A STRATEGY TO ISOLATE, IN VIVO, A EUKARYOTIC REPLICATION ORIGIN FOR THE ANALYSIS OF ITS ASSOCIATED PROTEINS

Francesca Romana Mariotti

PhD

The University of Edinburgh

2010

DECLARATION

I declare that this thesis was composed by myself and the research presented is my own unless otherwise stated.

Francesca Romana Mariotti

ACKNOWLEDGMENTS

- to *my parents* for leaving me free to decide, for always believing in me and supporting me in any choice I made. Thank you for your love.
- to *Maria* for accepting me in her lab and giving me this opportunity, which helped me grow as a scientist. Thank you also for understanding me when I needed to go back to Rome.
- to *Lea* for being my supervisor during the last year of my PhD. Thank you for for leaving me to decide about the experiments and for your encouragements when I was down.
- to *Fiona* for EVERYTHING. Thank you for always being beside me when I was happy and when I was sad. You always helped me when I had trouble with the projects and most importantly you gave me the strength to go on when I was crying and disappointed. You made me appreciate Edinburgh and Scotland. We shared really good times, such as dinners and concerts. I am really glad having you as a friend.
- to *Becky* for sharing this experience with me. We started our PhD together and we ended up being good friends sharing our happy and bad moments. We trusted on each other and this gave us the strength to face everything. I could have not desired a better "travel mate".
- to all the people in the Harrington lab for making me feel part of your group. Thanks to Elisa for the nights out and the fun. Thanks to Laura for your advice and for the "italian talks".
- to *Mone* for taking care of the "little Italian girl". Thank you for all the times we spent together. I had so much fun with you and Becky, going out, cooking and watching movies. Thank you for your "special parcels". One of the best thing of this experience is having found all of you as friends.
- to all my friends. In particular thank you to Elisa, Ste, Giulia and Paola. You have always been with me during all these years. Even though we have been far from each other for three years we are still really good friends. Most importantly thank you for listening to me even when I was a bit "crazy"!!!!
- to *Giorgio* for being my first mentor and for always trusting in me.

ABSTRACT

In all organisms the DNA content is precisely and efficiently replicated. DNA replication starts at specific regions called replication origins, where a pre-Replicative Complex (pre-RC) is assembled at the end of the G2/M phase, making chromatin competent for replication. At the transition between G1 and S phases the activation of two conserved protein kinases triggers the transition from pre-RC into the active form, the Replicative Complex (RC). Replication origins are not fired at the same time during S phase, thus we can define origins as "early" and "late" according to their time of activation. During the past years our knowledge about DNA replication has improved due to the identification of novel proteins. Despite a large body of work, it remains possible that additional factors remain unidentified. The discovery of new proteins by genetic means is difficult due to the presence of redundant mechanisms and to their requirement for cell survival. Thus the aim of my project is to develop an unbiased assay to study, in *Saccharomyces cerevisiae*, the proteome of a single eukaryotic replication origin and identify novel proteins involved in DNA replication. Because the chromosomal environment profoundly influences the behaviour of replication origins (such as timing, efficiency), it is therefore of interest to analyse the proteomics of replication origins in their endogenous context. Taking advantage of the pSR1 recombination system, I constructed two plasmids (pLT1/RsiteA and pLT2/RsiteB) that led to the integration of recombination sites into the yeast genome and the excision in vivo of the late replication origin ARS1413. These plasmids, allowing the excision of any flanking region, represent important genetic tools. The resulting episome has been isolated from the genome by density gradient purification in order to analyze by mass spectrometry the protein associated with the specific replication origin. This technique can be used to study the proteins involved in different steps of DNA replication. In fact cells arrested in G1 with α-factor will provide the proteome of the pre-RC while cells treated with HU will represent the proteome of late origins that are inhibited by the S-phase checkpoint.

ABBREVIATIONS

AE-BSF	4-(2-Aminoacetyl) benzenesulfonyl fluoride hydrochloride
APS	ammonium persulphate
bp	base pair
BSA	bovine serum albumin
ChIP	Chromatin Immuno Precipitation
Ci	curie
DNA	deoxyribonucleic acid
DNAse	deoxyribonuclease
dNTP	deoxynucleoside triphosphate
DTT	dithiothreitol
EDTA	ethylenediamine tetraacetic acid
FACS	fluorescence activated cell sorting
G418	geneticin
HU	hydroxyurea
IPTG	isopropyl-β-thiogalactopyranoside
kb	kilobase
kDa	kilo Dalton
l	liter
LB	Luria-Bertani
M	molar
m	milli 10^{-3}
min	minute

M	molarity
MW	molecular weight
n	nano 10^{-9}
OD	optical density
p	pico 10^{-12}
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PEG	polyethylene glycol
PK	proteinase K
PMSF	phenylmethylsulfonyl fluoride
RNaseA	ribonuclease A
rpm	revolutions per minute
SD	synthetic drop out
SDS	sodium dodecyl sulphate
sec	seconds
ssDNA	single strand DNA
TFA	trifluoroacetic acid
U	unit
V	volts
YNB	yeast nitrogen base
YPD	yeast peptone dextrose
X-Gal	5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside
μ	micro 10^{-6}
°C	degrees centigrade

TABLE OF CONTENTS

1 INTRODUCTION

1.1 GENOME REPLICATION

DNA replication is an essential process that enables the propagation of life. Genetic information contained in the genome must be accurately duplicated and segregated to provide each daughter cell with an intact complement of chromosomes.

DNA replication is initiated from multiple sites along the chromosomes at positions named replication origins. The initiation of DNA replication at multiple sites ensures that genome is completely duplicated during S-phase. At replication origins, the replication machinery is assembled and DNA unwound.

The cell cycle oscillates between two different states, which are regulated by the activity of conserved protein kinase complexes: the cyclin dependent kinase (CDK) and the Dbf4 dependent kinase Cdc7 (DDK) (Fig.1).

The first state occurs in G1 phase when CDK levels are low and a pre-Replicative Complex (pre-RC) is assembled at each replication origin (Fig.1 step 1). This process is referred to as "licensing" because it makes replication origins competent for replication initiation (reviewed in Takeda and Dutta, 2005).

At the transition between G1 and S-phase, an increase in CDK and DDK levels promotes transition of the pre-RC to its active form, the Replication Complex (RC), the key feature of the second chromatin state (Fig.1 step 2). This complex consists of DNA polymerases and of other factors required for the movement and the stability of replication forks.

Once the origin has fired, the replication forks proceed in opposite directions along the chromosomes while a post-replicative complex remains bound to the origin.

Considering the anti-parallel nature of the DNA and that DNA polymerase can only synthesize in the 5′-3′ direction, the two parental strands will be replicated by two different mechanisms: "leading" and "lagging" strand replication (Nick McElhinny, *et al*, 2008).

Once DNA replication is complete, the sister chromatids held together by the cohesin complex, are segregated between to daughter cells. Replication can begin again in a subsequent S phase (Fig.1 steps 4).

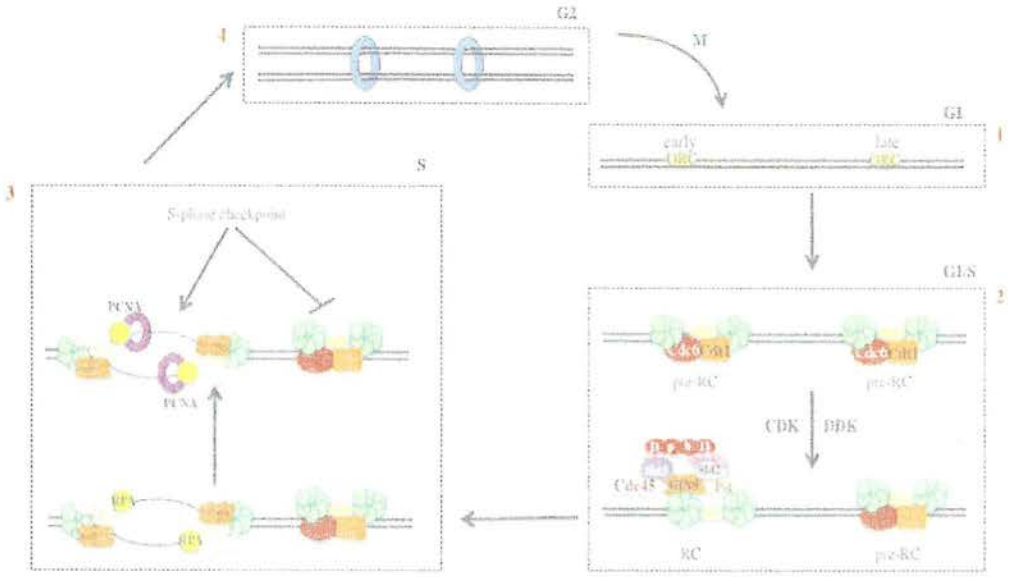


Fig.1 Chromosome replication (adapted from Diffley and Labib, 2002). The pre-RC complex is assembled at all early and late origins (step1). At the G1/S-phase transition, when CDK and DDK levels rise, pre-RC is converted to its active form, RC, and DNA is unwound (step 2). The two parental strands are replicated by different mechanisms due to the anti-parallel nature of DNA and the replication fork moves bidirectionally (step 3). Sites of DNA damage and stalled replication forks activate a checkpoint response that inhibits mitosis (step 3). Once replication is complete, sister chromatids are segregated to the daughter cells (steps 4).

The genome must be duplicated accurately and it is essential for cell survival that any region of the genome is replicated no more than once per cell cycle. Eukaryotic organisms have evolved various mechanisms to ensure their genomes are correctly replicated before each cell division. Limiting origin firing to only once per cell cycle is mainly achieved by restricting the licensing of replication origins to the G1 phase. In this way pre-RC assembly and initiation of DNA replication occur in mutually exclusive phases of the cell cycle (reviewed in Diffley, 2004). This regulation is controlled by CDK and DDK; both factors can promote activation of replication origins and inhibit the formation of new pre-RCs during S-phase. Origin firing is also

temporally regulated and origins can be classified as "early" and "late" depending on their time of activation during S-phase (Fig.1 step 1). It has been shown that the chromosomal context of an origin, rather than the nature of the origin itself, is one of the main determinants for the temporal program (Friedman, *et al*, 1996). In higher eukaryotes, a correlation between early and late origins and level of transcription has also been demonstrated (MacAlpine, *et al*, 2004; Woodfine, *et al*, 2004).

DNA replication is a potentially precarious process due to its complexity. Moreover, cells must also to cope with DNA damage or structural abnormalities, such as stalled replication forks, which can arise during DNA synthesis and lead to genomic instability if not properly repaired (Fig.1 step 3) (reviewed in Longhese, *et al*, 2003). To handle these lesions cells have evolved different mechanisms, called checkpoints, which arrest the cell cycle to allow for DNA to be repaired and correctly duplicated. The tight coordination between DNA replication and checkpoint activation is essential to guarantee that DNA replication occurs only once per cell cycle and that daughter cells receive an intact copy of the entire genome.

1.2 DNA REPLICATION IN BACTERIA AND ARCHAEA

The first model conceived to explain the events that lead to DNA replication is known as the "replicon model" (Jacob, *et al*, 1963). Studying the initiation of DNA replication in *Escherichia coli*, Jacob and colleagues proposed a model that resembled the transcriptional control of gene expression, whereby a *trans*-acting factor binds and regulates a *cis*-agent operator. This model defined the replicon as a genetic element that is replicated from a single replication origin (the *cis*-acting element), and recognized by a specific positive regulator protein (the *trans*-acting element). This control of initiation of DNA replication in *E. coli* implied the existence of specific features essential for genome duplication. Indeed, the replicon model requires the presence of specific sites where the DNA double helix is unwound as well as multiple factors, other than the DNA polymerases, which are involved in initiation of replication and synthesis of new DNA strands.

DNA replication in *E. coli* initiates from the single replication origin, named *oriC*, located within the circular chromosome. *OriC* of *E. coli* contains five 9 bp DnaA boxes, which are asymmetric consensus binding sites for the initiator protein DnaA (Fuller, *et al*, 1984), and a highly AT-rich region (Fig.2). Both these regions are critical elements for the activation of the replication origin. DnaA belongs to the AAA^{+}-ATPase family and its nucleotide-binding status influences the associations with the origin. DnaA-ATP binds to the five DnaA boxes as a monomer inducing a bend in the DNA structure (Schaper and Messer, 1995) and facilitating the melting of the adjacent AT-rich region (Fig.2).

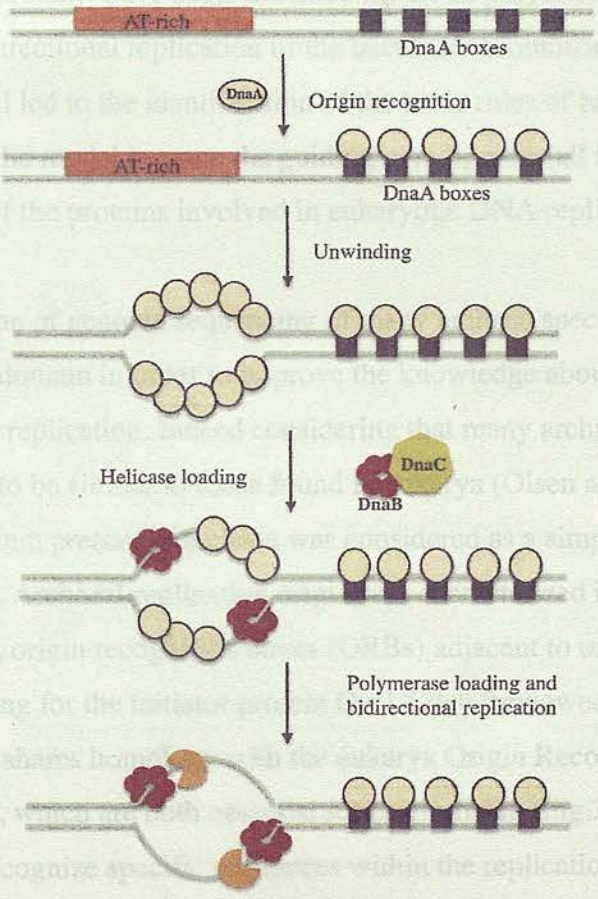


Fig.2 DNA replication in bacteria (adapted from Messer, 2002). Replication origins consist of several consensus sites, named DNA boxes, flanked by an AT-rich region. The initiator protein DnaA binds to and induces DNA unwinding of the replication origin. The replicative helicase, DnaB, is recruited to the replication origin through DnaC. Once DnaB and the DNA polymerase are loaded the DNA replication initiates and proceeds bidirectionally.

The formed single strand DNA (ssDNA) is then stabilized by the binding of other DnaA molecules. Once the replication origin has been unwound two DnaB hexamers are loaded on the opposite strand (Fang, *et al*, 1999) (Fig.2). However the recruitment of DnaB to *oriC* requires the action of another protein, called DnaC (Fig.2), which is a member of the AAA^+-family of ATPase. Indeed it has been shown that the role of ATP is crucial for DnaC action at *oriC* (Davey, *et al*, 2002). Davey and colleagues demonstrated that ATP controls the activity of the replicative helicase. Binding of DnaC to ATP increases the DnaC-ssDNA interaction and inhibits DnaB activity. However, after transfer of the helicase to *oriC*, DnaB and ssDNA trigger ATP hydrolysis relieving its inhibitory effect (Davey, *et al*, 2002). After the helicase proteins have been activated the DNA polymerases are recruited and initiate the bidirectional replication of the bacterial chromosome.

The replicon model led to the identification of the basic rules of bacterial DNA replication. Thus, the model became the guiding paradigm for all future characterizations of the proteins involved in eukaryotic DNA replication (reviewed in Mechali, 2001).

With the completion of genome sequencing of many archaea species attention has been given to this domain in order to improve the knowledge about the mechanisms that regulate DNA replication. Indeed considering that many archaeal replication proteins appeared to be similar to those found in eukarya (Olsen and Woese, 1997), the replication system present in archaea was considered as a simpler version of the eukaryotic system. Archaeal replication origins are characterized by the presence of an AT-rich region, origin recognition boxes (ORBs) adjacent to one or more copies of the gene encoding for the initiator protein Orc1/Cdc6 (reviewed in Barry and Bell, 2006). Orc1/Cdc6 shares homology with the eukarya Origin Recognition Complex (ORC) and Cdc6p, which are both essential for origin firing (Fig.3). The function of Orc1/Cdc6 is to recognize specific sequences within the replication origin (Robinson, *et al*, 2004) (Fig.3). As in eukarya the replicative helicase the hexameric Minichromosome Maintenance Complex (MCM), which assembles itself around single stranded DNA plasmid (Kelman, *et al*, 1999). Pivotal experiments carried out on archaeal MCM complexes contributed to unveiling the molecular mechanisms

that promote DNA unwinding (Kelman, *et al*, 1999). Kelman and colleagues demonstrated that the archaeal MCM complex contains helicase activity, similarly to the eukaryotic Mcm4/6/7 sub-complex (Fig.3). The recruitment of MCM to the replication origin requires the presence of additional proteins, such as Orc1/Cdc6 and Cdt1p in archaea and eukarya respectively (Fig.3). The GINS complex, as will be discussed later, has an essential function in the progression of DNA replication in eukaryotic cells (Takayama, *et al*, 2003) (Fig.3). It has been found that all archaea species encode two proteins that form a complex homologous to the eukaryotic GINS complex, (Marinsek *et al*, 2006) (Fig.3). The archaea GINS complex is required to couple the progression of the MCM helicase with the DNA synthesis on both DNA strands (Marinsek *et al*, 2006), a function that is conserved in eukaryotes.

