharton RP (1999). *E2F mediates developmental and cell cycle regulation of ORC1 in Drosophila.* EMBO J 18, 2435-48.

Avitahl N, Winandy S, Friedrich C, Jones B, Ge Y, Georgopoulos K (1999). *Ikaros sets thresholds for T cell activation and regulates chromosome propagation.* Immunity 10, 333-43.

Baker TA and Bell SP (1998). *Polymerases and the replisome: machines within machines.* Cell 92, 295-305.

Bell SP and Stillman B (1992). *ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex.* Nature 357, 128-34.

Bell SP, Kobayashi R, Stillman B (1993). *Yeast origin recognition complex functions in transcription silencing and DNA replication.* Science 262, 1844-9.

Bell SP, Mitchell J, Leber J, Kobayashi R, Stillman B (1995). *The multidomain structure of Orc1p reveals similarity to regulators of DNA replication and transcriptional silencing.* Cell 83, 563-8.

Berberich S, Trivedi A, Daniel DC, Johnson EM, Leffak M (1995). *In vitro replication of plasmids containing human c-myc DNA.* J Mol Biol 245, 92-109.

Biamonti G, Giacca M, Perini G, Contreas G, Zentilin L, Weighardt F, Guerra M, Della Valle G, Saccone S, Riva S, et al (1992a). *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, 3499-506.

Biamonti G, Perini G, Weighardt F, Riva S, Giacca M, Norio P, Zentilin L, Diviacco S, Dimitrova D, Falaschi A (1992b). *A human DNA replication origin: localization and transcriptional characterization.* Chromosoma 102, S24-31.

Bielinsky AK and Gerbi SA (1998). *Discrete start sites for DNA synthesis in the yeast ARS1 origin.* Science 279, 95-8.

Bielinsky AK and Gerbi SA (1999). *Chromosomal ARS1 has a single leading strand start site.* Mol Cell 3, 477-8.

Blow JJ, Laskey RA (1988). *A role for the nuclear envelope in controlling DNA replication within the cell cycle.* Nature 332, 546-8.

Botz J, Zerfass-Thome K, Spitkovsky D, Delius H, Vogt B, Eilers M, Hatzigeorgiou A, Jansen-Durr P (1996). *Cell cycle regulation of the murine cyclin E gene depends on an E2F binding site in the promoter.* Mol Cell Biol 16, 3401-9.

Bousset K and Diffley JF (1998). *The Cdc7 protein kinase is required for origin firing during S phase.* Genes Dev 12, 480-90.

Bramhill D and Kornberg A (1988). *A model for initiation at origins of DNA replication.* Cell 54, 915-8.

Brewer BJ and Fangman WL (1987). *The localization of replication origins on ARS plasmids in S. cerevisiae.* Cell 51, 463-71.

Brewer BJ and Fangman WL (1991). *Mapping replication origins in yeast chromosomes.* Bioessays 13, 317-22.

Brewer BJ, Diller JD, Friedman KL, Kolor KM, Raghuraman MK, Fangman WL (1993). *The topography of chromosome replication in yeast.* Cold Spring Harb Symp Quant Biol 58, 425-34.

Brown GW and Kelly TJ (1998). *Purification of Hsk1, a minichromosome maintenance protein kinase from fission yeast.* J Biol Chem 273, 22083-90.

Burhans WC, Selegue JE, Heintz NH (1986a). *Isolation of the origin of replication associated with the amplified Chinese hamster dihydrofolate reductase domain.* Proc Natl Acad Sci U S A 83, 7790-4.

Burhans WC, Selegue JE, Heintz NH (1986b). *Replication intermediates formed during initiation of DNA synthesis in methotrexate-resistant CHOC 400 cells are enriched for sequences derived from a specific, amplified restriction fragment.* Biochemistry 25, 441-9.

Burhans WC, Vassilev LT, Caddle MS, Heintz NH, DePamphilis ML (1990). *Identification of an origin of bidirectional DNA replication in mammalian chromosomes.* Cell 62, 955-65.

Burhans WC, Vassilev LT, Wu J, Sogo JM, Nallaseth FS, DePamphilis ML (1991). *Emetine allows identification of origins of mammalian DNA replication by*

imbalanced DNA synthesis, not through conservative nucleosome segregation. EMBO J 10, 4351-60.

Burkhart R, Schulte D, Hu D, Musahl C, Gohring F, Knippers R (1995). *Interactions of human nuclear proteins P1Mcm3 and P1Cdc46.* Eur J Biochem 228, 431-8.

Cardoso MC, Leonhardt H, Nadal-Ginard B (1993). *Reversal of terminal differentiation and control of DNA replication: cyclin A and Cdk2 specifically localize at subnuclear sites of DNA replication.* Cell 74, 979-92.

Carpenter PB, Mueller PR, Dunphy W (1996). *Role for a Xenopus Orc2-related protein in controlling DNA replication.* Nature 379, 357-60.

Carroll SM, DeRose ML, Kolman JL, Nonet GH, Kelly RE, Wahl GM (1993). *Localization of a bidirectional DNA replication origin in the native locus and in episomally amplified murine adenosine deaminase loci.* Mol Cell Biol 13, 2971-81.

Celniker SE, Sweder K, Srienc F, Bailey JE, Campbell JL (1984). *Deletion mutations affecting autonomously replicating sequence ARS1 of Saccharomyces cerevisiae.* Mol Cell Biol 4, 2455-66.

Chong JP, Mahbubani HM, Khoo CY, Blow JJ (1995). *Purification of an MCM-containing complex as a component of the DNA replication licensing system.* Nature 375, 418-21.

Chong JP, Thommes P, Blow JJ (1996). *The role of MCM/P1 proteins in the licensing of DNA replication.* Trends Biochem Sci 21, 102-6.
coamplification of competitive templates. Gene 122, 313-20.