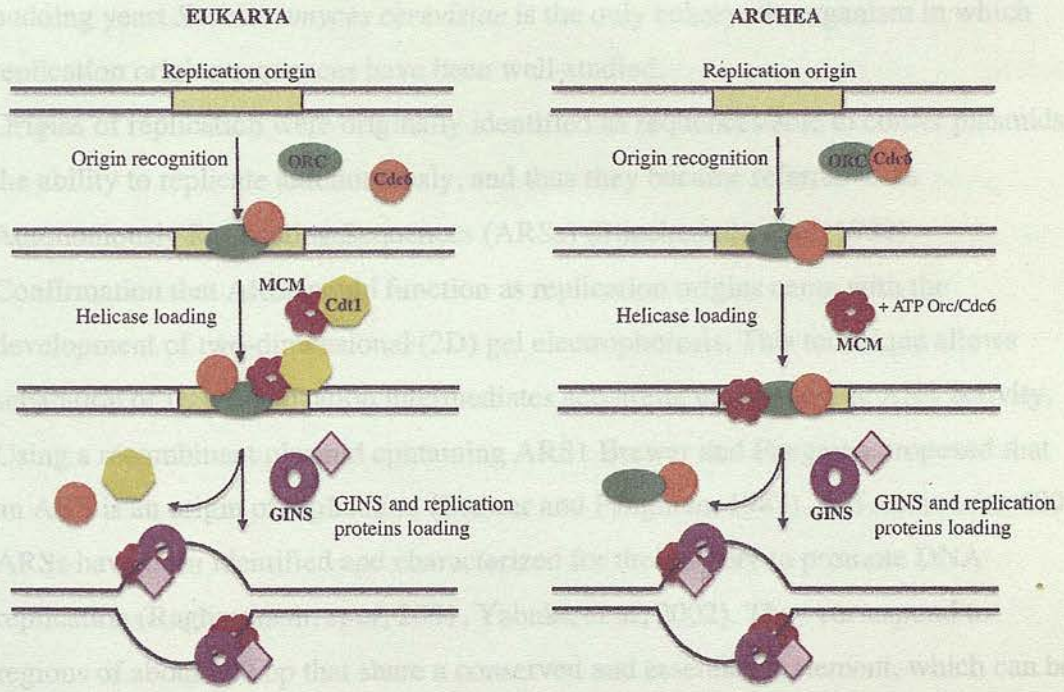


Fig.3 Models of DNA replication in Eukarya and Archaea (adapted from Kelman and Kelman, 2003). DNA replication initiates with the recognition of the replication origins by the initiator protein ORC and Orc/Cdc6 in Eukarya and Archea respectively. After origin recognition, the replicative helicase MCM, which is conserved in both domains of life, is loaded. Recruitment of additional factors, such as the GINS complex, is required to initiate and promote DNA replication.

Thus many principles and mechanisms of archaeal and bacterial DNA replication appear to be extendable to eukaryotes, supporting the idea that certain features of replication initiation might have been conserved during the evolution.

1.3 STRUCTURE OF EUKARYOTIC REPLICATION ORIGINS

Origins of replication act as initiation sites of DNA replication (Sclafani and Holzen, 2007). Studies in bacteria led to the identification and characterization of *oriC*, the single origin used by *E.coli* cells during genome replication. In contrast to bacterial cells in which DNA replication starts from a single origin, eukaryotes utilise hundreds of replication origins. However the presence of multiple initiation sites and the lack of consensus sequences have made the identification of replication origins in eukaryotes an arduous task (reviewed in Robinson and Bell, 2005). Indeed the budding yeast *Saccharomyces cerevisiae* is the only eukaryotic organism in which replication origins sequences have been well-studied.

Origins of replication were originally identified as sequences able to confer plasmids the ability to replicate autonomously, and thus they became referred to as Autonomously Replicating Sequences (ARSs) (Stinchcomb, *et al*, 1980). Confirmation that ARSs could function as replication origins came with the development of two-dimensional (2D) gel electrophoresis. This technique allows separation of DNA replication intermediates according to the specific ARS activity. Using a recombinant plasmid containing ARS1 Brewer and Fangman proposed that an ARS is an origin of replication (Brewer and Fangman, 1987). In *S. cerevisiae* 400 ARSs have been identified and characterized for their ability to promote DNA replication (Raghuraman, *et al*, 2001, Yabuki, *et al*, 2002). They correspond to regions of about 200 bp that share a conserved and essential A element, which can be flanked by multiple B domains (Marahrens and Stillman, 1992).

The A element consists of a 17 bp region that containes 11 bp matches with an AT-rich conserved sequence called the ACS (ARS consensus sequence A/TTTTAT/CA/GTTTA/T) (reviewed in Sclafani and Holzen, 2007). Different B elements have been identified (B1-B3) (Marahrens and Stillman, 1992) and, in

contrast to the A element, their sequences and number vary among replication origins. The A and B elements are binding sites for replication proteins, and thus are required for origin function (reviewed in Stillman, 2005). Indeed, a single mutation in the A element or simultaneous mutations in the three B elements abolished origin activity (Marahrens and Stillman, 1992). Footprinting analysis revealed that the pre-RC complex protects a region that consists of the A, B1 and B2 elements (Santocanale and Diffley, 1996). The A element is the main binding site of the Origin Recognition Complex (ORC) and the B1 element cooperates to form a bipartite site for ORC binding (Rao and Stillman, 1995). Origins that do not contain the B1 element can contain two ACSs that act in concert to form an ORC binding site (Bolon and Bielinsky 2006). The B2 element was originally proposed to promote the unwinding of DNA. Further studies have instead demonstrated that its sequence includes a partial inverted match to the ACS and that it is involved in pre-RC assembly, thereby facilitating the recruitment of the helicase complex (Wilmes and Bell, 2002). The B3 element is the binding site of the transcription factor Abf1p that acts as an enhancer of DNA replication (reviewed in Stillman, 2005). Abf1p also functions as a boundary for the correct positioning of nucleosomes surrounding replication origins (Venditti, *et al*, 1994). Even though yeast replication origins vary in structure and composition, they all contain conserved sequences. Thus, DNA replication in budding yeast seems to follow the replicon model a feature that is considered an exception amongst eukaryotes. Indeed in other eukaryotic organisms, origins of replication are located randomly throughout the genome and do not share conserved sequences. Therefore, replication in these organisms cannot be explained by the simple replicon model. However, some features, such as the high AT-rich content, have been identified in other eukaryotes. In *Schizosaccharomyces pombe,* replication origins are regions of at least 500 bp that do not have recognizable consensus elements but can be characterized by asymmetric AT-rich sequences (Segurado, *et al*, 2003). In mammals the human β-globin locus, located within Chr. 11, represents one of the best-characterized origins of replication. Its replication proceeds from a single initiation site that contains AT-rich regions and requires cooperation between at least two unique, non-redundant sites (Wang, *et al*, 2004).

Recently new insights on the structure of eukaryotic replication origins came from genome-wide studies on mammalian and *Drosophila melanogster* cells, which have shown a tight, positive correlation between gene density, transcription and origins of replication (Donaldson, 2005).

1.4 THE PRE-REPLICATIVE COMPLEX

Assembly of the pre-Replicative complex (pre-RC) on replication origins defines and renders origins competent for initiation of DNA replication. This event, which takes place between at the end of mitosis, involves the ordered loading of multiple factors including the Origin Recognition Complex (ORC), Cdc6p, Noc3p, Cdt1p and the Minichromosome Maintenance Complex (MCM). ORC is the first complex to interact with the ACS and the sequence-specific interaction is enhanced by ORC binding with Cdc6p. Together the ORC complex and Cdc6p are required to recruit the other components of the pre-RC complex (reviewed in Bell and Dutta, 2002).

1.5 THE ORIGIN RECOGNITION COMPLEX (ORC) AND CDC6

The ORC complex was originally purified as a multiprotein complex capable of recognizing replication origins and whose function correlated with origin activity (Bell and Stillman, 1992). The ORC complex is conserved among higher eukaryotes where it is essential for DNA replication (reviewed in Robinson and Bell, 2005). In DNA replication ORC is required to mark DNA replication origins and thus it has been proposed to act, in budding yeast, as the counterpart of the bacteria initiator protein DnaA. In budding yeast it is also involved in other cellular processes such as heterochromatin formation at the mating type loci (HM) and telomeres (Fox, *et al*, 1997).

The ORC complex consists of six subunits (Orc1-6) named according to the descending order of their relative mass (Chen, *et al*, 2008). Five of the ORC subunits are required for DNA binding, which occurs through the specific interaction of ORC with the A-rich strand of the ACS and with the adjacent B1 element (Rao and

Stillman,1995). Although the ORC complex is conserved in evolution the sequence-specificity of the DNA interaction was lost in eukaryotes as would be expected considering the absence of consensus elements that characterize replication origins in other eukaryotes (reviewed in Sclafani and Holzen, 2007). In humans and frogs the ORC complex binds preferentially to AT-rich regions, which could reflect the requirement of DNA origins to be unwound for initiation of DNA replication (reviewed in Sclafani and Holzen, 2007). In the fission yeast *S. pombe* the SpOrc4p, which is essential for ARS binding (Lee, *et al*, 2001) binds AT-rich DNA through its nine N-terminal AT-hook motifs. Moreover, ORC interacts preferentially with the T-rich strand and this association requires cooperation between different sequences within the replication origin (Kong and DePamphilis, 2002). Thus, even if all six ORC subunits are required for the correct assembly of the pre-RC, not all of them are necessary for ORC-DNA interaction. Indeed crosslinking studies in *S. cerevisiae* have demonstrated that the ORC complex recognizes replication origins through multiple, coordinated DNA interactions involving four of the six subunits. Orc6p is dispensable for sequence-specific DNA binding but its function is essential for the recruitment of other factors that take part in pre-RC assembly (Lee and Bell, 1997). DNA cross-linking experiments further elucidated the interaction between the ORC subunits and replication origins: Orc1p and Orc4p bind in the proximity of the A element while Orc2p and Orc3p bind near the B1 element. Based on recent experiments, a model has been proposed to explains how the Orc1-6 proteins interact within the complex and the ORC complex binds the A and B elements of yeast replication origins (Chen, *et al*, 2008). Electron microscopy analysis of each ORC subunit fused to maltose binding protein (MBP), in addition to immunoprecipitation experiments using the tagged proteins, led to the elucidation of yeast ORC architecture (Fig.4). Orc1p, Orc4p and Orc5p reside in the upper half of the complex, while Orc2p, Orc3p and Orc6p occupy the lower half. The Orc1 protein then interacts with Cdc6p, an essential factor for pre-RC assembly that was identified in a screen for suppressors of ORC mutations (Liang, *et al*, 1995). Thus, according to the ORC structure, the A element should bind to the upper half of the ORC complex where Orc1p and Orc4p are situated (Chen, *et al*, 2008).

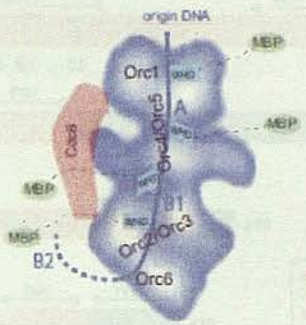


Fig.4 *S. cerevisiae* ORC acrchitecture (Chen, *et al*, 2008). Mass spectrometry and immunoprecipitation data have revealed the position of each ORC subunits within the complex. Orc1p, Orc4p and Orc5p localize in the upper half while Orc2p, Orc3p and Orc6p occupy the lower half of the structure. MBP: fused maltose binding protein; WHD: DNA winged-helix domains located in the C-terminus of the Orc1-5 subunits.

Orc6p is positioned at the bottom of the ORC complex where it can interact with DNA in the proximity of the B2 element as previously demonstrated (Lee and Bell, 1997).

Rearrangement of the ORC-DNA interaction might be caused by interaction with Cdc6p. Indeed, it is known from footprinting analysis that Cdc6p binding induces a conformational change in ORC resulting in a lengthening of the DNA contact surface (Speck, *et al*, 2005).

According to their structure Cdc6p and five of the six ORC subunits (Orc1-5) belong to the superfamily of AAA^{+} ATPase (ATPase Associated with various cellular Activities) which are characterized by their ability to bind ATP (reviewed in Duncker, *et al*, 2009). The main feature of the ATPase proteins is the presence of two distinct domains: AAA+ and WH (Fig.5). The former contains the Walker A (WA) and Walker B (WB) motifs, which are required for binding and hydrolysis of ATP. Two other motifs have been identified inside this domain: sensor I and sensor II. Interacting with the phosphate group of bound ATP they can discriminate between ATP and ADP and thus play a role in the subsequent conformational change (reviewed in Bell and Dutta, 2002).

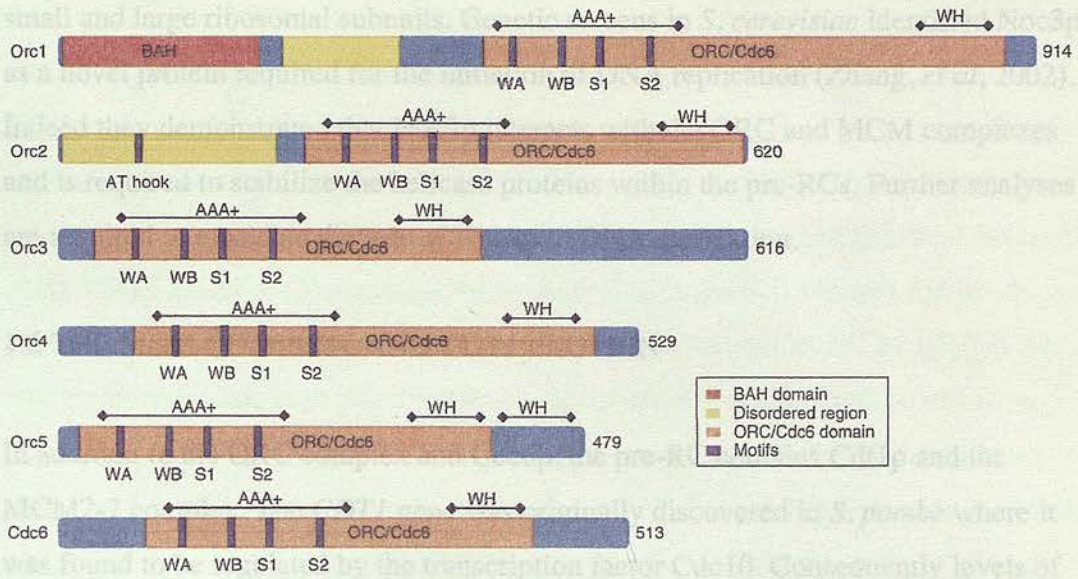


Fig.5 Comparison of *S. cerevisiae* Orc1-5 and Cdc6 proteins (Duncker, *et al*, 2009). The location of the WH and AAA+ domains in each protein have been indicated. The AAA+ domain contains Walker A (WA) and Walker B (WB) Sensor I (S1) and Sensor II (S2) motifs.

The second domain is structurally related to the Winged Helix Domain (WHD). This domain is found in various transcription factors and is involved in DNA binding. The Orc1 protein has an additional motif (Bromo-Adjacent Homology, BAH) in the N-terminal region, which has recently been demonstrated to be involved in the selection of replication origin (Muller, *et al*, 2010).

The interaction of ORC subunits with ATP is required at different times during pre-RC assembly. Indeed the interaction of ORC with replication origins occurs in an ATP-dependent manner (Bell and Stilmann, 1992). The specificity and stability of the ORC-DNA association is dependent on the interaction of ORC with the Cdc6 protein (Speck, *et al*, 2005). The ATPase activities of Orc1p and Cdc6p are required for subsequent steps in the pre-RC assembly. Indeed Cdc6p ATPase activity promotes the recruitment of the first MCM complex while hydrolysis of ATP by Orc1p is essential for the reiterative loading of the MCM complex (Bowers, *et al*, 2004). The recruitment of Cdc6p and the MCM complex is also dependent on the Noc3 protein (Zhang, *et al*, 2002). This protein is a member of the family of four

nucleolar complex-associated proteins which are required for the maturation of small and large ribosomal subunits. Genetic screens in *S. cerevisiae* identified Noc3p as a novel protein required for the initiation of DNA replication (Zhang, *et al*, 2002). Indeed they demonstrated that Noc3p interacts with the ORC and MCM complexes and is required to stabilize the helicase proteins within the pre-RCs. Further analyses are required to elucidate the role of Noc3p in DNA replication.

1.6 THE MCM COMPLEX AND CDT1 PROTEIN

In addition to the ORC complex and Cdc6p, the pre-RC contains Cdt1p and the MCM2-7 complex. The *CDT1* gene was originally discovered in *S. pombe* where it was found to be regulated by the transcription factor Cdc10. Consequently levels of Cdt1p peak in G1 and then decay with cell cycle progression (reviewed in Bell and Dutta, 2002). This protein is conserved among eukaryotes and is required during the initial steps of DNA replication, particularly in pre-RC assembly. Cdt1p interacts with subunits of the MCM complex and promotes their binding to replication origins (You and Masai, 2008). The MCM2-7 complex, which consists of six AAA^+-ATPases subunits, is likely the replicative helicase. It is recruited to replication origins as a component of the pre-RC and then activated with the onset of S-phase (reviewed in Cvetic and Walter, 2006).

The components of the helicase complex were identified in a genetic screen for factors involved in mini-chromosome maintenance in *S. cerevisiae*. A subset of the candidate mutations, which impaired plasmid stability, corresponded to six paralogue genes named *MCM2-7*; all of these genes are essential for viability (Schwacha and Bell, 2001). Recently, two other Mcm proteins were identified, Mcm8p and Mcm9p, that are present in human and *Drosophila melanogaster* (Maiorano *et al*, 2005a; Lutzman and Mechali, 2008). Studies in other organisms are shedding light on the roles of these proteins in DNA replication. Contrasting functions have been proposed for Mcm8p. Experiments on frog samples have shown that Mcm8p does not take part in pre-RC assembly but it is involved in fork movement, while human Mcm8p is

needed for the recruitment of Cdc6p and is important for origin licensing (Maiorano, *et al*, 2005a).

The Mcm9 protein is present only in vertebrates (Lutzman and Mechali, 2008). In *Xenopus Laevis* Mcm9p is involved in the assembly of pre-RCs and has been shown to bind chromatin before the MCM2-7 complex and in an ORC-dependent manner. Moreover it acts as an activating partner of Cdt1p thereby preventing excess loading of the Cdt1p inhibitor Geminin during the licensing reaction (Lutzman and Mechali, 2008). Finally another component of the human MCM-complex, MCM-BP (MCM binding protein), has been discovered by tandem affinity purification experiments (Sakwe, *et al*, 2007). MCM-BP interacts with three subunits of the MCM complex (4,6,7) and localizes to replication origins during G1. Thus, it has been proposed that MCM-BP can form an alternative MCM complex that contributes to the loading or stabilization of some MCM complexes on chromatin. The structure of the eukaryotic MCM complex has been deduced from studies on the archeal Mth-MCM complex of *Methanobacterium thermoautotrophicum*. As mentioned above, all the subunits of the eukaryotic helicase complex share a core domain, which contains the canonical catalytic AAA^+-ATPase regions, such as the Walker A and the Walker (reviewed in Maiorano, *et al*, 2006). The N-terminal region is involved in oligomerization and DNA binding while the C-terminal domain is less well characterized and thus its structure and function have still to be determined.

The Mcm2 and Mcm3 proteins contain a Nuclear Localization Sequence (NLS) and a Nuclear Export Sequence (NES) that are essential to regulate the intracellular localization of Mcm2-7p (Liku, *et al*, 2005). This is especially important in the budding and fission yeasts where the MCM2-7 complex is assembled in the cytoplasm. In other organisms the MCM proteins are constitutively localized in the nucleus and DNA binding is regulated by the cell cycle.

The six subunits of the Mcm2-7 complex form a double hexamer resembling a ring-like structure with a large central, positively charged channel, that can accommodate double-stranded (dsDNA) or single-stranded (ssDNA) DNA. According to the structure of the MCM2-7 complex, elucidated by biochemical data, the ATP moieties are located at the interface between each subunit. In this way one protein bind ATP

and the other provides the Arginine residue essential for the ATPase activity (Davey, *et al*, 2003). This organization is common among the AAA+ complexes, and has also been observed in the ORC complex (Bowers, *et al*, 2004). The ring shape of the MCM complex and its involvement in both pre-RC assembly and DNA replication suggested that it could function as the replicative helicase. However validation of the MCM2-7 complex as the replicative helicase has been an arduous task due to the contrasting results of *in vitro* assays. Initial experiments demonstrated that the Mcm4/6/7 sub-complexes had DNA helicase activity, but this activity was weak and not processive. Moreover, addition of the other three subunits (Mcm2/3/5) inhibited the helicase activity (Lee and Hurwitz, 2001). A model to explain these results has been proposed that the Mcm2/3/5 subunits represent the regulatory complex while the Mcm4/6/7 proteins are the catalytic domain providing the 3'-5' DNA helicase activity (Schwacha and Bell, 2001).

1.7 ASSEMBLY OF THE PRE-RC COMPLEX

The simplicity of the bacterial replicon model is lost in eukaryotic organisms. Compared to bacteria, which have a single chromosome, eukaryotes have more complex genomes with hundreds (as in yeast) or thousands (as in human) of replication origins that, except in budding yeast, lack consensus sequences. Indeed, *S. cerevisiae* represents the only eukaryotic organism in which initiation of DNA replication resembles the replicon model. The identification of conserved replication origins sequences has led to the elucidation of the mechanisms that promote assembly of the pre-RC. The first step in the sequential loading of proteins is the interaction of the ORC complex with replication origins through its association with Cdc6p. Once the ORC-Cdc6 complex becomes stably bound to DNA, the Cdt1 protein is loaded and the cooperation between these factors promotes recruitment of multiple MCM complexes.

1.8 ORC-CDC6 INTERACTION WITH REPLICATION ORIGINS

The ORC complex and Cdc6p are required for the successful initiation of DNA replication. Two-dimensional gel analysis, which can be used to follow replication origin activity, demonstrated that *orc2-1* and *orc5-1* temperature sensitive (ts) mutants were defective for formation of active replication origins (Liang, *et al*, 1995). A screen for genes that could rescue the replication defects led to the identification of the *CDC6* gene, and a synthetic lethality assay demonstrated the *in vivo* interaction between ORC and Cdc6p. The ORC complex is required during G1 and early S phases for recognition of ARS sites and initiation of DNA replication; it is dispensable for the elongation phase of DNA replication (Gibson, *et al*, 2006). The sequence-specific binding of ORC to replication origin requires both Cdc6p and ATP. Footprinting analyses have demonstrated that the pre-RC exhibits different profiles according to the phase of the cell cycle (Diffley, *et al*, 1995). In the G2 phase a post-RC (including ORC) is bound to replication origins (A and B1 elements), in agreement with the continuous presence of ORC at the origins throughout the cell cycle. In the G1 phase a different pattern was observed that indicated protection of the B2 element as well, due to the loading of other replication factors. Indeed it has been shown that binding of Cdc6p to the ORC complex can determine the extended footprint observed in G1 (Speck, *et al*, 2005).
The ability of Orc1p to hydrolyze ATP does not interfere with the selection of replication origins by ORC (Bowers, *et al*, 2004); indeed only ATP binding, stabilized by ORC-ARS association, regulates origin recognition (Klemm, *et al*, 1997). While replication origins can inhibit the ATPase activity of ORC, ssDNA can stimulate hydrolysis of ATP (Lee, *et al*, 2000). The ORC complex is able to bind ssDNA in a sequence-independent manner and this association causes an increase in ATPase activity and a conformational change in the ORC complex resulting in a more bent structure (Lee, *et al*, 2000). Thus, it has been proposed that the binding of ATP by Orc1p stimulates assembly of a higher complex while ATP hydrolysis promotes pre-RC disassembly (Klemm and Bell, 2001).

The ORC complex is not able to bind DNA per se in a stable and sequence-specific manner. Thus Cdc6p is required to increase the specificity of ORC for replication origin sequences. Moreover, Cdc6p and ORC1-6 must be bound to ATP in order to interact with one another (Speck, *et al*, 2005). The function of this interaction is to increase the specificity of the ORC complex for ARS sequences. Thus, Cdc6p stimulates the rate of ORC dissociation from non-specific DNA sequences in order to restrict ORC binding to functional origins of replication (Mizushima, *et al*, 2000). Indeed, in the presence of non-origin sequences or mutations in the A element, Cdc6p ATPase activity promotes disassembly of the ORC-DNA complex (Speck and Stillman, 2007). On the other hand Cdc6p stabilizes the association of ORC with replication origin seqeunces. Therefore, Cdc6p ATPase activity is essential for DNA replication because of its role in the selection of DNA sequences that promotes MCM loading and pre-RC assembly.

Electron micrograph data on the budding yeast have allowed reconstruction of the Cdc6p-ORC complex structure. These proteins forms a complex with a ring-like complex that might facilitate the subsequent loading of MCM complexes (Speck, *et al*, 2005).

1.9 LOADING OF THE REPLICATIVE HELICASE COMPLEX

Two opposing models have been proposed to explain the role of ORC in pre-RC assembly. In one interpretation, ORC acts as a "landing-pad" at origins of replication for pre-RC components (reviewed in Bell, 2002). The interaction of ORC complex with components of the pre-RC supports the hypothesis of a passive role for ORC in pre-RC assembly. In the other model, ORC not only nucleates the pre-RC components to the initiation sites but it also actively takes part in their loading (Bowers, *et al*, 2004). The ring-like structure of the ORC-Cdc6p complex and their ATPase activities support this model where ORC and Cdc6p might act as clamp loaders, a class of proteins that couple ATP hydrolysis with loading of circular protein rings around DNA (Bowers, *et al*, 2004). Orc1p, which can both bind and

hydrolyze ATP, must be in the ATP-bound state to facilitate the interaction between DNA and ORC (Klemm, *et al*, 1997).

Recent studies have also demonstrated that Orc1p and Cdc6p ATPase activities are essential for the reiterative recruitment of the Mcm2-7 helicases onto replication origins (Bowers, *et al*, 2004; Randell, *et al*, 2006). Cdt1p recruitment to replication origins is instead dependent on the smallest subunit of the ORC complex, Orc6p (Chen, *et al*, 2007), which is required to maintain the stability of the pre-RC. In cells depleted of Orc6p the association of the MCM complex with the chromatin is reduced (Da-Silva and Duncker, 2007) due to defects in Cdt1p recruitment onto replication origins (Chen *et al*, 2007). The Cdt1 protein associates transiently with the other components of the pre-RC, and these associations are controlled by the hydrolysis of ATP by Cdc6p (Randell, *et al*, 2006). However this ATP hydrolysis occurs only in the presence of ORC and DNA ensuring that Cdc6p does not act on Cdt1p proteins that have not been loaded onto the replication origins. The ORC-DNA dependent hydrolysis of ATP by Cdc6p is coupled with the dissociation of Cdt1p from origins and the consequent loading of the MCM complex (Fig.4) (Randell, *et al*, 2006). According to the "MCM paradox", at least two MCM complexes are recruited to each replication origin. Biochemical and electron microscopy analyses demonstrated that during pre-RC formation Mcm2-7 assembles into double hexamers (Evrin, *et al*, 2009); this reiterative loading is dependent on the ATPase activity of the Orc1 protein (Bowers, *et al*, 2004). Thus the Orc1p ATPase is dispensable for the first loaded MCM complex, which is instead dependent on Cdc6p ATPase activity. However ATP hydrolysis by Orc1p is required for the subsequent repetitive loading of Mcm2-7 at replication origins. The assembly of the pre-RC is a stepwise process coordinated by the different ATPase proteins involved (Fig. 6). After the sequence specific binding of ORC to replication origins, through its interaction with Cdc6p, the MCM-Cdt1p complex is recruited to the DNA. Hydrolysis of ATP by Cdc6p promotes the dissociation of Cdt1p and the closure of the MCM2-7 ring around the dsDNA. The subsequent ATP hydolysis by Orc1p catalyzes reiterative loading of other MCM complexes.

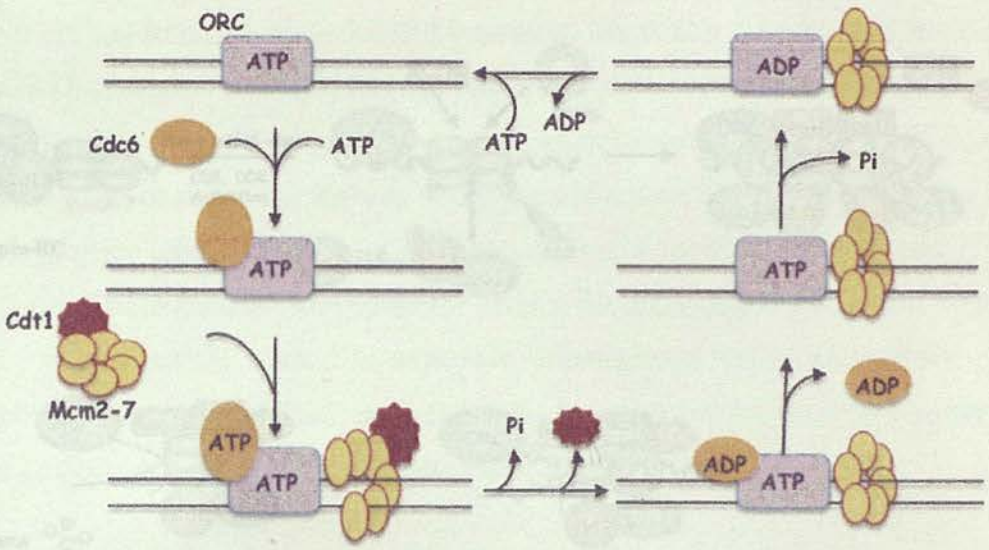


Fig.6 Model for pre-RC assembly (adapted from Randell, *et al*, 2006). ATP bound ORC interacts with replication origins. This association is stabilized by Cdc6-ATP. ORC and Cdc6p are required to recruit Cdt1p and Mcm2-7. The Cdc6p ATPase promotes the first round of Mcm2-7 loading resulting in the dissociation of Cdt1p. The hydrolysis of ATP destabilizes Cdc6p, which then dissociates. The ORC ATPase promotes reiterative loading of MCM complexes.

1.10 TRANSITION INTO S-PHASE

DNA replication is a multistep process initiated by the assembly of pre-RC complexes on replication origins between the end of mitosis and the G1 phase of the cell cycle. Pre-RC assembly is referred to as licensing because it renders the origins competent for initiation of DNA replication. However, only a subset of origins will fire during S-phase. At the G1/S transition, the pre-RC is converted into its active form, the Replication Complex (RC), through the recruitment of different factors. The Mcm10 protein is the first to interact with the pre-RC and, is required for the loading of Cdc45p (Fig.7) (Sawyer, *et al*, 2004). The GINS complex, which is required for the loading Cdc45p and DNA polymerase, interacts with the Cdc45 protein and the MCM complex. Together, these proteins, represent the core of the

replisome complex (Fig.7) (Gambus *et al*, 2006) and, are essential to promote the unwinding of DNA and the subsequent loading of DNA polymerases.

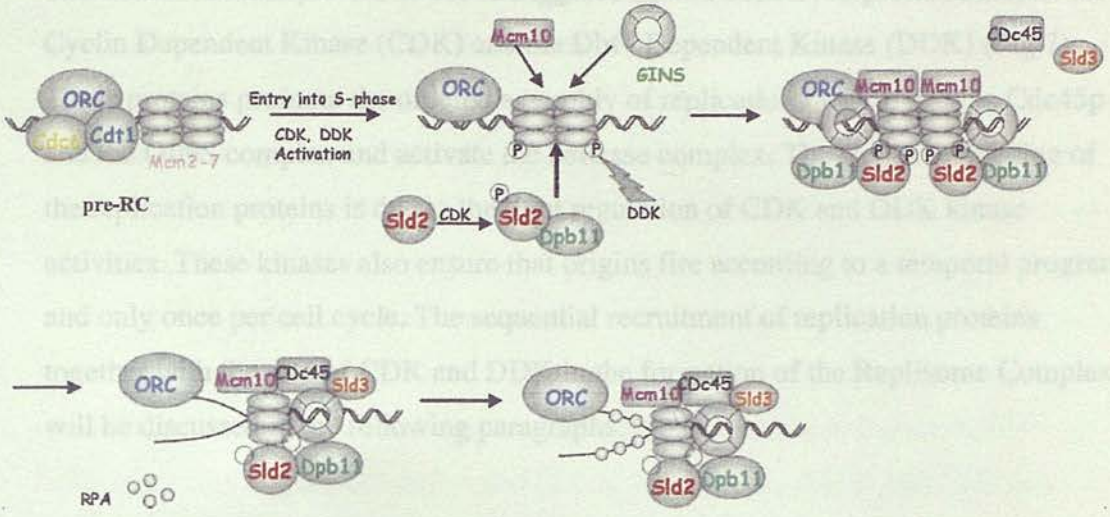


Fig.7 Transition from pre-RC to RC and initiation of DNA replication (adapted from Takeda and Dutta, 2005). Activation of the pre-RC is triggered by CDK and DDK activities. Phosphorylation of their targets promotes loading of various replication factors. The core of the replisome complex consists of Mcm2-7, Cdc45p and GINS. Their tight association is essential to maintain fork stability. As replication forks move in opposite directions, the post-RC remains bound to the fired replication origin

Initiation of DNA replication is also dependent on Dpb11p, Sld2p and Sld3p. Dpb11p and Sld3p are required for the loading of Cdc45p and DNA polymerases (reviewed in Takeda and Dutta, 2005). The following recruitment of both Replication Protein A (RPA) and PCNA proteins stabilizes the ssDNA formed during DNA unwinding and enhances the processivity of the DNA polymerases, respectively. The assembled replisome complex moves with the replication fork while a post-Replicative Complex, constituted by the ORC complex, remains on the origin (Fig.7) (Aparicio, *et al*, 1997). The interactions between the different proteins

that form the replication complex maintain the stable association of DNA polymerases with DNA and prevent the collapse of the fork upon activation of the S-phase checkpoint.