Cocker JH, Piatti S, Santocanale C, Nasmyth K, Diffley JF (1996). *An essential role for the Cdc6 protein in forming the pre-replicative complexes of budding yeast.* Nature 379, 180-2.

Coffman FD, Georgoff I, Fresa KL, Sylvester J, Gonzalez I, Cohen S (1993). *In vitro replication of plasmids containing human ribosomal gene sequences: origin localization and dependence on an aprotinin-binding cytosolic protein.* Exp Cell Res 209, 123-32.

Coleman TR, Carpenter PB, Dunphy WG (1996). *The Xenopus Cdc6 protein is essential for the initiation of a single round of DNA replication in cell-free extracts.* Cell 87, 53-63.

Coue M, Amariglio F, Maiorano D, Bocquet S, Mechali M (1998). *Evidence for different MCM subcomplexes with differential binding to chromatin in Xenopus.* Exp Cell Res 245, 282-9.

Crossley M and Orkin SH (1993). *Regulation of the beta-globin locus.* Curr Opin Genet Dev. 1993 3, 232-7.

Dahmann C, Diffley JF, Nasmyth KA (1995). *S-phase-promoting cyclin-dependent kinases prevent re-replication by inhibiting the transition of replication origins to a pre-replicative state.* Curr Biol 5, 1257-69.

Delgado S, Gomez M, Bird A, Antequera F (1998). *Initiation of DNA replication at CpG islands in mammalian chromosomes.* EMBO J 17, 2426-35.

DePamphilis ML (1993). *Eukaryotic DNA replication: anatomy of an origin.* Annu Rev Biochem 62, 29-63.

DePamphilis ML (1996). *Origins of DNA replication.* In DePamphilis ML, editor. DNA replication in eukaryotic cells, 45-89. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

DePamphilis ML (1997). *The search for origins of DNA replication.* Methods 13, 211-9.

DePamphilis ML (1999). *Replication origins in metazoan chromosomes: fact or fiction?* Bioessays 21, 5-16.

Deshpande AM and Newlon CS (1992). *The ARS consensus sequence is required for chromosomal origin function in Saccharomyces cerevisiae.* Mol Cell Biol 12, 4305-13.

Detweiler CS and Li JJ (1998). *Ectopic induction of Clb2 in early G1 phase is sufficient to block prereplicative complex formation in Saccharomyces cerevisiae.* Proc Natl Acad Sci U S A 95, 2384-9.

Diffley JF (1996). *Once and only once upon a time: specifying and regulating origins of DNA replication in eukaryotic cells.* Genes Dev 10, 2819-30.

Diffley JF (1998). *Replication control: choreographing replication origins.* Curr Biol 8, R771-3.

Diffley JF and Cocker JH (1992). *Protein-DNA interactions at a yeast replication origin.* Nature 357, 169-72.

Diffley JF and Stillman B (1988). *Purification of a yeast protein that binds to origins of DNA replication and a transcriptional silencer.* Proc Natl Acad Sci U S A 85, 2120-4.

Diffley JF, Cocker JH, Dowell SJ, Rowley A (1994). *Two steps in the assembly of complexes at yeast replication origins in vivo.* Cell 78, 303-16.

Dijkwel PA and Hamlin JL (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, 3715-22.

Dijkwel PA and Hamlin JL (1995). *The Chinese hamster dihydrofolate reductase origin consists of multiple potential nascent-strand start sites.* Mol Cell Biol 15, 3023-31.

Dijkwel PA, Vaughn JP, Hamlin JL (1991). *Mapping of replication initiation sites in mammalian genomes by two-dimensional gel analysis: stabilization and enrichment of replication intermediates by isolation on the nuclear matrix.* Mol Cell Biol 11, 3850-9.

Dillin A, Rine J (1998). *Roles for ORC in M phase and S phase.* Science 279, 1733-7.

Dimitrova DS, Giacca M, Demarchi F, Biamonti G, Riva S, Falaschi A (1996). *In vivo protein-DNA interactions at human DNA replication origin.* Proc Natl Acad Sci U S A 93, 1498-503.

Diviacco S, Norio P, Zentilin L, Menzo S, Clementi M, Biamonti G, Riva S, Falaschi A, Giacca M (1992). *A novel procedure for quantitative polymerase chain reaction by coamplification of competitive templates.* Gene 122, 313-20

Donaldson AD, Blow JJ (1999). *The regulation of replication origin activation.* Curr Opin Genet Dev 9, 62-8.

Donaldson AD, Fangman WL, Brewer BJ (1998). *Cdc7 is required throughout the yeast S phase to activate replication origins.* Genes Dev 12, 491-501.

Donaldson AD, Raghuraman MK, Friedman KL, Cross FR, Brewer BJ, Fangman WL (1998). *CLB5-dependent activation of late replication origins in S. cerevisiae.* Mol Cell 2, 173-82.

Donovan S and Diffley JF (1996). *Replication origins in eukaroytes.* Curr Opin Genet Dev 6, 203-7.

Donovan S, Harwood J, Drury LS, Diffley JF (1997). *Cdc6p-dependent loading of Mcm proteins onto pre-replicative chromatin in budding yeast.* Proc Natl Acad Sci U S A 94, 5611-6.

Dowell SJ, Romanowski P, Diffley JF (1994). *Interaction of Dbf4, the Cdc7 protein kinase regulatory subunit, with yeast replication origins in vivo.* Science 265, 1243-6.

Drury LS, Perkins G, Diffley JF (1997). *The Cdc4/34/53 pathway targets Cdc6p for proteolysis in budding yeast.* EMBO J 16, 5966-76.

Dutta A and Bell SP (1997). *Initiation of DNA replication in eukaryotic cells.* Annu Rev Cell Dev Biol 13, 293-332.