The transition from pre-RC to RC is triggered by two conserved protein kinases: the Cyclin Dependent Kinase (CDK) and the Dbf4 Dependent Kinase (DDK) (Fig.7). These proteins promote the ordered assembly of replication factors, such as Cdc45p and the GINS complex and activate the helicase complex. The step-wise loading of the replication proteins is due to the tight regulation of CDK and DDK kinase activities. These kinases also ensure that origins fire according to a temporal program and only once per cell cycle. The sequential recruitment of replication proteins together with the role of CDK and DDK in the formation of the Replisome Complex will be discussed in the following paragraphs.

1.11 MCM10 AND CDC45 PROTEINS

Mcm10p is the first initiation factor recruited to the pre-RC. Mcm10p remains at replication origins until they fire and then is displaced and travels with the replication fork (Ricke and Bielinsky, 2004).

The *MCM10* gene was originally discovered in a screen to isolate mutants that were defective in DNA replication initiation (Merchant, *et al*, 1997). Merchant and colleagues observed cell cycle arrest and reduced origin activation in the *mcm10-1* mutant. By two-hybrid assay it was shown that the product of the *MCM10* gene physically interacted with multiple subunits of the MCM complex. These data suggested that Mcm10p was involved in the initiation of DNA replication. Mcm10p and the Mcm7 subunit interact, and this association stabilizes the MCM complex (Homesley, *et al*, 2000). This interaction also occurs in fission yeast, in which Mcm10p promotes the phosphorylation and thus the activation of Mcm2-7 by DDK (Lee, *et al*, 2003).

Mcm10p colocalizes with Orc1p (Kawasaki, *et al*, 2000) and interacts genetically with elongation factors, such as DNA polymerases subunits (Kawasaki *et al*, 2000). The association of Mcm10p with the ORC and MCM complexes places Mcm10p in

an ideal position to ensure that initiation of DNA replication occurs next to the ORC binding site.

Further studies revealed that Mcm10p is required not only for replication initiation but also for progression of the replication fork. Immunoprecipitation experiments demonstrated that the Mcm10 protein binds to DNA polymerase α (Pol α). This interaction functions to stabilize the Cdc17 subunit of the primase and its interaction with chromatin (Ricke and Bielinsky, 2004). According to a model proposed by Ricke and Bielinsky the Mcm10/pol α-primase complex is recruited to the replication fork upon DNA unwinding and loading of the RPA protein. Moreover, Kawasaki and colleagues demonstrated that *mcm10-1* mutants cannot complete S-phase following treatment with Hydroxyurea (HU), a compound that depletes the pool of dNTPs and causes activation of the S-phase checkpoint. Gregan and colleagues confirmed that inactivation of Cdc23p, the homolog of Mcm10p in *S. pombe*, impairs completion of DNA replication after HU treatment (Gregan, *et al*, 2003).

The Mcm10 protein physically interacts with Cdc45p, and this association is lost in *mcm10-1* mutants. Loading of Cdc45p by Mcm10p occurs in a cooperative way, and is essential for initiation of DNA replication (Sawyer, *et al*, 2004). The Cdc45 protein, highly conserved among the eukaryotes, is a crucial factor for the transition to replication initiation and the stability of the replication fork (Zou, *et al*, 1997). The *CDC45* gene was discovered to be cell cycle regulated and with stable mRNA detected specifically at the G1/S stage (Hardy, 1997). Temperature sensitive Cdc45p mutants were defective for S-phase entry and origin firing demonstrating that Cdc45p is essential for the efficient initiation of DNA synthesis at chromosomal replication origins (Zou, *et al*, 1997). The presence of Cdc45p at replication origin is also required for recruitment of two pivotal factors: RPAp and DNA polymerase α. ChIP experiments have demonstrated that in the absence of a functional Cdc45 protein, Pol α cannot be loaded onto the replication origins (Zou and Stillman, 2000). The association of RPA with the origin is mutually dependent on Cdc45p, and these factors are recruited at the same time during replication. Therefore, the interaction between these proteins, during the formation of the RC complex, is

critical for origin activation. Cdc45p has a major role in controlling origin activation and its association with replication origin correlates with origin firing (Zou and Stillman, 2000). Cdc45p loading is temporally regulated. It is recruited to early origins at the onset of S-phase and to late replication origins as DNA replication proceeds (Zou and Stillman, 2000). Recruitment of Cdc45p is dependent on CDK activity, ensuring Cdc45p will not be loaded before completion of pre-RC assembly (Zou and Stillman, 1998). Indeed, Cdc45p is not associated with origins of replication in α-factor arrested cells indicating that binding of Cdc45p to chromatin is temporally distinct from pre-RC assembly (Zou and Stillman, 1998). However, formation of the pre-RC seems to be a prerequisite for Cdc45p loading. The role of Cdc45p in controlling origin activation is particularly important in respect to the checkpoint surveillance, which determines the inhibition of late firing origins. ChIP experiments have shown that this control is due to a block in the recruitment of Cdc45p onto late replication origins (Aparicio, *et al*, 1999). Moreover, its presence at the replication fork is essential to ensure completion of DNA replication after stalling of the replication fork (Tercero, *et al*, 2000).

Cdc45p coimmunoprecipitates with subunits of the MCM complex (Zou and Stillman, 1998). These proteins associate with origins at the early stage of DNA replication and with non origin regions as the S-phase proceeds, suggesting that they move with the replication fork (Aparicio, *et al*, 1997). These data revealed a role for Cdc45p during the elongation step of DNA replication, which has been confirmed by experiments demonstrating an essential role of Cdc45p in replication fork progression (Tercero, *et al*, 2000). Thus, Cdc45p is required during initiation of DNA replication for the loading of DNA polymerase, and during elongation for maintaining the stable association between DNA polymerase and the other replication factors.

1.12 DPB11, SLD2 AND SLD3 REPLICATION FACTORS

Dpb11p, another crucial factor for assembly of the RC, was discovered in a screen to find interactors of DNA polymerase ε (Polε) (Araki, *et al*, 1995). Genetic crosses

between conditional mutants have revealed that Dpb11p interacts with Pol2p and Dpb2p, two subunits of the DNA polymerase holoenzyme. The *DPB11* gene is essential for cell growth in budding yeast and is involved in the S-phase checkpoint as demonstrated by sensitivity of the *dbp11-1* strain to HU (Araki, *et al*, 1995). Homologues of Dpb11p have been identified in fission yeast (Cut5/Rad4), *Xenopus* (Cut5/Mus101) and mammals (TopBP1) where they are all involved in DNA replication and checkpoint activation (reviewed in Takeda and Dutta, 2005). Dpb11p interacts with Polε and is required for its recruitment to replication origins. Indeed, they associate with ARS regions in a mutually dependent manner and loading of Dpb11p requires the presence of both the MCM complex and RPA protein (Masumoto, *et al*, 2000). Sld2p and Sld3p were indentified, in a screen for genes that were synthetically lethal with *dpb11-1*, as the principal targets of Dpb11p (Kamimura, *et al*, 1998). A putative Sld2p homologue in higher eukaryotes, RecQ4L, binds to *Xenopus* Cut5 (yDpb11p) (Matsuno, *et al*, 2006). An Sld3p counterpart has not yet been discovered. Sld2p contains 11 clustered phosphorylation sites bearing the canonical sequence CDK recognition sequence Ser/Thr-Pro, which is a binding site for the BRCT 3-4 domains of Dpb11p (Tak, *et al*, 2006). By two hybrid assay it was demonstrated that the C-terminal domain of Dpb11p, bearing BRCT 3-4, interacted with the Sld2p region containing the phosphorylation sites. The N-terminal domain of Dpb11p, which contains the BRCT 1-2 domains, binds Sld3p (Zegerman and Diffley, 2007). As for Sld2p, the interaction between Sld3p and Dpb11p is dependent on the prior CDK-mediated phosphorylation of two residues within the Sld3p phosphorylation cluster (Tanaka, *et al*, 2007). Both Sld2p and Sld3p are essential for the initiation of DNA replication, as demonstrated by the reduced frequency of initiation observed in *sld3-5* mutants (Kamimura *et al*, 2001). The role of Sld3p in DNA replication is not exclusively associated with its interaction with Dpb11p; Sld3p is also required for the loading of other proteins components of the RC complex. Sld3p physically interacts with, and recruits, Cdc45p s in a mutually dependent manner in both budding and fission yeasts (Kamimura *et al*, 2001; Yabuuchi, *et al*, 2006).

1.13 THE GINS COMPLEX

As described before, the *SLD2* and *SLD3* genes were discovered in a screen to identify genes that conferred synthetic lethality in combination with *dpb11-1*; the *SLD5* gene was identified in the same analysis (Kamimura, *et al*, 1998). Sld5p has recently been demonstrated to be a component of the GINS complex (Go, Ichi, Nii and San), which consists of the Sld5, Psf1, Psf2 and Psf3 proteins (Takayama, *et al*, 2003). GINS is a heterotetrameric complex that is highly conserved among eukaryotes. Each subunit interacts with two other components of the complex in the following order: Sld5-Psf1-Psf3-Psf2 (Chang, *et al*, 2007). The resulting GINS complex has a ring like structure with a central hole that can encircle DNA and function as a clamp loader for the other replication proteins (Chang, *et al*, 2007). The GINS complex is essential for both initiation of DNA replication and progression of the replication fork (reviewed in Labib and Gambus, 2007). Indeed one of the main functions of the complex is to maintain the associations between the proteins that take part in the assembly of the RC. The GINS complex binds to replication origins in S-phase, and *SLD5* and *PSF1* mutants are defective in the early steps of DNA replication demonstrating its role in the establishment of the replication fork (Takayama, *et al*, 2003). In humans, the stability of the GINS components depends on the formation of the full complex, which associates with the origins during S-phase (Aparicio, *et al*, 2009). The complex associates with multiple replication factors, as demonstrated by immunoprecipitation experiments, and thus it has been proposed to function as a scaffold for the assembly of replication complexes. Indeed, the Psf1 subunit associates with the Sld3 protein, and the interaction of Psf1p with replication origins is required for Cdc45p loading (Takayama, *et al*, 2003). The binding of Sld3p and Cdc45p to replication origins occurs prior to recruitment of the GINS complex (Kanemaki and Labib, 2006). However, while Sld3p dissociates from the Replication Complex, Cdc45p and GINS move with the replication fork. Similar interactions have been identified in fission yeast where the GINS complex and the Dpb11 protein require Sld3p for origin association while Sld3p can bind origins independently of other factors

(Yabuuchi, *et al*, 2006). Moreover, loading of Cdc45p requires the presence of GINS and Sld3p. Based on these data, a model has been proposed in which Sld3p is recruited to origins before other proteins and dictates the recruitment of the other replication factors (Yabuuchi, *et al*, 2006). Dpb11p forms a complex with Sld2p and DNA polymerase ε. Thus, it has been supposed that GINS mediates the interaction between Cdc45-Sld3 and Dpb11-DNA Polymerase ε complexes. In this scenario, GINS may act as a pivotal factor to promote the association of DNA polymerases with replication origins (Takayama, *et al*, 2003). However the GINS complex is also essential to maintain the stability of the Replisome Progression Complex (RPC) whose main components are GINS, Cdc45p and Mcm2-7 (Gambus *et al*, 2006). GINS is also required to establish and maintain the association of Mcm2-7 with the other components of the replisome such as Mrc1p, Ctf4p and the FACT complex (Gambus *et al*, 2006) (Fig.8).

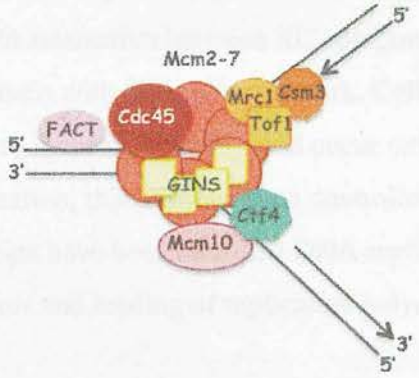


Fig.8 Model of RPC in *S. cerevisiae* (adapted from Labib and Gambus, 2007). GINS, Mcm2-7 and Cdc45p form the core of the RPC complex. Their association is required to maintain the stable association of the other component of the replisome machinery such as, Mrc1/Csm3/Tof1, Ctf4p and FACT.

The RPC complex, which is assembled during initiation and disassembled at the end of S-phase, was discovered in budding yeast through Mcm4p and Sld5p pull down experiments (Gambus, *et al*, 2006).

At the same time, the eukaryotic counterpart of the RPC complex was discovered. *Drosophila* Cdc45p pull down experiments identified a high molecular weight complex, referred to as CMG (Cdc45/Mcm2-7/GINS) (Moyer, *et al*, 2006). Binding of CMG to replication origins depends on the Psf3p subunit of the GINS complex (Ilves, *et al*, 2010). In human cells the formation of the CMG requires the presence of Mcm10p, Ctf4p and RecQL4p (homologue of ySld2p) (Im *et al*, 2009). The core of this complex is formed, as in budding yeast, by Cdc45p and the GINS and MCM complexes. CMG exerts 3'-5' helicase activity on ssDNA (Moyer, *et al*, 2006).
It has been hypothesized that the activation of the replicative helicases is due to the interaction of Mcm2-7 with Cdc45p and GINS. Indeed, the rate of ATP hydrolysis by Mcm2-7 is increased when the complex is associated with Cdc45p and GINS (Ilves, *et al*, 2010). Activation of the MCM helicase by Cdc45p and GINS is proposed to occur through a conformational change in the MCM subunits that increases ATPase activity (Ilves, *et al*, 2010). Thus, the transition from pre-RC to RC is based on the sequential loading of multiple replication factors among which Cdc45p and the MCM and GINS complexes represent the main components of the assembled replisome. The tight interaction between RC components is essential to form a stable complex that travels with the replication fork. Cell cycle control via CDK and DDK kinases ensures that RC assembly will occur only at the beginning of the S-phase after pre-RC formation; this allows a time-controlled initiation of DNA replication. Once all the proteins have been recruited DNA replication is triggered by the unwinding of double strands and loading of replicative polymerases.

1.14 CDK AND DDK FUNCTIONS IN RC ASSEMBLY

As mentioned above CDK and DDK activities are under cell cycle control in order to ensure that initiation of DNA replication only occurs during the S-phase.
The activation of DDK depends on the association of the kinase subunit Cdc7 with the cyclin Dbf4 (Jackson, *et al*, 1993). The DDK complex is activated in S-phase when levels of Dbf4p peak, and inhibited in other phases of the cell cycle when

Dbf4p is degraded by the Anaphase Promoting Complex (APC) (Ferreira, et al, 2000).

In contrast multiple CDK complexes can be assembled in different phases of the cell cycle. In budding yeast only one kinase, Cdc28, associates with different cyclin subunits while in mammals four classes of cyclins can interact with various kinase proteins. In *S. cerevisiae* Cdc28p associates with Cln1p-3p in G1, with Clb5p-6p in S-phase and with Clb1p-4p during mitosis (reviewed in Bloom and Cross, 2007). Differential regulation of cyclin gene transcription and protein posttranslational modifications ensure that each complex is formed in a specific phase of the cell cycle. CDK and DDK are regulated independently from each other by similar mechanisms. Both kinases are required for initiation of DNA replication. S-phase CDK complex (S-CDK) is activated in late G1 and is required (but not sufficient) to trigger the onset of S-phase. The transition to S-phase also requires the kinase activity of DDK (Nougarede, *et al*, 2000). Genetic and molecular evidence suggests that CDK and DDK interact to regulate DNA replication (Sclafani, *et al*, 2002) however they are activated independently from each other and they act through different mechanisms. While DDK acts at individual replication origins, CDK functions as a global regulator. For this reason CDK has been proposed to be the S-phase promotion factor (SPF) (Pasero, *et al*, 1999).

1.15 CDK TARGETS

During initiation of DNA replication the CDK complex promotes Cdc45p loading (Zou and Stillman, 1998), stimulates DDK phosphorylation, and in turn the activation of the helicase complex (Francis, *et al*, 2009). It also has a role in restricting assembly of pre-RC and RC at two separate points during cell cycle, ensuring that DNA synthesis occurs only once per cell cycle.

The minimal targets of CDK activity have been recently identified and they are the Sld3 and Sld2 proteins (Zegermanand Diffley, 2007; Tanaka, *et al*, 2007). Sld2p binds the BRCT3-4 domains in Dpb11p (Fig.9).

Moreover, assembly of the Dpb11-Sld2 complex is dependent on the prior CDK-dependent phosphorylation of a cluster of Ser/Thr residues (Tak, *et al*, 2006). Tak and colleagues demonstrated *in vivo* that CDK-dependent phosphorylation of the Thr84 residue is essential for the Sldp2-Dpb11p interaction. CDK phosphorylation of other canonical sites renders the Thr84 residue accessible to posttranslational modification. These multiple phosphorylation steps dictate a high threshold of CDK activity that ensures the association between Sld2p and Dpb11p will not occur prior to S-phase (Tak *et al*, 2006).

Sld3p is phosphorylated at the T600 residue by CDK (Zegerman and Diffley, 2007). This modification increases after release from α-factor arrest and promotes the interaction between Sld3p and the N-terminal domain of Dpb11p where the BRCT1-2 domains are located (Fig.9). Using *SD fusion* mutants in which a truncated Dpb11p was fused to a 3'-mutated Sld3p, Zegerman and Diffley showed that the forced interaction between these replication proteins can bypass the requirement for the Sld3p CDK sites. These data implied that the only essential function of CDK phosphorylation of both Sld3p and Dpb11p is to promote formation of the Sld3-Dpb11 complex. Moreover, by integrating the *SLD2-T84D* phospho-mimicking copy into *SD fusion* mutants, Zegerman and Diffley demonstrated that cells could bypass the requirement for S-CDK and initiate DNA replication. This finding identified Sld3p and Sld2p as the minimal targets for CDK activity.

Tanaka and colleagues confirmed these results (Tanaka, *et al*, 2007). In addition they isolated the *JET1* mutation of *CDC45* that affects the interaction between Dpb11p and Sld3p and allows Sld3p phosphorylation to be bypassed; Jet1p binds to Sld3p and promotes its association with Dpb11p independently of CDK phosphorylation.

Recently, CDK was found to promote the formation of the novel preloading complex (pre-LC) containing Sld2p, Dpb11p, GINS and Polε (Muramatsu, *et al*, 2010). Pre-LC assembly is dependent on CDK phosphorylation of Sld2p, but independent of DNA replication, DDK and association with replication origins. Thus it has been hypothesized that CDK controls the initiation of DNA replication by regulating pre-LC assembly.

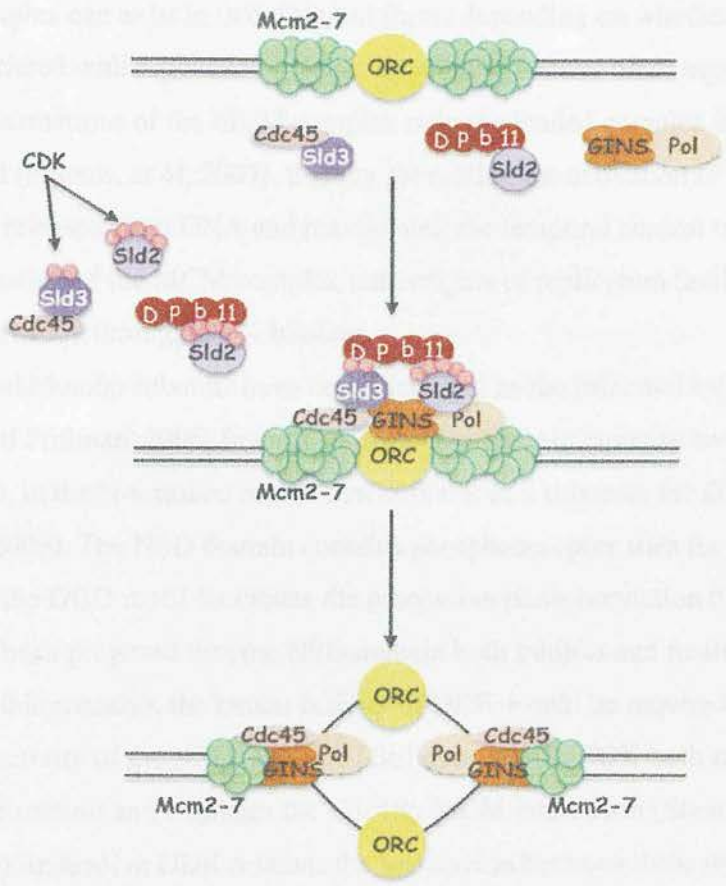


Fig.9 Model for CDK regulation of DNA replication (adapted from Tanaka, *et al*, 2007). Transition from pre-RC to RC is triggered by CDK and DDK activities. The essential targets of CDK essential are the Sld2 and Sld3 proteins. CDK-dependent Sld2p and Sld3p phosphorylation promotes their association with the Dpb11p BRCT3-4 and BRCT1-2 domains, respectively. These associations promote the initiation of DNA replication.

1.16 DDK TARGETS

DDK activity depends on the interaction of the Cdc7 kinase with the Dbf4 cyclin, whose protein levels peak at the beginning of S-phase. The binding of DDK to replication origins requires the presence of the ORC complex while Cdc6p and Mcm2-7 helicases are dispensable (Pasero, *et al*, 1999). Cdc45p requires the activity

of the DDK complex for stable binding with chromatin (Zou and Stillman, 2000) and for association with the MCM and GINS complexes. These interactions are essential for the activation of the helicase activity.

The MCM complex can exist in two different forms depending on whether they are loaded or associated with replication origins. Even though DDK binds equally well to the two conformations of the MCM complex only the loaded complex is phosphorylated (Francis, *et al*, 2009), thereby preventing the activation of complexes that have been released from DNA and maintaining the temporal control over DNA replication. Loading of the MCM complex onto origins of replication facilitates their consequent activation through DDK binding.

The Mcm4p and Mcm6p subunits have been identified as the principal targets of DDK (Sheu and Stillman, 2006; Francis, *et al*, 2009). Mcm4p contains two domains, NSD and DDD, in the N-terminal region that define it as a substrate for DDK (Sheu and Stillman, 2006). The NSD domain contains phosphoacceptor sites for DDK activity, while the DDD motif facilitates the processive phosphorylation of the NSD domain. It has been proposed that the NSD domain both inhibits and facilitates DNA replication. In this scenario, the kinase activity of DDK would be required to relieve the inhibitory activity of the NSD (Sheu and Stillman, 2010). DDK both modifies the Mcm4 helicase subunit and regulates the Cdc45p-MCM interaction (Sheu and Stillman, 2006). Indeed, in DDK mutants the association between these replication factors is not stable, demonstrating that Mcm4p phosphorylation is important for the presence of Cdc45p interaction at replication forks.

1.17 LOADING OF POLYMERASES AND INITIATION OF DNA REPLICATION

Synthesis of new DNA strands occurs through the coordinated actions of DNA polymerases. Their loading onto replication origins is the final step of DNA replication initiation. Two types of polymerases are required to duplicate the genome: the primase, which initiates DNA synthesis, and the replicative polymerases, which synthesize the genome (reviewed in Baker and Bell, 1998).

Three polymerases have been identified as components of the replication fork; the DNA polα-primase (Polα) and DNA polymerases ε (Polε) and δ (Polδ). These enzymes colocalize at replication foci and participate in DNA replication (Hiraga, *et al*, 2005) (Fig.10). The three DNA polymerases have different functions during elongation and their loading is dependent on replisome components previously assembled into the Replication Complex. The interaction of Dpb11p with Polε is essential for polymerase loading onto origins (Masumoto, *et al*, 2000). The GINS complex is also required for Polε loading. The C-terminal region of the Pol2 subunit of the polymerase interacts with the GINS complex and this binding facilitates DNA synthesis *in vitro* (Seki, *et al*, 2006) (Fig.10). Recruitment of the primase complex is facilitated by Cdc45p (Zou and Stillman, 2000) and its interactions with Mcm10p and the GINS complex. Loading of Polε and Polα occurs practically at the same time (Hiraga, *et al*, 2005), but while binding of Polα to replication proteins requires the presence of Polε and Dpb11p, the association of Polε does not depend on Polα (Masumoto, *et al*, 2000). This ordered loading ensures that the all the polymerases are loaded prior to initiation of DNA synthesis. The DNA polymerases perform distinct functions during genome duplication despite sharing a conserved catalytic core. The primase-polymerase α complex initiates DNA replication by synthesizing a short RNA primer of 12 nt that is then elongated by the polymerase activity of Polα, in order to create an RNA-DNA primer (reviewed in Baker and Bell, 1998). Polδ and Polε utilize this primer to synthesize new DNA strands. The DNA polymerases can only synthesize DNA in the 5'-3' direction; therefore, considering the opposite polarities of the double helix, the two strands are duplicated by two different systems. One strand, the leading strand, is duplicated in a continous manner, while the other one, the lagging strand, is copied in short (approximately 200 bp) Okazaki fragments (Pursell, *et al*, 2007; Nick McElhinny, *et al*, 2008). In addition to polymerase activity, Polε and Polδ have also a 3'-5' exonuclease activity that is required to proofread newly synthesized DNA strands and ensure high fidelity DNA replication. Only recently it has been demonstrated how the labor at the replication fork is shared between these two DNA polymerases. Taking advantage of amino acid mutations that impair the proofreading activity, Pursell and colleagues

found that Polε participates in the leading strand DNA synthesis (Pursell, *et al*, 2007). The same technique was applied to unveil which polymerase was responsible for replication of the lagging strand. By measuring the rate of mismatch incorporation, it was shown that Polδ is required to elongate the Okazaki fragments. Thus, it is the DNA polymerase responsible for lagging strand synthesis (Nick McElhinny, *et al*, 2008). Maturation of elongated Okazaki fragments requires the activity of two additional proteins, Fen1p and Dna2p, which work in association with Polδ (Rossi and Bambara, 2008).

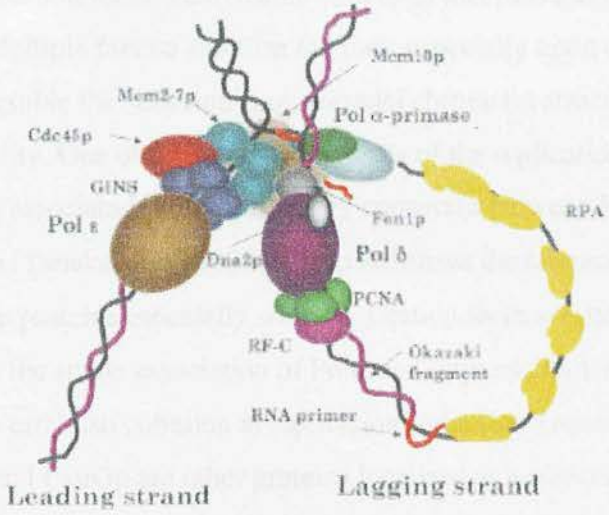


Fig.10 Model for eukaryotic replication forks (from Seki, *et al*, 2006). In this model Mcm2-7 encircles the leading strand adjacent to Polε, GINS and Cdc45. Leading strand synthesis is promoted by the cooperative activity of Polδ, Polα-primase, Fen1p, RFC, PCNA and Dna2p. Loading of RPA is required to stabilize ssDNA.

Processivity of the polymerase activity of both Polε and Polδ and the switch between Polα and the replicative polymerases on the leading strand, are dependent on loading of the PCNA protein onto the RNA-DNA hybrid (reviewed in Moldovan, *et al*,

2007). PCNA is a sliding clamp protein, which has a ring structure that can encircle the DNA. Loading of PCNA onto the DNA is ATP-dependent and requires the presence of clamp loader proteins (Gomes and Burgers, 2001). In eukaryotic cells RF-C, consisting of five AAA^+-ATPases subunits (Rfc1-5), is the clamp loader responsible for loading of PCNA onto the DNA.

The interaction between the Pol32 subunit of Polδ and PCNA is essential to ensure that the switch between Polα and Polδ is continuous during the synthesis of the lagging strand (Johansson, *et al*, 2004).

1.18 OTHER COMPONENTS OF THE REPLICATION FORK

Elongating replication forks are dynamic structures that move along the chromosome. Multiple factors stabilize the fork, especially upon activation of the checkpoint, assemble the cohesion ring, remodel chromatin structure and maintain replisome integrity. One of the main constituents of the replication fork is Ctf4p, which is a Polα associated protein and a key connector between RPC components and the primase (Tanaka, *et al*, 2009). Ctf4p stabilizes the interaction of Polα and the other replication proteins especially when replication forks are stalled. Loss of Ctf4p function affects the stable association of Polα and impairs DNA replication. Ctf4p is also required to establish cohesion at replication forks (Lengronne, *et al*, 2006).

Mrc1p, Tof1p and Csm3p are other proteins localized at replication forks and are required to stabilize stalled replication fork in the presence of the S-phase checkpoint or replication fork barriers (Katou, et al, 2003; Calzada, et al, 2005).

Immunoprecipitation experiments have demonstrated that Mrc1p and Ctf4p interact with the Dia2 protein (Morohashi, *et al*, 2009), which is a component of the ubiquitin- proteasome system. The presence of Dia2p at the replication fork is required to control S-phase entry and progression of the replication machinery.

The FACT complex (Spt16p-Pob3p) is a transcriptional elongation factor involved in chromatin remodeling that might be involved in nucleosome depositioning during DNA replication.

1.19 REGULATION OF REPLICATION ORIGIN ACTIVATION

The utilization of multiple origins ensures the efficient replication of the relatively large eukaryotic genomes. Checkpoint mechanisms, which detect the presence of DNA damage, guarantee the fidelity of genome duplication. Moreover, it is crucial that DNA replication is restricted to only once per cell cycle.

1.20 THE TWO-STATE MODEL

Cell fusion experiments led to the first insight into how initiation is restricted to once per cell cycle (Rao and Johnson, 1970). With the aim of discovering the determinants involved in regulation of DNA replication, Rao and Johnson fused mammalian cells that had reached different phases of the cell cycle; the resulting heterokaryons contained two different nuclei and shared the same cytoplasm. The fusion of G1 and S-phase cells resulted in premature replication in the G1 nucleus. In contrast the fusion of the post replicative G2 nucleus with an S-phase nucleus failed to initiate DNA replication. Therefore, it was concluded that the S-phase nucleus contained diffusible factors required for initiation of genome duplication. Moreover whereas G1 nuclei could respond to these factors, G2 nuclei were refractory to further DNA replication. These data demonstrated that cells can differentiate between replicated and unreplicated DNA. Additional experiments in *Xenopus* oocytes confirmed these results and led to a model in which origins of replication oscillate between two non overlapping phases: licensing and initiation of DNA replication (reviewed in Blow and Dutta, 2005). Thus, control of DNA replication is accomplished by progression through different states of the cell cycle, which are closely related to CDK levels (reviewed in Diffley, 2004). The first phase occurs in G1 when all origins exhibit a pre-replicative pattern, referred to as "licensing" because it makes chromatin competent for replication. In the next phase, which spans S, G2 and M the CDK and CDK complexes trigger initiation of DNA replication (Nougarede, et al, 2000). Therefore, origin activation can occur only under certain conditions that also inhibit relicensing. The CDK-dependent temporal segregation between pre-RC assembly

and initiation of DNA replication ensures that replication occurs only once per cell cycle.

1.21 PREVENTION OF RE-REPLICATION

Restriction of DNA replication to once per cell cycle is achieved, in eukaryotic organisms, by inhibition of re-licensing within the same cell cycle through the regulation of CDK levels and the inhibition of new pre-RC assembly.

1.22 REGULATION OF CDK LEVELS

Cyclins are regulated both at the transcriptional level, through control of gene expression and at the protein level by the targeting of CDK complexes for ubiquitin-mediated proteolysis. These mechanisms provide for the periodic accumulation and degradation of cyclins. During metaphase M-CDK phosphorylates and activates the Anaphase Promoting Complex (APC) APC^{Cdc20}, which targets mitotic CDKs for degradation leading to exit from mitosis (Fig.11).

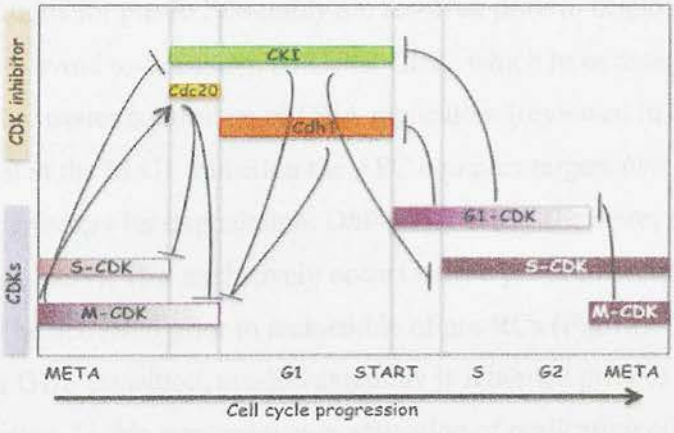


Fig.11 Cell cycle regulation of CDK levels (adapted from Diffley, 2004). Purple boxes indicate the different CDK complexes, according to the specific phase of the cell cycle. The APC complex, with its activators Cdh1p and Cdc20p, is involved in cyclin degradation. CKI also inhibits CDK activity. The tight regulation of these complexes ensures restriction of replication to once per cell cycle.