Ellis DJ, Jenkins H, Whitfield WG, Hutchison CJ (1997). *GST-lamin fusion proteins act as dominant negative mutants in Xenopus egg extract and reveal the function of the lamina in DNA replication.* J Cell Sci 110, 2507-18.

Elsasser S, Lou F, Wang B, Campbell JL, Jong A (1996). *Interaction between yeast Cdc6 protein and B-type cyclin/Cdc28 kinases.* Mol Biol Cell 7, 1723-35.

Engel JD (1993). *Developmental regulation of human beta-globin gene transcription: a switch of loyalties?* Trends Genet 9, 304-9.

Epner E, Forrester WC, Groudine M (1988). *Asynchronous DNA replication within the human beta-globin gene locus.* Proc Natl Acad Sci U S A 85, 8081-5.

Ermakova OV, Nguyen LH, Little RD, Chevillard C, Riblet R, Ashouian N, Birshtein BK, Schildkraut CL (1999). *Evidence that a single replication fork proceeds from early to late replicating domains in the IgH locus in a non-B cell line.* Mol Cell 3, 321-30.

Ewen ME, Sluss HK, Sherr CJ, Matsushime H, Kato J, Livingston DM (1993). *Functional interactions of the retinoblastoma protein with mammalian D-type*

Fields BN, Knipe DM, Howley MH (1996). *Virology-third edition.* Lippincott-Raven, Philadelphia.

Forrester WC, Epner E, Driscoll MC, Enver T, Brice M, Papayannopoulou T, Groudine M (1990). *A deletion of the human beta-globin locus activation*

region causes a major alteration in chromatin structure and replication across the entire beta-globin locus. Genes Dev 4, 1637-49.

Fox CA, Loo S, Dillin A, Rine J (1995). *The origin recognition complex has essential functions in transcriptional silencing and chromosomal replication.* Genes Dev 9, 911-24.

Francon P, Maiorano D, Mechali M (1999). *Initiation of DNA replication in eukaryotes: questioning the origin.* FEBS Lett 452, 87-91.

Gale JM, Tobey RA, D'Anna JA (1992). *Localization and DNA sequence of a replication origin in the rhodopsin gene locus of Chinese hamster cells.* J Mol Biol 224, 343-58.

Gavin KA, Hidaka M, Stillman B (1995). *Conserved initiator proteins in eukaryotes.* Science *270*, 1667-1671.

Gencheva M, Anachkova B, Russev G (1996). *Mapping the sites of initiation of DNA replication in rat and human rRNA genes.* J Biol Chem 271, 2608-14.

Geng Y, Eaton EN, Picon M, Roberts JM, Lundberg AS, Gifford A, Sardet C, Weinberg RA (1996). *Regulation of cyclin E transcription by E2Fs and retinoblastoma protein.* Oncogene 12, 1173-80.

Gerbi SA and Bielinsky AK (1997). *Replication initiation point mapping.* Methods 13, 271-80.

Giacca M, Zentilin L, Norio P, Diviacco S, Dimitrova D, Contreas G, Biamonti G, Perini G, Weighardt F, Riva S, et al (1994). *Fine mapping of a replication origin of human DNA.* Proc Natl Acad Sci U S A 91, 7119-23.

Giacca M, Pelizon C, Falaschi A (1997). *Mapping replication origins by quantifying relative abundance of nascent DNA strands using competitive polymerase chain reaction.* Methods 13, 301-12.

Gilbert DM (1998). *Replication origins in yeast versus metazoa: separation of the haves and the have nots.* Curr Opin Genet Dev 8, 194-9.

Gilbert DM, Miyazawa H, DePamphilis ML (1995). *Site-specific initiation of DNA replication in Xenopus egg extract requires nuclear structure.* Mol Cell Biol 15, 2942-54.

Gogel E, Langst G, Grummt I, Kunkel E, Grummt F (1996). *Mapping of replication initiation sites in the mouse ribosomal gene cluster.* Chromosoma 104, 511-8.

Goldberg M, Jenkins H, Allen T, Whitfield WG, Hutchison CJ (1995). *Xenopus lamin B3 has a direct role in the assembly of a replication competent nucleus: evidence from cell-free egg extracts.* J Cell Sci 108, 3451-61.

Gossen M, Pak DT, Hansen SK, Acharya JK, Botchan MR (1995). *A Drosophila homolog of the yeast origin recognition complex.* Science 270, 1674-7.

Guo ZS, Heine U, DePamphilis ML (1991). *T-antigen binding to site I facilitates initiation of SV40 DNA replication but does not affect bidirectionality.* Nucleic Acids Res 19, 7081-8.

Handeli S, Klar A, Meuth M, Cedar H (1989). *Mapping replication units in animal cells.* Cell 57, 909-20.

Handeli S, Weintraub H (1992). *The ts41 mutation in Chinese hamster cells leads to successive S phases in the absence of intervening G2, M, and G1.* Cell 71, 599-611.

Hardy CF (1997). *Identification of Cdc45p, an essential factor required for DNA replication.* b) Gene 187, 239-46.

Hardy CF, Dryga O, Seematter S, Pahl PM, Sclafani RA (1997). *mcm5/cdc46-bob1 bypasses the requirement for the S phase activator Cdc7p.* a) Proc Natl Acad Sci U S A 94, 3151-5.

Harland RM and Laskey RA (1980). *Regulated replication of DNA microinjected into eggs of Xenopus laevis.* Cell 21, 761-71.

Hartwell LH (1976). *Sequential function of gene products relative to DNA synthesis in the yeast cell cycle.* J Mol Biol 104, 803-17.