In addition, M-CDK proteolysis promotes the accumulation of the APC^{Cdh1} complex which contributes to cyclin and APC^{Cdc20} proteolysis (Kramer, et al, 2000). Collectively, APC^{Cdh1} and Cyclin depdendent Kinase Inhibitors (CKI), such as Sic1p, inactivate CDK activity during the G1 phase of cell cycle (Fig.11). During this unique window of time when CDK levels are low pre-RC assembly can occur.
The presence of growth signals promotes expression of the *CLN1-3* genes and accumulation of G1 cyclins, which inactivate APC^{Cdh1} and target CKI for degradation (Fig.11). Coordination of these processes allows for synthesis of S-CDK in S-phase, and M-CDK later in the cell cycle (reviewed in Bloom and Cross, 2007). Activation of S-CDK, which can occur only at the end of G1 when APC complex and Sic1p are absent, triggers the initiation of DNA replication and inhibits re-licensing of replication origins

1.23 INSULATING LICENSING

Licensing and initiation of DNA replication occur in two non-overlapping phases through the regulation of CDK levels.
However, the intermediate levels of cyclins present during transition periods could lead to overlap between licensing and replication initiation Therefore, it is essential that other conditions for pre-RC assembly are resolved prior to origin firing (and vice versa) in order to avoid re-licensing. APC and CDK, which have antagonistic functions, are the master controllers of DNA replication (reviewed in Porter, 2008). In budding yeast at the M/G1 transition the APC complex targets two essential S-phase promoting factors for degradation: Dbf4p and Clb5p (Ferreira, *et al*, 2000; Shirayama, *et al*, 1999). This exclusively occurs before pre-RC assembly ensuring origins will not be activated prior to reassemble of pre-RCs (Fig.12).
Similarly at the G1/S transition, pre-RC assembly is inhibited prior to the initiation of DNA replication. In this way premature activation of replication origins is avoided (Fig.12). Both these mechanisms allow for the insulation of pre-RC assembly from origin firing.

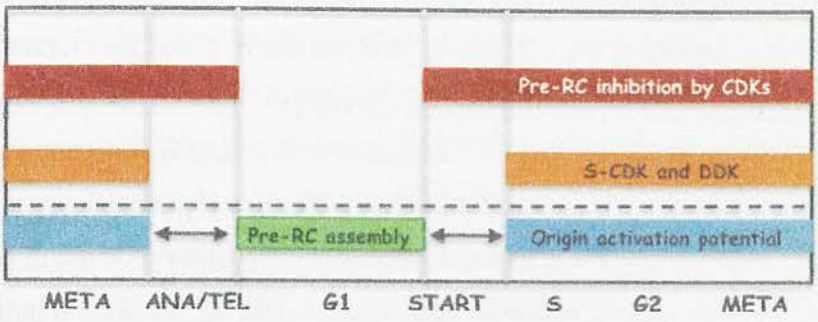


Fig.10 Insulation of licensing and replication initiation in *S. cerevisiae* (adapted from Diffley, 2004). Degradation of S-phase promoting factors prior to pre-RC assembly ensures that origins will not fire before licensing. Conversely inhibition of pre-RC before S-CDK activation avoids premature activation of replication origins.

1.24 STRATEGIES TO PREVENT RE-LICENSING

Considering that pre-RC assembly licenses replication origins, limitation of re-replication is achieved both through inactivation of existing pre-RCs and prevention of new pre-RC assembly until mitosis is complete. Strategies to prevent re-licensing can vary between organisms. Mechanisms to downregulate the formation of pre-RCs include CDK phosphorylation, targeting proteins for degradation and nuclear export.

1.25 RE-LICENSING PREVENTION IN *S. cerevisiae*

A high degree of redundancy in the mechanisms developed to prevent re-licensing has been observed in *S. cerevisiae*. The four strategies identified, are all based on CDK-dependent phosphorylation and inhibition of pre-RC components. Only when all pathways are down-regulated re-replication occurs (Nguyen, *et al*, 2001). Restriction of DNA replication occurs through regulation of the unstable Cdc6p protein, which is synthesized in each cell cycle accumulating during early G1. Three different mechanisms control Cdc6p activity: protein degradation, repression of

CDC6 transcription and inhibition of the interaction with ORC. Three different modes of Cdc6p proteolysis are responsible for the Cdc6p degradation in G1 and S-phase (Drury, *et al*, 2000). While the first "mode" has not been fully understood, both "mode 2" and "mode 3" depend on CDK and SCF^{Cdc4} (Skp, Cullin, F-box containing complex) (Fig.13). However, SCF^{Cdc4} mediated Cdc6p proteolysis is dependent on phosphorylation of Cdc6p by CDK. Two CDK consensus sites were identified in the N-terminal and C-terminal domains and both are required for Cdc6p degradation (Perkins, *et al*, 2001). These sites modulate the rate of SCF^{Cdc4}-dependent proteolysis during the cell cycle.

The mitotic cyclin Clb2 also regulates Cdc6p activity. Clb2p binds to the phosphorylated N-terminal domain of Cdc6p (Mimura, *et al*, 2004) thereby impairing the association between Cdc6p and the ORC complex and inhibiting assembly of new pre-RC (Fig.13). Levels of Cdc6p are also regulated at the transcriptional level. CDK phosphorylation of the transcriptional activator Swi5p prevents it from entering into the nucleus, thereby inhibiting expression of the *CDC6* gene. These three different mechanisms regulate Cdc6p activity. However genome-wide analysis demonstrated that re-replication can occur only when at least two pathways controlling pre-RC components (ORC, Cdc6p, MCM2-7, Cdt1p) are down-regulated (Green, *et al*, 2006).

Prevention of re-replication is also achieved by inhibition of ORC activity and, particularly, the Orc2 and Orc6 subunits. Through SDS-page analysis it was shown that these proteins, hypophosphorylated in G1, become phosphorylated after the START decision point and remain hyperphosphorylated until the subsequent G1 phase (Nguyen, *et al*, 2001). These modifications depend on CDK function and the S-phase cyclin complex (Fig.13) (Nguyen, *et al*, 2001). A role of Clb5-Cdc28 in ORC regulation has also been described, demonstrating that this cyclin is required to control Orc6p activity (reviewed in Arias and Walter, 2007).

Regulation of the MCM complex is accomplished by controlling the localization of helicase subunits during the cell cycle. Mcm4p is nuclear during G1 and becomes cytoplasmatic later in the cell cycle (Labib, *et al*, 1999). Both G1 and S-phase cyclins are responsible for nuclear export of Mcm4p. Further analyses showed that

the mitotic cyclin Clb2 modifies the nuclear localization sequence (NLS) and the nuclear export sequence (NES), thereby regulating intracellular compartmentalization (Liku, *et al*, 2005). The NLS and NES sequences of Mcm2p and Mcm3p form a regulatory transport domain, and CDK activity functions as a switch that controls import and export of the MCM complex (Liku, *et al*, 2005). Notably, only Mcm2-7 that is not bound to chromatin can be exported (Nguyen *et al*, 2000). The regulation of Cdt1p is tightly connected to that of Mcm2-7 and thus its nuclear localization changes during the cell cycle. Cdt1p is nuclear in G1 but completely cytoplasmatic later on. In the absence of Mcm2-7 Cdt1p cannot accumulate in the nucleus. Conversely, in the absence of Cdt1p nuclear import of the MCM complex is impaired (Tanaka and Diffley, 2002). Moreover Cdt1p export is dependent on CDK activity.

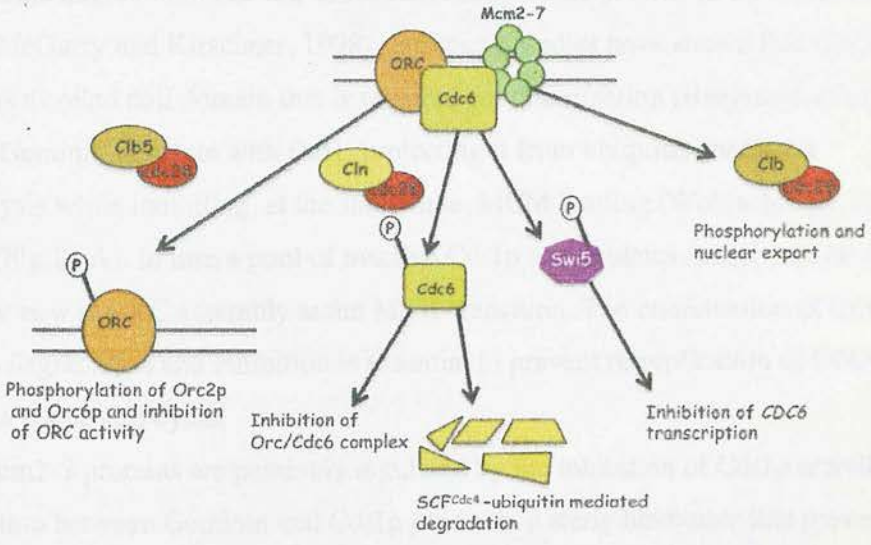


Fig.13 Re-licensing inhibition in budding yeast (adapted from Bloom and Cross, 2007). *S. cerevisiae* restricts initiation of DNA replication through CDK-dependent regulation of all the pre-RC components. Cdc6p is the main pre-RC target. Its phosphorylation inhibits ORC binding and acts as a target for proteasome-mediated degradation. CDK phosphorylation of the transcriptional activator Swi5p downregulates *CDC6* expression. CDK kinase activity on Mcm2-7 promotes its export from the nucleus; phosphorylation of Orc2p and Orc6p inhibits ORC activity.

1.26 RE-LICENSING PREVENTION IN METAZOA

In metazoa, restriction of DNA replication to once per cell cycle is achieved through regulation of pre-RC assembly.

Cdt1p is the main regulator of DNA replication in metazoa. Indeed overexpression of Cdt1p is associated with re-licensing and re-replication cells (Maiorano *et al*, 2005b). Cdt1p function is regulated by degradation of the protein and inhibition of the interaction with the MCM complex. While the Cdt1 protein levels accumulate at the end of mitosis and during G1, it is targeted for proteolysis at the G1/S-phase transition (Nishitani, *et al*, 2006). Two E3 ubiquitin ligase complexes, Cul4-Ddb1 and SCF, are responsible for Cdt1p proteolysis in mammals (Fig.14 A). However phosphorylation of Cdt1p by CDK is required for ubiqutin dependent degradation by the proteasome (Nishitani, *et al*, 2006). Cdt1p activity is also regulated by Geminin, an unstable protein that is the major trans-acting inhibitor of licensing (Li and Blow, 2005). APC^{Cdc20} – mediated ubiquitination of Geminin targets the protein for proteasome degradation and this allows assembly of the pre-RC in the following cell cycle (McGarry and Kirschner, 1998). Structural studies have shown that Geminin contains a coiled coil domain that is required for dimerization (Benjamin, *et al*, 2004). Geminin interacts with Cdt1, protecting it from ubiquitin-mediated proteolysis while inhibiting, at the same time, MCM loading (Wohlschlegel, *et al*, 2000) (Fig.14 A). In turn a pool of inactive Cdt1p accumulates, which can be readily used for new pre-RC assembly at the M/G1 transition. The combination of Cdt1p protein degradation and inhibition is essential to prevent re-replication of DNA within a single cell cycle.

The Mcm2-7 proteins are passively regulated by the inhibition of Cdt1p activity. The interaction between Geminin and Cdt1p produces a steric hindrance that prevents the association of the MCM complex with Cdt1p. This interaction impairs the loading of Mcm2-7 onto replication origins (reviewed in Machida, *et al*, 2005) (Fig.14 B).

ORC activity is regulated by cell cycle changes in the largest subunit Orc1p. In contrast to the levels of the other subunits that remain constant throughout the cell cycle, levels of Orc1p rise during G1 and decrease in the following phases (Mendez,

et al, 2002). Deletion of Orc1p is caused by ubiquitin-dependent proteolysis, which also requires CDK activity (Mendez, *et al*, 2002) (Fig.14 C). Cdc6p, which accumulates in the nucleus during G1, is regulated by protein degradation and nuclear export (Coverley, *et al*, 2000). The proteolysis of Cdc6p in S-phase nuclei is triggered by CyclinA-Cdk2 phosphorylation (Coverley, *et al*, 2000) (Fig.14 C). However, only free Cdc6p and not the chromatin bound fraction is targeted for CDK-dependent degradation. CDK also regulates Cdc6p localization. Phosphorylation in proximity of the NLS sequence determines the nuclear export during S-phase.

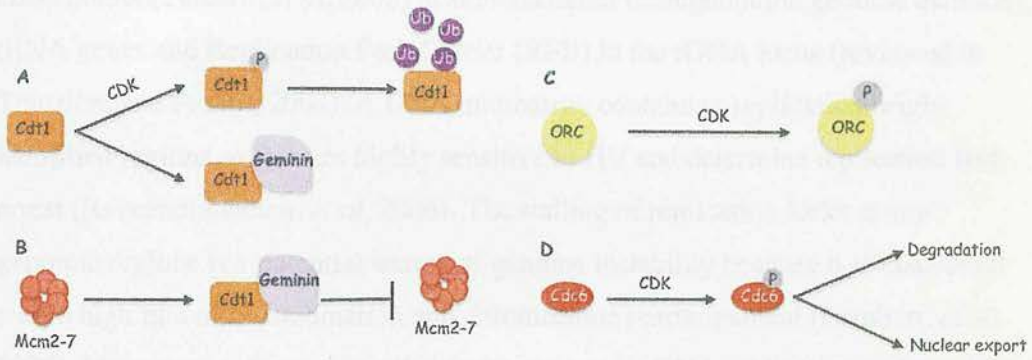


Fig. 14 Re-licensing inhibition in metazoa. A) CDK phosphorylation targets Cdt1p for ubiquitin dependent proteasome degradation. Cdt1p binding to Geminin stabilizes Cdt1p during G2 and M and protects it from degradation. **B)** The Geminin-Cdt1p interaction impair loading of Mcm2-7. **C)** Orc1p subunit is target for degradation upon CDK phosphorylation. **D)** Cdc6p is targeted for proteolysis and nuclear export; both these processes require prior CDK phosphorylation.

1.27 THE CHECKPOINT MECHANISMS

It has been underlined that the complete duplication of the genetic material, and accurate segregation of the chromosomes are essential for cell survival. Cells have evolved surveillance mechanisms, called checkpoints, to preserve genome integrity and ensure that each daughter cell receives an undamaged copy of the genome (reviewed in Willis and Rhind, 2009). These checkpoints are a complex network of signal-transduction pathways that can detect and respond to DNA damage caused by

genotoxic agents, as replication stress, or abnormal events during normal cell division. Several checkpoints respond to different events and are employed as an overall cellular response to regulate progression through cell cycle.

1.28 THE S-PHASE CHECKPOINT

With the intrinsic complexity of the DNA replication process comes the possibility of accumulation of genetic defects. Progression of the replication fork can be impaired by DNA lesions caused by exogenous and endogenous agents. The replication forks may also halt stochastically and independently of replication checkpoints (Tercero, *et al*, 2003) at sites scattered throughout the genome including tRNA genes and Replication Fork Barrier (RFB) in the rDNA locus (reviewed in Tourriere and Pasero, 2007). A DNA microarray containing replication origin identified regions, which are highly sensitive to HU and determine replication fork arrest (Raveendranathan, *et al*, 2006). The stalling of replication forks at any genomic regions is a potential source of genome instability because it is associated with a high rate of recombination and chromosome rearrangement (Lambert, *et al*, 2005). Moreover, collapse of stalled forks can impair DNA replication and result in un-replicated DNA regions. Therefore, all organisms have evolved a surveillance mechanism, called the S-phase checkpoint, which senses the presence of stalled forks and regions of bulky DNA and coordinates DNA replication with DNA repair. The hallmark of the S-phase checkpoint is the slowing of DNA replication and inhibition of late origin firing after damage is detected (Tercero and Diffley, 2001; Santocanale and Diffley, 1998). This surveillance mechanism is also required to maintain the integrity of replication forks, by preventing disassembly of the replisome in response to replication arrest (Lopes, *et al*, 2001) and to facilitate fork progression once the stress is removed (Tourriere and Pasero, 2007).

1.29 PROTEINS INVOLVED IN THE S-PHASE CHECKPOINT

The checkpoint response involves the activation of different protein kinases that sense replication stress and amplify the signal throughout the cell. In the simplest

model, a sensor protein activates a mediator factor upon detection of DNA damage or stalled forks. The mediator transduces the signal to effector kinases that amplify the signal and transduce a wide variety of cellular responses.

The keystone sensor kinases of the S-phase checkpoint are the Mec1 and Tel1 proteins, in *S. cerevisiae*, and ATR (ataxia-talangiectasia mutated-and Rad3-related) and ATM (ataxia-talangiectasia mutated) in mammals (Table.1).

Protein function	S.cerevisiae	Mammals
ATM/ATR-kinase	Mec1 and Tel1	ATM and ATR
ATR-interactiong proteins	Ddc2	ATRIP
Mediators	Rad9 and Mrc1	Claspin
Effector kinases	Rad53 and Chk1	Chk1 and Chk2
RFC-like proteins	Rad24 and RFC2-5	Rad17 and RFC2-5
PCNA-like proteins	Ddc1, Rad17 and Mec3	Rad9, Rad1 and Hus1
DSB recognition/ processing	MRX: Mre11, Rad50 and Xrs2	MSN: Mre11, Rad50 and Nbs1

Table 1. Proteins involved in S-phase checkpoint in yeast and mammals (adapted from Longhese, *et al*, 2003). Proteins are described in the text.

While Mec1p, and ATR, are activated in the presence of UV damage, double-strand breaks (DSBs) and stalled forks, Tel1p and ATM respond mainly to DSBs (Usui, *et al*, 2001).