Hateboer G, Wobst A, Petersen BO, Le Cam L, Vigo E, Sardet C, Helin K (1998). *Cell cycle-regulated expression of mammalian CDC6 is dependent on E2F.* Mol Cell Biol 18, 6679-97.

Hatton KS, Dhar V, Brown EH, Iqbal MA, Stuart S, Didamo VT, Schildkraut CL (1988). Replication program of active and inactive multigene families in mammalian cells. Mol Cell 8, 2149-58.

Hay RT and DePamphilis ML (1982). *Initiation of SV40 DNA replication in vivo: location and structure of 5' ends of DNA synthesized in the ori region.* Cell 28, 767-79.

Hay RT, Hendrickson EA, DePamphilis ML (1984). *Sequence specificity for the initiation of RNA-primed simian virus 40 DNA synthesis in vivo.* J Mol Biol 175, 131-57.

Heintz NH and Hamlin JL (1982). *An amplified chromosomal sequence that includes the gene for dihydrofolate reductase initiates replication within specific restriction fragments.* Proc Natl Acad Sci U S A 79, 4083-7.

Heintz NH and Stillman BW (1988). *Nuclear DNA synthesis in vitro is mediated via stable replication forks assembled in a temporally specific fashion in vivo.* Mol Cell Biol 8, 1923-31.

Hendrickson EA, Fritze CE, Folk WR, DePamphilis ML (1987). *The origin of bidirectional DNA replication in polyoma virus.* EMBO J 6, 2011-8.

Hendrickson M, Madine M, Dalton S, Gautier J (1996). *Phosphorylation of MCM4 by cdc2 protein kinase inhibits the activity of the minichromosome maintenance complex.* Proc Natl Acad Sci U S A. 93, 12223-8.

Hennessy KM, Lee A, Chen E, Botstein D (1991). *A group of interacting yeast DNA replication genes.* Genes Dev 5, 958-69.

Hernandez P, Martin-Parras L, Martinez-Robles ML, Schvartzman JB (1993). *Conserved features in the mode of replication of eukaryotic ribosomal RNA genes.* EMBO J 12, 1475-85.

Higuchi R, Krummel B, Saiki RK (1988). *A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions.* Nucleic Acids Res 16, 7351-67.

Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR (1989). *Site-directed mutagenesis by overlap extension using the polymerase chain reaction.* Gene 77, 51-9.

Hochstrasser M (1996). *Protein degradation or regulation: Ub the judge.* Cell 84, 813-5.

Hokak P, Jackson DA, Cook PR (1996). *Role of nuclear structure in DNA replication.* In Blow JJ, editor. Eukaryotic DNA replication. Oxford IRL press, 124-42.

Hopwood B and Dalton S (1996). *Cdc45p assembles into a complex with dc46p/Mcm5p, is required for minichromosome maintenance, and is essential for chromosomal DNA.* Proc Natl Acad Sci U S A. 1996 93, 12309-14.

Hua XH and Newport J (1998). *Identification of a preinitiation step in DNA replication that is ndependent of origin recognition complex and cdc6, but dependent on cdk2.* J Cell Biol 40, 271-81.

Huberman JA, Spotila LD, Nawotka KA, el-Assouli SM, Davis LR (1987). *The in vivo replication origin of the yeast 2 microns plasmid.* Cell 51, 473-81.

Hyrien O, Maric C, Mechali M (1995). *Transition in specification of embryonic metazoan DNA replication origins.* Science 270, 994-7.

Ishimi Y (1997). *A DNA helicase activity is associated with an MCM4, -6, and -7 protein complex.* J Biol Chem. 272, 24508-13.

Ishimi Y, Komamura Y, You Z, Kimura H (1998). *Biochemical function of mouse minichromosome maintenance 2 protein.* J Biol Chem 273, 8369-75.
J Biol Chem 269, 23351-6.

Jackson AL, Pahl PM, Harrison K, Rosamond J, Sclafani RA (1993). *Cell cycle regulation of the yeast Cdc7 protein kinase by association with the Dbf4 protein.* Mol Cell Biol 13, 2899-908.

Jacob F, BrennerS, CuzinF (1963). *On the regulation of DNA replication in Bacteria.* Cold Spring harbor Symp. Quant. Biol. 28, 329-33.

Jallepalli PV, Brown GW, Muzi-Falconi M, Tien D, Kelly TJ (1997). *Regulation of the replication initiator protein p65cdc18 by CDK phosphorylation.* Genes Dev 11, 2767-79.

Jallepalli PV, Kelly TJ (1996). *Rum1 and Cdc18 link inhibition of cyclin-dependent kinase to the initiation of DNA replication in Schizosaccharomyces pombe.* Genes Dev 10, 541-52.

Jallepalli PV, Tien D, Kelly TJ (1998). *sud1(+) targets cyclin-dependent kinase-phosphorylated Cdc18 and Rum1 proteins for degradation and stops unwanted diploidization in fission yeast.* Proc Natl Acad Sci U S A 95, 8159-64.

Johnson DG, Schwarz JK, Cress WD, Nevins JR (1993). *Expression of transcription factor E2F1 induces quiescent cells to enter S phase.* Nature 365, 349-52.

Jonsson ZO and Hubscher U (1997). *Proliferating cell nuclear antigen: more than a clamp for DNA polymerases.* Bioessays 19, 967-75.

Kalejta RF, Li X, Mesner LD, Dijkwel PA, Lin HB, Hamlin JL (1998). *Distal sequences, but not ori-beta/OBR-1, are essential for initiation of DNA replication in the Chinese amster DHFR origin.* Mol Cell 2, 797-806.

Kato J, Matsushime H, Hiebert SW, Ewen ME, Sherr CJ (1993). *Direct binding of cyclin D to the retinoblastoma gene product (pRb) and pRb phosphorylation by the cyclin D-dependent kinase CDK4.* Genes Dev 7, 331-42.

Kearsey SE and Labib K. (1998). *MCM proteins: evolution, properties, and role in DNA replication.* Biochim Biophys Acta *1398*, 113-36.

Kelly RE, DeRose ML, Draper BW, Wahl GM (1995). *Identification of an origin of bidirectional DNA replication in the ubiquitously expressed mammalian CAD gene.* Mol Cell Biol 15, 4136-48.

Kim J, Sif S, Jones B, Jackson A, Koipally J, Heller E, Winandy S, Viel A, Sawyer A, Ikeda T, Kingston R, Georgopoulos K (1999) *Ikaros DNA-binding proteins direct formation of chromatim remodeling complexes in lymphocytes.* Immunity 10, 345-55.

Kimura H, Takizawa N, Nozaki N, Sugimoto K (1995). *Molecular cloning of cDNA encoding mouse Cdc21 and CDC46 homologs and characterization of the products: physical interaction between P1(MCM3) and CDC46 proteins.* Nucleic Acids Res 12, 2097-104.

Kitsberg D, Selig S, Keshet I, Cedar H (1993).*Replication structure of the human beta-globin gene domain.* Nature 366, 588-90.

Kobayashi T, Rein T, DePamphilis ML (1998). *Identification of primary initiation sites for DNA replication in the hamster dihydrofolate reductase gene initiation zone.* Mol Cell Biol 18, 3266-77.

Koipally J, Renold A, Kim J, Georgopoulos K (1999). *Repression by Ikaros and Aiolos is mediated through histone deacetylase.* EMBO J 18, 3090-100.

Kominami K and Toda T (1997). *Fission yeast WD-repeat protein pop1 regulates genome ploidy through ubiquitin-proteasome-mediated degradation of the CDK inhibitor Rum1 and the S-phase initiator Cdc18.* Genes Dev 11, 1548-60.

Kornberg A and Baker T: *DNA replication- second edition.* Freeman, WH and company, New York (1992).

Krude T, Jackman M, Pines J, Laskey RA (1997). *Cyclin/Cdk-dependent initiation of DNA replication in a human cell-free system.* Cell 88, 109-19.

Kubota Y, Mimura S, Nishimoto S, Masuda T, Nojima H, Takisawa H (1997). *Licensing of DNA replication by a multi-protein complex of MCM/P1 proteins in Xenopus eggs.* EMBO J 16, 3320-31.

Kumar S, Giacca M, Norio P, Biamonti G, Riva S, Falaschi A (1996). *Utilization of the same DNA replication origin by human cells of different derivation.* Nucleic Acids Res 24, 3289-94.

Lalande M (1990). *A reversible arrest point in the late G1 phase of the mammalian cell cycle.* Exp Cell Res 186, 332-9.

Laskey RA, Gorlich D, Madine MA, Makkerh JPS, Romanowski P (1996). *Regulatory roles of the nuclear envelope.* Exp Cell Res 229, 204-11.

Lawlis SJ, Keezer SM, Wu JR, Gilbert DM (1996). *Chromosome architecture can dictate site-specific initiation of DNA replication in Xenopus egg extracts.* J Cell Biol 135, 1207-18.

Leatherwood J (1998). *Emerging mechanisms of eukaryotic DNA replication initiation.* Curr Opin Cell Biol 10, 742-8.

Leatherwood J, Lopez-Girona A, Russell P (1996). *Interaction of Cdc2 and Cdc18 with a fission yeast ORC2-like protein.* Nature 379, 360-3.

Lei M, Kawasaki Y, Tye BK (1996). *Physical interactions among Mcm proteins and effects of Mcm dosage on DNA replication in Saccharomyces cerevisiae.* Mol Cell Biol 16, 5081-90.

Lei M, Kawasaki Y, Young MR, Kihara M, Sugino A, Tye BK (1997). *Mcm2 is a target of regulation by Cdc7-Dbf4 during the initiation of DNA synthesis.* Genes Dev 11, 3365-74.

Leno GH, Downes CS, Laskey RA (1992). *The nuclear membrane prevents replication of human G2 nuclei but not G1 nuclei in Xenopus egg extract.* Cell 69, 151-8.

Levine AJ, Kang HS, Billheimer FE (1970). *DNA replication in SV40 infected cells. I. Analysis of replicating SV40 DNA.* J Mol Biol 50, 549-68.

Lewin B (1994). *Genes V.* Oxford University Press, Oxford.

Li JJ and Herskowitz I (1993). *Isolation of ORC6, a component of the yeast origin recognition complex by a one-hybrid system.* Science 262, 1870-4.

Liang C, Spitzer JD, Smith HS, Gerbi SA (1993). *Replication initiates at a confined region during DNA amplification in Sciara DNA puff II/9A.* Genes Dev 7, 1072-84.

Lin HB, Falchetto R, Mosca PJ, Shabanowitz J, Hunt DF, Hamlin JL (1996). *Mimosine targets serine hydroxymethyltransferase.* J Biol Chem 271, 2548-56.

Lin S and Kowalski D (1997). *Functional equivalency and diversity of cis-acting elements among yeast replication origins.* Mol Cell Biol 17, 5473-84.

Little RD, Platt TH, Schildkraut CL (1993). *Initiation and termination of DNA replication in human rRNA genes.* Mol Cell Biol 13, 6600-13.

Loo S, Fox CA, Rine J, Kobayashi R, Stillman B, Bell S (1995). *The origin recognition complex in silencing, cell cycle progression, and DNA replication.* Mol Biol Cell 6, 741-56.

Lu L and Tower J (1997). *A transcriptional insulator element, the su(Hw) binding site, protects a chromosomal DNA replication origin from position effects.* Mol Cell Biol 17, 2202-6.