ATR-related proteins require an accessory protein to be recruited to damage foci. Ddc2p has been demonstrated, by immunoprecipitation, to interact with Mec1p *in vivo* (Paciotti, *et al*, 2000) independently of other checkpoint proteins. The associations of Mec1p with Ddc2p in yeast, and ATR with ATRIP (ATR interacting protein) in metazoa are essential for the recruitment of these complexes to sites of DNA damage (Zou and Elledge, 2003). The RFC-like and PCNA-like complexes are also required for the activation of downstream targets. In *Xenopus* the PCNA-like complex, in association with TopBP1p (Dpb11p), directly activates ATR (Kumagai, *et al*, 2006). In yeast, the PCNA-like complex and Dpb11p function synergistically to activate Mec1p (Navadgi-Patil and Burgers, 2008) demonstrating that proteins involved in DNA replication can also function in the S-phase checkpoint. The sensor complexes Mec1/Ddc2 and ATR/ATRIP are necessary but not sufficient to activate the kinases involved in the checkpoint response and thus mediator proteins are required. In *S. cerevisiae* the main mediator kinases are Rad9 (BRCA1 in mammals) (Table. 1) and Mrc1 (Claspin in mammals). Mrc1p acts as a mediator for Rad53p activation only in response to DNA replication stress; however Mrc1p is also required in normal DNA replication (Osborn and Elledge, 2003). Rad53p (Chk2 in mammals) and Chk1p (Chk1) are the effector kinases that transduce signals to other cellular components. Rad53p seems to have a major role in the S-phase checkpoint response, and Chk1p is instead the main effector kinase in higher eukaryotes (reviewed in Segurado and Tercero, 2009). While Chk1p appears to be needed for the DNA damage response during the G2/M transition, a new role for this kinase in the response to replication stress has been proposed (Segurado and Diffley, 2008). Indeed Chk1p is required to stabilize replication forks in the absence of a functional Rad53 protein.

The involvement of Mec1p and Rad53p in the response to replication stress has been widely demonstrated (Lopes, *et al*, 2001; Tercero and Diffley, 2001). Their activation promotes two hallmark responses of the S-phase checkpoint: stabilization of replication forks and regulation of late origin firing (Santocanale and Diffley, 1998).

Tel1p, and the corresponding mammalian protein ATM respond primarily to DSBs. The response of Tel1p to DBSs requires the interaction with the MRX complex, which consists of Rad50/Mre11 and Xrs2p in budding yeast or Nsb1p in mammals (Table 1). The Tel1-MRX and Mec1p pathways act in parallel and both converge on the same Rad53p effector kinase.

1.30 ACTIVATION OF THE S-PHASE CHECKPOINT

Stalled replication forks are one of the main sources of genomic instability. The structure of stalled forks can vary according to the nature of the replication block. Treatment of cells with HU blocks DNA synthesis without interfering with replicative helicases. In constrast DNA lesions induced by the alkylating agent methyl methanesulfonate (MMS) or UV light cause the uncoupling of leading and lagging strand synthesis (reviewed in Tourriere and Pasero, 2007). Moreover cells must be able to discriminate between stalled forks at natural "fragile sites" and forks that arrest due to interference with the DNA replication process.

Different stalled fork structures can be differentiated by the amount of single-stranded DNA (ssDNA). Excess ssDNA is not generated at forks paused at natural sites, such as Replication Fork Barrier (RFB), (reviewed in Nedelcheva-Veleva, *et al*, 2006) and represents a hallmark of the S-phase checkpoint (Fig.15). In the presence of replication stress ssDNA is generated due to uncoupling of leading and lagging strand synthesis or a block in DNA polymerase activity. The long stretch of ssDNA is coated by RPA (Fig.15), a complex implicated in the S-phase checkpoint activation. Two RPA subunits, Rpa1p and Rpa2p, are phosphorylated by Mec1p in response to DNA damage or replication fork stalling (Brush, *et al*, 1996). RPA colocalizes with damaged foci and is essential in both budding yeast and mammals for recruitment of the Mec1/Ddc2 and ATR/ATRIP complexes respectively (Zou and Elledge, 2003) (Fig.15). Loading of multiple RPA complex on ssDNA stimulates ATRIP recruitment. Mec1p localizes exclusively at stalled forks, as demonstrated by ChIP experiment on cells treated with HU. Mrc1p, which associates with chromatin as cells begin DNA synthesis, contains different phosphorylation sites (SQ/TQ) that

are targets for Mec1p kinase activity. Phosphorylation of these residues is important for the recruitment of Mec1p, the stable binding of Mrc1p to replication fork (Naylor, *et al*, 2009) and to maintain cell viability during replication stress. Indeed cell cycle arrest is inhibited in *mrc1*^{AQ} mutants (deletion of Mec1 phosphorylation sites) treated with HU due to impaired Rad53p activation (Osborn and Elledge, 2003). These data also indicate that Mec1-dependent Mrc1p phosphorylation is required to activate Rad53p (Fig.15). Rad53p activation occurs only if functional replication forks have been established and correlates with entry into, and duration of S-phase (Tercero, *et al*, 2003). *Δclb5Δclb6* and *Δclb5* mutants, which have delayed entry into, and extended S-phase, respectively, when treated with MMS exhibit a delayed and prolonged Rad53p response.

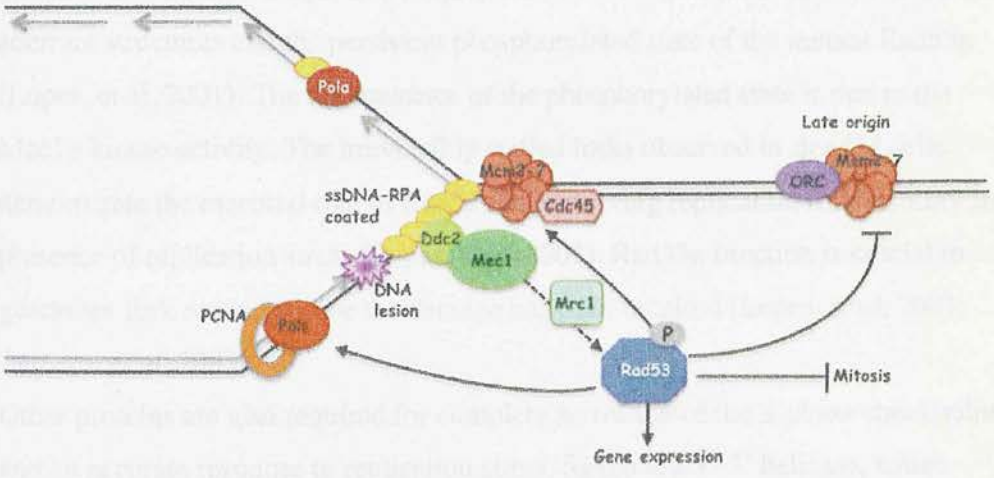


Fig. 15 Schematic representation of the S-phase checkpoint response (adapted from Segurado and Tercero, 2009). Stalling of replication forks in the presence of drugs or DNA lesions results in the exposure of ssDNA. Coating of ssDNA by RPA triggers the recruitment of the Mec/Ddc2 complex. Mrc1p, which is phosphorylated by Mec1p, promotes the activation of Rad53p. The effector kinase is then required to block mitosis, stabilize the replisome complex and inhibit late origin activation.

1.31 STABILIZATION OF STALLED REPLICATION FORKS

Mec1p and Rad53p responses to DSBs and stalled replication fork activating different pathways that lead to upregulation of dNTPs levels, expression of genes involved in replication stress or damaged DNA and mitotic arrest. Most importantly, checkpoint activation slows down the rate of replication fork progression (Tercero and Diffley, 2001), prevents the collapse of stalled replication forks (Lopes, *et al*, 2001) and inhibits firing of late origins (Santocanale and Diffley, 1998). Defects in both replication fork progression and activation of late origins are detected in *Δrad53* cells treated with HU (Lopes, *et al*, 2001). While WT cells are able to recover from HU treatment, indicating that stalled forks can be rescued, *Δrad53* cells are unable to complete DNA replication even when HU is removed. Taking advantage of a dominant negative version of Rad53p Lopes and colleagues demonstrated that the inability to properly resume DNA replication correlates with the accumulation of aberrant structures and the persistent phosphorylated state of the mutant Rad53p (Lopes, et al, 2001). The maintenance of the phosphorylated state is due to the Mec1p kinase activity. The irreversibly stalled forks observed in *Δrad53* cells demonstrate the essential role of Rad53p in preserving replication fork stability in the presence of replication stress (Lopes, et al, 2001). Rad53p function is crucial to guarantee fork recovery once the damage has been repaired (Lopes, *et al*, 2001; Tercero, *et al*, 2003).

Other proteins are also required for complete activation of the S-phase checkpoint and an accurate response to replication stress. Sgs1p is a 3'-5' helicase, which localizes to replication foci containing normal or stalled replication forks (Cobb, *et al*, 2003). Sgs1p is involved in normal replication progression and as well as S-phase checkpoint response. Sgs1p is required for Rad53p kinase activation as shown by the HU hypersensitivity of *Δsgs1* cells and is also essential to maintain the stability of DNA polymerases at stalled forks (Cobb, *et al*, 2003). ChIP analysis on Polα and Polε levels have demonstrated that, in the absence of Sgs1p, these proteins are depleted at replication forks. Cobb and colleagues have also shown that Sgs1p cooperates with Mec1p to stabilize the replisome, (Cobb, *et al*, 2005). Indeed a drop

in DNA polymerase levels can also be observed in *Δsgs1Δmec1-100* mutants due to displacement of DNA polymerases and RPA from damaged foci, rather than to Rad53p inactivation (Cobb, *et al*, 2005).

Chromatin remodeling complexes, such as Ino80p and Isw2p, are also involved in the S-phase checkpoint. Exposure of cells to DNA-damaging agents induces Mec1/Tel1 dependent phosphorylation of the Ies4 subunit of the Ino80 complex (Morrison, *et al*, 2007). Moreover, ChIP-on-ChIP analysis has shown that Ino80p associates not only with origin bound complexes but also with stalled forks, and its loss delays the completion of DNA replication after fork arrest (Shimada, *et al*, 2008). A role for Ino80p and Isw2p in DNA replication has been confirmed by whole genome replication profiling (Vincent, *et al*, 2008). Defects in replication fork progression in double mutants, in the presence of MMS, has led to the conclusion that both complexes are required for efficient replication of late replicating regions.

The sensor kinase Mrc1p has both checkpoint-dependent and independent functions in DNA replication. DNA combing experiments have shown that in *mrc1* mutants exposed to HU DNA replication cannot resume at stalled forks, demonstrating that Mrc1p is required for completion of DNA replication (Tourriere, *et al,* 2005). Mrc1p interacts with Pol2p (Lou, *et al*, 2008) and this association places Mrc1p close to the active site of DNA polymerase where it might sense replication stress and respond by activating the checkpoint response. Mrc1p is also a component of replication forks in unperturbed S-phase (Fig.16) and is required for the normal rate of replication as demonstrated by the slower progression observed, with dense isotope substitution, in *mrc1* mutants (Hodgson, *et al*, 2007). Mrc1p forms a complex with Tof1p and Csm3p, which moves with replication forks and is required to stabilize the replisome complex in the presence of replication stress (Bando, *et al*, 2009). Deletion of Tof1p or Mrc1p in cells treated with HU leads to uncoupling of Cdc45p and helicase activity (Katou, *et al*, 2003).

Stalling of replication fork at natural sites, such as RFB, requires the presence of Tof1p; Mrc1p is instead dispensable to maintain stable replisome in this scenario (Fig.16) (Calzada, *et al*, 2005; Hodgson, *et al*, 2007). Indeed in *Δmrc1* mutants replication forks still pause at an RFB and no uncoupling of DNA synthesis and

helicase activity can be detected. However, it has been shown that Mrc1p in association with Tof1p might instead be required for replication fork stalling at tRNA fragile sites (Tourriere, *et al*, 2005). Fork progression through RFB sites depends on the presence of the Rrm3 helicase (Fig.16) (Calzada, *et al*, 2005). Integrity of replication forks at these sites per se is a checkpoint independent mechanism because neither Mec1p nor Rad53p are required for pausing or recovery of forks at RFBs (Calzada, *et al*, 2005).

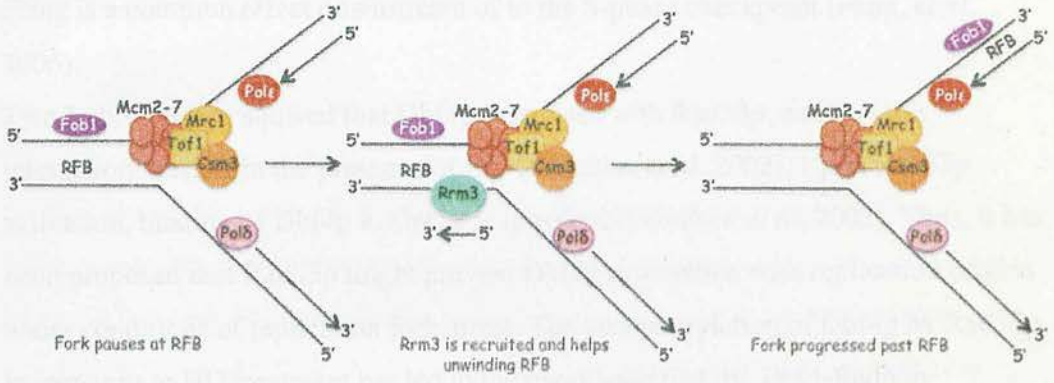


Fig.16 Model for pausing of replication forks at RFBs (adjusted from Calzada, *et al*, 2005). Stability of replication forks paused at RFBs required Tof1p, Csm3p and Fob1 (replication fork block) proteins. Recruitment of Rrm3p helicase promotes DNA unwinding and progression of the fork through the RFB.

1.32 INHIBITION OF LATE ORIGIN FIRING

Early replication origins fire in cells treated with HU; however, in the same scenario no replication intermediates can be visualized at late origins (Santocanale and Diffley, 1998). These results demonstrate that the activation of the S-phase checkpoint in response to HU treatment inhibits the firing of late origins. This inhibition occurs at the conversion from pre-RC to RC and requires both Mec1p and Rad53p as demonstrated by the analysis of replication intermediates in *mrc1*-1 and

rad53-11 mutants (Santocanale and Diffley, 1998). Genome-wide analysis confirmed that sites of ssDNA associate with replication origins (Feng, *et al*, 2006). Analysis of the ssDNA profiles of WT cells treated with HU, revealed that in WT cells ssDNA is not exclusively restricted to origins. This result indicates that the replication forks progress, albeit at a very slow rate. However, in *rad53* mutants ssDNA structures could be detected only at replication origins suggesting a complete stalling of replication forks. ssDNA formation occurred at restricted origins in WT cells and at almost every origin in *Δrad53* cells demonstrating that Rad53p is required to prevent the activation of a subset of origins in the presence of HU (Feng, *et al*, 2006). Similar analysis on fission yeast indicated that inhibition of late origin firing is a common effect downstream of to the S-phase checkpoint (Feng, *et al*, 2006).

Two-hybrid assays showed that Dbf4p associates with Rad53p, and that this interaction persists in the presence of HU (Duncker, *et al*, 2002). Upon Rad53p activation, binding of Dbf4p to Orc2p is impaired (Duncker, *et al*, 2002). Thus, it has been proposed that Rad53p might prevent Dbf4p association with replication origins under conditions of replication fork arrest. The phosphorylation of Dbf4p by Rad53p in response to HU treatment has led to the hypothesis that the Dbf4-Rad53p interaction is involved in late origin inhibition (Santocanale and Diffley, 1998). Recent data have led to new insights on the regulation of origins in the presence of replication stress. Two groups have demonstrated that, apart from Dbf4p, the Sld3 protein is an important target of Rad53p (Lopez-Mosqueda, et al, 2010; Zegerman and Diffley, 2010). Phosphorylation of these replication proteins is required to inhibit both the CDK- and the DDK- dependent pathways that are essential for the inhibition of late origin firing .

Mrc1p is involved in both DNA replication and regulation of origin firing, downstream of the S-phase checkpoint activation. Mrc1p associates with replication origins as cells begin DNA synthesis. This interaction is dependent on Dbf4p as demonstrated by the absence of Mrc1p on late replication origins in HU treated cells (Osborn and Elledge, 2003). Δ*mrc1* cells have a prolonged S-phase (over two-fold

longer) and exhibit increased activation of dormant origins, demonstrating a role for Mrc1p as a regulator of the overall replication program (Koren, *et al*, 2010). The chromatin context of replicating DNA is a major determinant of origin firing. Indeed, deletion of the histone deacetylase Rpd3 promotes the earlier activation of late origins (Vogelauer, *et al*, 2002). Moreover, in *rpd3* mutants treated with HU, late origins can escape inhibition of the S-phase checkpoint (Aparicio, *et al*, 2004). This phenotype is not due to inactivation of the checkpoint, as demonstrated by intact Rad53p phosphorylation, and is observable in the presence of HU but not MMS. This specificity has been explained by a delay in Rad53p activation in HU cells compared to MMS treated cells; this delay could contribute to the escape of late origins from checkpoint control (Aparicio, *et al*, 2004).