Lu ZH, Sittman DB, Romanowski P, Leno GH (1998). *Histone H1 reduces the frequency of initiation in Xenopus egg extract by limiting the assembly of prereplication complexes on sperm chromatin.* Mol Biol Cell 9, 1163-76.

Madine MA, Khoo CY, Mills AD, Mushal C, Laskey RA (1995). *The nuclear envelope prevents reinitiation of replication by regulating the binding of MCM3 to chromatin in Xenopus egg extracts.* Curr Biol 5, 1270-9.

Mahbubani HM, Chong JP, Chevalier S, Thommes P, Blow JJ (1997). *Cell cycle regulation of the replication licensing system: involvement of a Cdk-dependent inhibitor.* J Cell Biol 136, 125-35.

Marahrens Y and Stillman B (1992). *A yeast chromosomal origin of DNA replication defined by multiple functional elements.* Science 255, 817-23.

McGarry TJ, Kirschner MW (1998). *Geminin, an inhibitor of DNA replication, is degraded during mitosis.* Cell 93, 1043-53.

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 11, 233-42.

Michaelson JS, Ermakova O, Birshtein BK, Ashouian N, Chevillard C, Riblet R, Schildkraut CL (1997). *Regulation of the replication of the murine immunoglobulin heavy chain gene locus: evaluation of the role of the 3' regulatory region.* Mol Cell Biol 17, 6167-74.

Moir D and Botstein D (1982). *Determination of the order of gene function in the yeast nuclear division pathway using cs and ts mutants.* Genetics 100, 565-77.

Nawotka KA and Huberman JA (1988). *Two-dimensional gel electrophoretic method for mapping DNA replicons.* Mol Cell Biol 8, 1408-13.

Newlon CS (1997). *Putting it all together: building a prereplicative complex.* Cell 91, 717-20.

Newlon CS and Theis JF (1993). *The structure and function of yeast ARS elements.* Curr Opin Genet Dev 3, 752-8.

Newlon CS, Collins I, Dershowitz A, Deshpande AM, Greenfeder SA, Ong LY, Theis JF (1993). *Analysis of replication origin function on chromosome III of Saccharomyces cerevisiae.* Cold Spring Harb Symp Quant Biol 58, 415-23.

Niller HH, Glaser G, Knuchel R, Wolf H (1995). *Nucleoprotein complexes and DNA 5'-ends at oriP of Epstein-Barr virus.* J Biol Chem 270, 12864-8.

Ohtani K, DeGregori J, Leone G, Herendeen DR, Kelly TJ, Nevins JR G (1996). *Expression of the HsOrc1 gene, a human ORC1 homolog, is regulated by cell proliferation via the E2F transcription factor.* Mol Cell Biol 16, 6977-84.

Ohtani K, DeGregori J, Nevins JR (1995). *Regulation of the cyclin E gene by transcription factor E2F1.* Proc Natl Acad Sci U S A 92, 12146-50.

Okazaki R, Okazaki T, Sakabe K, Sugimoto K, Sugino A (1968). *Mechanism of DNA chain growth. Possible discontinuity and unusual secondary structure of newly synthesized chains.* Proc Natl Acad Sci U S A 59, 598-605.

Owens JC, Detweiler CS, Li JJ (1997). *CDC45 is required in conjunction with CDC7/DBF4 to trigger the initiation of DNA replication.* Proc Natl Acad Sci 94, 12521-6.

Pak DT, Pflumm M, Chesnokov I, Huang DW, Kellum R, Marr J, Romanowski P, Botchan MR (1997).*Association of the origin recognition complex with heterochromatin and HP1 in higher eukaryotes.* Cell 91, 311-23.

Pasero P, Braguglia D, Gasser SM (1997). *ORC-dependent and origin-specific initiation of DNA replication at defined foci in isolated yeast nuclei.* Genes Dev 11, 1504-18.

Pelizon C, Diviacco S, Falaschi A, Giacca M (1996). *High-resolution mapping of the origin of DNA replication in the hamster dihydrofolate reductase gene domain by ompetitive PCR.* Mol Cell Biol 16, 5358-64.

Perkins G, Diffley JF (1998). *Nucleotide-dependent prereplicative complex assembly by Cdc6p, a homolog of eukaryotic and prokaryotic clamp-loaders.* Mol Cell 2, 23-32.

Piatti S, Bohm T, Cocker JH, Diffley JF, Nasmyth K (1996). *Activation of S-phase-promoting CDKs in late G1 defines a "point of no return" after which Cdc6 synthesis cannot promote DNA replication in yeast.* Genes Dev 10, 1516-31.

Piatti S, Lengauer C, Nasmyth K (1995). *Cdc6 is an unstable protein whose de novo synthesis in G1 is important for the onset of S phase and for preventing a 'reductional' anaphase in the budding yeast Saccharomyces cerevisiae.* EMBO J 14, 3788-99.

Qin XQ, Livingston DM, Kaelin WG Jr, Adams PD (1994). *Deregulated transcription factor E2F-1 expression leads to S-phase entry and p53-mediated apoptosis.* Proc Natl Acad Sci U S A 91, 10918-22.

Quintana DG, Hou ZH, Thome KC, Hendricks M., Saha P, Dutta A (1997). *Identification of HsORC4, a member of the human origin of replication recognition complex.* J. Biol. Chem. *272*, 28247-28251.

Quivy JP and Becker PB (1993). *An improved protocol for genomic sequencing and footprinting by ligation-mediated PCR.* Nucleic Acids Res *21*, 2779-2781.

Raghuraman MK, Brewer BJ, Fangman WL (1997). *Cell cycle-dependent establishment of a late replication program.* Science 276, 806-9.

Rao H, Marahrens Y, Stillman B (1994). *Functional conservation of multiple elements in yeast chromosomal replicators.* Mol Cell Biol 14, 7643-51.

Rao PN and Johnson RT (1970). *Mammalian cell fusion: studies on the regulation of DNA synthesis and mitosis.* Nature 225, 159-64.

Rein T, Natale DA, Gartner U, Niggemann M, DePamphilis ML, Zorbas H (1997). *Absence of an unusual "densely methylated island" at the hamster dhfr ori-beta.* J Biol Chem 272, 10021-9.

Rein T, Zorbas H, DePamphilis ML (1997). *Active mammalian replication origins are associated with a high-density cluster of mCpG dinucleotides.* Mol Cell Biol 17, 416-26.

Rein T, Kobayashi T, Malott M, Leffak M, DePamphilis ML (1999). *DNA methylation at mammalian replication origins.* J Biol Chem 274, 25792-800.

Ritzi M, Baack M, Musahl C, Romanowski P, Laskey RA, Knippers R (1998). *Human minichromosome maintenance proteins and human origin recognition complex 2 protein on chromatin.* J Biol Chem. 1998 273, 24543-9.

Rivin CJ and Fangman WL (1980). *Replication fork rate and origin activation during the S phase of Saccharomyces cerevisiae.* J Cell Biol 85, 108-15.

Romanowski P, Madine MA, Rowles A, Blow JJ, Laskey RA (1996). *The Xenopus origin recognition complex is essential for DNA replication and MCM binding to chromatin.* Curr Biol 6, 1416-25.

Roufa DJ and Marchionni MA (1982). *Nucleosome segregation at a defined mammalian chromosomal site.* Proc Natl Acad Sci U S A 79, 1810-4.

Rowles A, Chong JP, Brown L, Howell M, Evan GI, Blow JJ (1996). *Interaction between the origin recognition complex and the replication licensing system in Xenopus.* Cell 87, 287-96.

Royzman I, Austin RJ, Bosco G, Bell SP, Orr-Weaver TL (1999). *ORC localization in Drosophila follicle cells and the effects of mutations in dE2F and dDP.* Genes Dev 13, 827-40.

Saha P, Chen J, Thome KC, Lawlis SJ, Hou ZH, Hendricks M, Parvin JD, Dutta A (1998). *Human CDC6/Cdc18 associates with Orc1 and cyclin-cdk and is selectively eliminated from the nucleus at the onset of S phase.* Mol Cell Biol. 18, 2758-67.

Saluz HP and Jost JP (1987). *A laboratory guide to genomes sequencing: the direct sequencing of native uncloned DNA.* Birkhauser, Boston, MA.

Santocanale C, Diffley JF (1996). *ORC- and Cdc6-dependent complexes at active and inactive chromosomal replication origins in Saccharomyces cerevisiae.* EMBO J 15, 6671-9.

Santocanale C and Diffley JF (1997). *Genomic footprinting of budding yeast replication origins during the cell cycle.* Methods Enzymol 283, 377-90.

Santocanale C and Diffley JF (1998). *A Mec1- and Rad53-dependent checkpoint controls late-firing origins of DNA replication.* Nature 395, 615-8.

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, 4340-51.

Shirahige K, Hori Y, Shiraishi K, Yamashita M, Takahashi K, Obuse C, Tsurimoto T, Yoshikawa H (1998). *Regulation of DNA-replication origins during cell-cycle progression.* Nature 395, 618-21.

Spann TP, Moir RD, Goldman AE, Stick R, Goldman RD (1997). *Disruption of nuclear lamin organization alters the distribution of replication factors and inhibits DNA synthesis.* J Cell Biol 136, 1201-12.

Sterner JM, Dew-Knight S, Musahl C, Kornbluth S, Horowitz JM (1998). *Negative regulation of DNA replication by the retinoblastoma protein is mediated by its association with MCM7.* Mol Cell Biol 18, 2748-57.

Stillman, B. (1993). *DNA replication.* Replicator renaissance. Nature *366*, 506-507.

Stinchcomb DT, Struhl K, Davis RW (1979). *Isolation and characterisation of a yeast chromosomal replicator.* Nature 282, 39-43.

Tada S, Chong JPJ, Mahbubani HM, Blow JJ (1999). *The RLF-B component of the replication licensing system is distinct from Cdc6 and functions after Cdc6 binds to chromatin.* Curr Biol 9, 211-4.

Taira T, Iguchi-Ariga SM, Ariga H (1994). *A novel DNA replication origin identified in the human heat shock protein 70 gene promoter.* Mol Cell Biol 4, 6386-97.

Tanaka T, Knapp D, Nasmyth K (1997). *Loading of an Mcm protein onto DNA replication origins is regulated by Cdc6p and CDKs.* Cell 90, 649-60.

Tanaka T and Nasmyth K (1998). *Association of RPA with chromosomal replication origins requires an Mcm protein, and is regulated by Rad53, and cyclin- and Dbf4-dependent kinases.* EMBO J 17, 5182-91.

Tasheva ES and Roufa DJ (1994a). *A mammalian origin of bidirectional DNA replication within the Chinese hamster RPS14 locus.* Mol Cell Biol 14, 5628-35.

Tasheva ES and Roufa DJ (1994b). *Densely methylated DNA islands in mammalian chromosomal replication origins.* Mol Cell Biol 14, 5636-44.

Thommes P, Kubota Y, Takisawa H, Blow JJ (1997). *The RLF-M component of the replication licensing system forms complexes containing all six MCM/P1 polypeptides.* EMBO J 16, 3312-9.

Todorov IT, Attaran A, Kearsey SE (1995). *BM28, a human member of the MCM2-3-5 family, is displaced from chromatin during DNA replication.* J Cell Biol 129, 1433-45.

Todorovic V, Falaschi A, Giacca M. *Replication origins in mammalian chromosomes: The happy few.* Frontiers in Biosciences. In press.

Toledo F, Lachages AM, Mayau V, Debatisse M (1999). *Initiation of DNA replication at the Chinese hamster origin oriGNAI3 relies on local sequences and/or chromatin structures, but not on transcription of the nearby GNAI3 gene.* Nucleic Acids Res 27, 1600-8.

Tribioli C, Biamonti G, Giacca M, Colonna M, Riva S, Falaschi A (1987). *Characterization of human DNA sequences synthesized at the onset of S-phase.* Nucleic Acids Res 15, 10211-32.

Tuan D, Solomon W, Li Q, London IM (1985). *The "beta-like-globin" gene domain in human erythroid cells.* Proc Natl Acad Sci U S A 82, 6384-8.

Tugal T, Zou-Yang XH., Gavin K, Pappin D, Canas B, Kobayashi R, Hunt T, Stillman B (1998). *The Orc4p and Orc5p subunits of the Xenopus and human origin recognition complex are related to Orc1p and Cdc6p.* J. Biol. Chem. *273*, 32421-29.

Udvadia AJ, Templeton DJ, Horowitz JM (1995). *Functional interactions between the retinoblastoma (Rb) protein and Sp-family members: superactivation by Rb requires amino acids necessary for growth suppression.* Proc Natl Acad Sci U S A 92, 3953-7.

Vassilev L and Johnson EM (1990). *An initiation zone of chromosomal DNA replication located upstream of the c-myc gene in proliferating HeLa cells.* Mol Cell Biol 10, 4899-904.

Vassilev L and Johnson EM (1990). *An initiation zone of chromosomal DNA replication located upstream of the c-myc gene in proliferating HeLa cells.* Mol Cell Biol 10, 4899-904.

Vassilev LT, Burhans WC, DePamphilis ML (1990). *Mapping an origin of DNA replication at a single-copy locus in exponentially proliferating mammalian cells.* Mol Cell Biol 10, 4685-9.

Venditti P, Costanzo G, Negri R, Camilloni G (1994). *ABF1 contributes to the chromatin organization of Saccharomyces cerevisiae ARS1 B-domain.* Biochim Biophys Acta 1219, 677-89.

Virta-Pearlman VJ, Gunaratne PH, Chinault AC (1993). *Analysis of a replication initiation sequence from the adenosine deaminase region of the mouse genome.* Mol Cell Biol 13, 5931-42.

Walker JE, Saraste M, Runswick MJ, Gay NJ (1982). *Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold.* EMBO J 1, 945-51.

Walter J, Sun L, Newport J (1998). *Regulated chromosomal DNA replication in the absence of a nucleus.* Mol Cell 1, 519-29.

Wang S, Dijkwel PA, Hamlin JL (1998). *Lagging-strand, early-labelling, and two-dimensional gel assays suggest multiple potential initiation sites in the Chinese hamster dihydrofolate reductase origin.* Mol Cell Biol 18, 39-50.

Wang TS (1991). *Eukaryotic DNA polymerases.* Annu Rev Biochem 60, 513-52.

Watson JD, Crick FH (1953). *Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid.* Nature, 4356.

Weinert T (1998). *DNA damage and checkpoint pathways: molecular anatomy and interactions with repair.* Cell 94, 555-8.

Weintraub SJ, Prater CA, Dean DC (1992). *Retinoblastoma protein switches the E2F site from positive to negative element.* Nature 358, 259-61.

Williams RS, Shohet RV, Stillman B (1997). *A human protein related to yeast Cdc6p.* Proc Natl Acad Sci U S A 94, 142-7.

Wolffe A (1998). *Chromatin structure and functions.* Third edition. Academic press, London.

Wu JR, and Gilbert DM (1996). *A distinct G1 step required to specify the Chinese hamster DHFR replication origin.* Science 271, 1270-2.

Wu JR and Gilbert DM (1997). *The replication origin decision point is a mitogen-independent, 2-aminopurine-sensitive, G1-phase event that precedes restriction point control.* Mol Cell Biol 17, 4312-21.

Wu JR, Yu G, Gilbert DM (1997). *Origin-specific initiation of mammalian nuclear DNA replication in a Xenopus cell-free system.* Methods 13, 313-24.

Yan H, Gibson S, Tye BK (1991). *Mcm2 and Mcm3, two proteins important for ARS activity, are related in structure and function.* Genes Dev 5, 944-57.

Yoon Y, Sanchez JA, Brun C, Huberman JA (1995). *Mapping of replication initiation sites in human ribosomal DNA by nascent-strand abundance analysis.* Mol Cell Biol 15, 2482-9.

Zhang C, Jenkins H, Goldberg MW, Allen TD, Hutchison CJ (1996). *Nuclear lamina and nuclear matrix organization in sperm pronuclei assembled in Xenopus egg extract.* J Cell Sci 109, 2275-86.

Zhao Y, Tsutsumi R, Yamaki M, Nagatsuka Y, Ejiri S, Tsutsumi K (1994). *Initiation zone of DNA replication at the aldolase B locus encompasses transcription promoter region.* Nucleic Acids Res 22, 5385-90.

Zou L and Stillman B (1998). *Formation of a preinitiation complex by S-phase cyclin CDK-dependent loading of Cdc45p onto chromatin.* Science 280, 593-6.

Zou L, Mitchell J, Stillman B (1997). *CDC45, a novel yeast gene that functions with the origin recognition complex and Mcm proteins in initiation of DNA replication.* Mol Cell Biol 17, 553-63.

Zwerschke W, Rottjakob HW, Kuntzel H (1994). *The Saccharomyces cerevisiae CDC6 gene is transcribed at late mitosis and encodes a ATP/GTPase controlling S phase initiation.* J Biol Chem. 269, 23351-6.