1.33 SPATIAL AND TEMPORAL REGULATION OF ORIGIN FIRING

Origins of replication are clustered together. Within each domain, origins are activated almost synchronously; however a variation in the time of firing can be observed between different clusters (Shirahige, *et al*, 2003; Blow, *et al*, 2001; MacAlpine, *et al*, 2004). Individual replication origins can be distinguished as early and late according to the time in which they fire during S-phase (Yabuki, *et al*, 2002). Thus, each origin is characterized by a specific replication time (T_{rep}) that indicates the time at which half of the specific replication fragment has reached maximal replication in a population of cells (Friedman, *et al*, 1997). However, replication timing is not an intrinsic property of origins (Ferguson and Fangman 1992). Cis-acting elements, such as chromosome position (Raghuraman, *et al*, 1997) and chromatin structure (reviewed in Donaldson, 2005), can determine early or late activation. Checkpoint proteins (Labit, *et al*, 2008) or CDK and DDK complexes (Donaldson, *et al*, 1998; McCune, *et al*, 2008; Sclafani *et al*, 2002) can regulate origin firing, acting as trans-acting factors.

Pre-RC assembly is the first step in the initiation of DNA replication and is required to render origins competent for firing. Even if the pre-RC is assembled on every origin, only a subset will fire in the following S-phase, and not all of them will be

activated with the same efficiency. Efficiency has been defined as the percentage of cells in which an origin is activated during any given S-phase (Raghuraman and Brewer, 2010). Notably, the T_{rep} value and efficiency of a replication origin are not tightly correlated (Weinreich *et al*, 2004). Indeed, origins that fire late S-phase can be efficiently activated (Friedman, *et al*, 1996). Moreover, T_{rep} is affected by various parameters and cannot discriminate between origins that actually fire and those that are instead passively replicated (reviewed in Raghuraman and Brewer, 2010). Genome-wide analyses in yeasts and metazoa have shown the distribution of replication origin and as well as mechanisms that regulate their temporal program.

1.34 TEMPORAL PROGRAM OF ORIGIN FIRING

The coordinated activation of replication origins is essential to ensure complete duplication of the genetic material. Various techniques have been developed to identify origin locations and study their replication dynamics in different model organisms. Combination of the collected data has led to the construction of genomic origin maps and the general view that origins are activated according to a temporal program that is imposed by chromosomal context and under the control of trans-acting factors (reviewed in Tuduri, *et al*, 2010). In eukaryotic cells, the temporal program is stably transmitted to daughter cells (Hiratani, *et al*, 2009; Lucas and Feng, 2003). However exceptions from the temporal program can be found among the different organisms.

1.35 REPLICATION ORIGIN ACTIVATION IN YEASTS

S. cerevisiae is the eukaryotic organism in which replication origins have been most widely studied with regard to their sequences and their temporal regulation. Comparison and phylogenetic conservation of replication origin elements between four yeast species, has allowed the localization of origins at base pair resolution (Nieduszynski, *et al*, 2006). Replication origins are associated with intergenic regions with the exception of ARS604 and ARS605 which are located within an

Open Reading Frame (ORF). The average distance between origins is 20-40 kbp. A closer distance, less than 7 kbp, can hinder origin activation by an "origin interference" mechanism. In this context DNA replication is initiated from one origin or the other with equal possibility (Marahrens and Stillman, 1992; Brewer and Fangman, 1993). Application of 2D gel analysis led to the identification of most *S. cerevisiae* replication origins at a chromosomal level. Chromosomes are organized into blocks of DNA sequences that are replicated in a fixed sequential order (Shirahige, *et al*, 1993) with early replicating regions interspersed with late activated domains (Raghuraman, *et al*, 1997). Chromosome (Chr) VI is one of the best-characterized genomic regions in budding yeast. According to the efficiency of the nine origins identified it is supposed to consist of five replicons: three in the left arm and two in the right arm (Yamashita, *et al*, 1997). A large variation in replication time has been observed near the telomeres on Chr. VI where slow fork movement is due to a combination of efficient replication origins (early and late) (Friedman, *et al*, 1997). However, these conclusion were based on chromosomal-level analysis of specific replication origins. The first demonstration of the temporal regulation of origin firing in budding yeast came from more extensive analyses at a the genome-wide level (Raghuraman, *et al*, 2001; Yabuki, *et al*, 2002; Wyrick, *et al*, 2001). These studies have identified most of the replication origins and elucidated how genome replication is under a temporal control. Raghuraman and colleagues used a variation of the Meselson and Stahl dense isotope transfer technique to study the kinetics of replication across the entire yeast genome (Raghuraman, *et al*, 2001). This analysis confirmed the existence of 332 effective origins of replication which are scattered throughout the genome. A replication profile for the whole genome was constructed by plotting the replication times as a function of the position on the chromosome (the replication profile of Chr. VI is reported in Fig.17). Peaks identify regions that replicated before neighboring domains and tall and short peaks correlate with early and late replication, respectively.

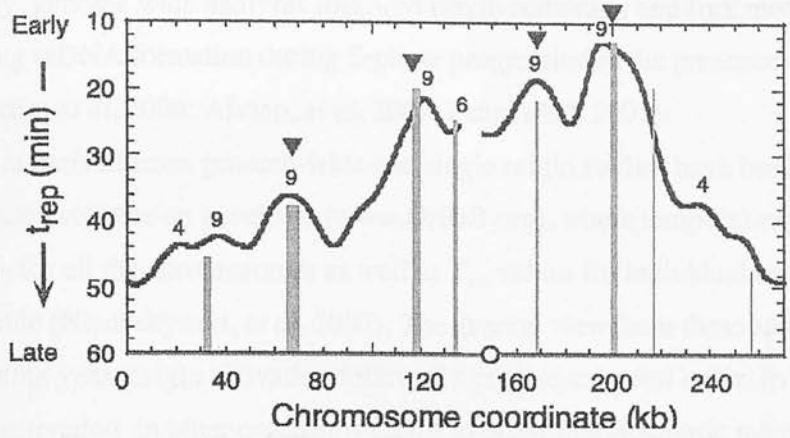


Fig.17 Replication profile of Chr. VI (Raghuraman, *et al*, 2001). The small circle on the X-axis represents the position of the centromere. The peaks correspond to replication origins and their height is correlated with the time of origin firing; the valleys are termination zones.

The replication profiles can predict the location of efficient origins and the relative time of replication for different genomic fragments. The array data demonstrated that *S. cerevisiae* origins fire according to a specific temporal program with some regions, such as centromeres, replicating early, and others, such as telomeres, replicating later in S-phase.

With a different approach Kitada and colleagues mapped early and late origins by using a microarray to monitor changes in DNA copy number, from one to two, during replication (Yabuki, *et al*, 2002). By performing the analysis in the presence of HU they identified the location of ARSs that are activated early in S-phase. The data confirmed the early replication of centromeric regions. They also found that early origins tend to cluster together within the genome resulting in regional differences in the time of replication completion. Wyrick and colleagues studied the location of replication origins throughout the genome by immunoprecipitation of the ORC and MCM complexes (Wyrick, *et al*, 2001). The sequencing or hybridization to microarray of the coimmunoprecipitated DNA revealed the genomic sites of pre-RC assembly. They identified 429 sites that they predicted to contain origins. Eighty of the predicted sites in Chr. X corresponded to active origins of replication. More

recently, genome wide analyses followed origin activation and fork movement by mapping ssDNA formation during S-phase progression in the presence or absence of HU (Feng, *et al*, 2006; Alvino, *et al*, 2007; Feng, *et al*, 2007).

The data derived from genome-wide and single origin studies have been collected and made available on a website (www.OriDB.org), where temporal replication profiles for all the chromosomes as well as T_{rep} values for individual origins are accessible (Nieduskynski, *et al*, 2007). The general view from these studies was that in budding yeast origin activation followed a precise temporal order. In contrast, origin activation in other organisms seems to occur in a stochastic manner. Interestingly, a recent DNA combing analysis of individual Chr. VI molecules from synchronized cells, *in S. cerevisiae*, suggested that origin firing does not follow a precise temporal program (Czajkowsky, *et al*, 2008). These data have questioned the existence of a staggered temporal program, which seems to be a observed exclusively in S. *cerevisiae*.

In the fission yeast *S. pombe* replication origins are predicted to be AT-rich islands that are spaced every 30 kbp and preferentially located in intergenic regions (Segurado, *et al*, 2003; Feng, *et al*, 2006). The firing of replication origin in *S. pombe* was analyzed by measuring the distribution of active origins using DNA combing; this technique involves the hybridization of fluorescent probes to stretched DNA molecukes (Patel, *et al*, 2006). Pulse-labeling DNA with a thymidine analogue, such as Bromodeoxyuridine (BrdU) allows visualization of the regions surrounding origins of replication. The profiles indicated that origins fire stochastically without regard to their neighbors indicating that firing is randomly distributed across the genome. The absence of temporal regulation and neighbors influence in origin activation has also been confirmed by analyses of single origins (Patel, *et al*, 2006). High resolution mapping, in *S. pombe*, of ssDNA formation (Feng, *et al*, 2006) and Orc1p and Mcm6p binding sites, in wild type and checkpoint-deficient *cds1Δ* cells both treated with HU (Hayashi, *et al*, 2007) identified the presence of distinct domains that replicate in different moment during the S-phase. Early and late replication origins tend to cluster separately into large chromosome regions accordingly to their time of activation, with the former being more efficient

compared to the latter. As for budding yeast, timing and efficiency in *S. pombe*, are not an intrinsic property of the origin but depend on local chromatin context (reviewed in Costa and Blow, 2007). Rate limiting trans-acting factors, such as pre-RC components and DDK, can influence both timing and efficiency due to a competition between origins for these proteins (Patel, *et al*, 2008; Wu and Nurse, 2009).

1.36 TEMPORAL PROGRAM IN HIGHER EUKARYOTES

In mammals replication origins do not share conserved sequences and form clusters distributed over 40 kbp. Replication timing in human cells is established early in G1 and is reestablished at every cell cycle (reviewed in Hiratani, *et al*, 2009). As in *S. cerevisiae* timing of origin activation is not an intrinsic property of origins; rather it is correlated with genomic structure. The first genome-wide analysis of replication of the human genome was performed by Woodfine and colleagues who performed genomic array hybridization to quantify relative DNA copy number changes (Woodfine, *et al*, 2004). Using this method the ratio between replicated (S-phase) and unreplicated (G1 phase) genomic material provides a measurement of the replication timing of a specific sequence. In contrast to *S. cerevisiae* (Raghuraman, *et al*, 2001) replication time is correlated with two chromatin features: gene density and GC content. Gene poor regions replicate late while a exon-rich regions correlate with early replication (Jeon, *et al*, 2005). GC-rich regions are also predictors of early replication (Woodfine, *et al*, 2004); this is in agreement with cytological banding studies where Giemsa G-dark regions, which are characterized by a high AT-rich content and gene poor density replicate late during S-phase. A high resolution genomic DNA microarray was performed to analyze the replication profile of Chr. 22 in two different cell lines (White, *et al*, 2004). As expected late replicating regions overlapped with G bands. Moreover, a correlation was found between transcriptional activity and early duplication, which also applied to gene-poor regions.

Correlation between early replicating and transcriptionally active regions seems to be conserved among eukaryotes as demonstrated by analysis of origin activity and gene expression on *Drosophila* Chr. 2 (MacAlpine, *et al*, 2004). Early and late replication origins were mapped using a similar approach to that of White and colleagues. Moreover, analysis of RNA Polymerase II (Pol II) enrichment along the chromosome revealed that the replication time is sharply defined by the transcriptional profile. Indeed, a higher density of Pol II correlates with early replication. A correlation between replication origin activity and transcription initiation was also observed in human cells (Karnani, *et al*, 2010). Indeed, an enrichment of replication initiation sites is observed in genomic regions 5 kbp downstream or upstream of transcription start sites (TSS). Consistent with these data origins of replication are enriched in proximity of Pol II binding sites (Karnani, *et al*, 2010).

However it must be considered that a general association between transcriptional activity and replication lines could not be observed in the different cell lines. Indeed, changes in replication time are more likely to associated with transcription competence rather than transcription per se (reviewed in Hiratani and Gilbert, 2009). Thus, early replication can define a chromatin organization that is necessary but not sufficient to induce transcription.

In *Xenopus* DNA replication is less correlated with gene expression. In this organism origins are spaced 5-15 kbp apart and grouped into clusters that on average cover approximately 40 kbp (Blow, *et al*, 2001). DNA fiber analysis revealed that clustered origins are activated almost synchronously, while each cluster of origins fires at different times throughout S-phase (Blow, *et al*, 2001). Moreover, analyses of egg embryos demonstrated that replication time is deterministic at the level of large chromosomal domains but stochastic at the level of single origins (Labit, *et al*, 2008).

1.37 THE RANDOM GAP PROBLEM

Two opposing models for DNA replication timing have arisen from the collective studies of DNA replication: a defined pattern of genome replication resembling the

replicon model, and the stochastic origin firing (Labit, *et al*, 2008). In the latter scenario, the firing of *S. cerevisiae* replication origins according to a defined temporal program (Raghuraman *et al*, 2001; Yabuki, *et al*, 2002) would be an exception (Rhind, 2006). However, the recent study of Chr. VI provided evidence of stochastic origin firing and thus called into doubt the existence of a temporal program (Czaykowski, *et al*, 2008). Nonetheless, replication in eukaryotes cannot occur exclusively according to a stochastic pattern. Evidences has suggested that the genome is replicated in a specific window of time and thus eukaryotes must have a defined pattern for origin firing (Donaldson, 2005). Indeed stochastic firing would lead to the "random gap" problem (Fig.18). A random distribution and activation of replication origins would result in large gaps of DNA between replication forks and prolong the completion of S-phase (Rhind, 2006).

Various solutions have been proposed to reconcile random distribution with efficient replication. In *Xenopus*, origins of replication are spaced 5-15 kbp apart, and this spacing provides a simple solution to the random gap problem while still allowing origins to be positioned randomly with respect to specific DNA sequences (Blow, *et al*, 2001).

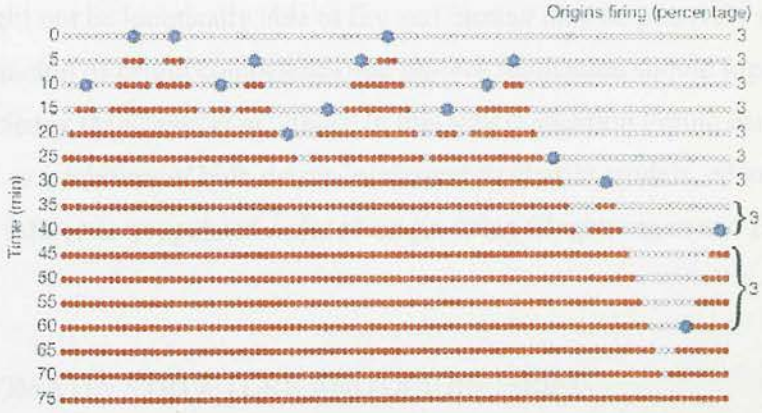


Fig.18 Representation of the random gap problem (Rhind, 2006). Fired origins are represented by blue circles; red and white circles indicate replicated and unreplicated regions, respectively. Random origin distribution and stochastic firing of inefficient origins wold lead to large gaps between origins and a longer S-phase.

S-phase can be completed in a determined window of time if origins fire stochastically and with more efficiency with the proceeding of S-phase (Rhind, 2006). This scenario could occur in the case where origins compete for rate limiting factors and thus can fire in different moments during S-phase (Raghuraman and Brewer, 2010). As one example of this replication model, in *S. pombe*, the level of replication proteins influences origin efficiency. For example, cells regulate the amount of the DDK subunit Dfp1p in order to control origin efficiency and maximize genome stability (Patel, *et al*, 2008). Analyses of the replication pattern in budding yeast cells beraing a mutant Clb5p, the S-phase cyclin, is responsible for late origin activation (Donaldson, *et al*, 1998), have demonstrated the occurrence of staggered temporal regulation (McCune, *et al*, 2008).

By changing some parameters, the deterministic model can fit with the variation of origin firing observed in the stochastic model (Rhind, 2006). According to McCune and colleagues, the word "efficiency" has been used with different and confusing meanings that have led to discordant interpretation of replication timing profiles. On one side, "efficiency" indicates an empirical value that describes "the percentage of cells in a population observed to have fired a particular origin". McCune and colleagues propose that origin "competence" should be considered as the percentage of cells in which an origin is biochemically competent to fire; however competent origins might not be kinetically able to fire and instead may be passively replicated. The combination of origin competence and passive replication should together define origin efficiency (McCune, *et al*, 2008). In this way replication timing can be considered as a mixture of both deterministic and stochastic models, where stochastic events can influence a regulated order of origin firing (Raghuraman and Brewer, 2010).

1.38 CHROMATIN STRUCTURE AND ORIGIN FIRING

The properties that characterize replication origins, such as timing and efficiency, are not intrinsic features of the origins but rather depend on chromosome location and on surrounding chromatin structures. In particular telomere proximity (Ferguson, *et al*,

1991; Ferguson and Fangman, 1992), nuclear localization (Heun, *et al*, 2001), nucleosome positioning (Eaton, *et al*, 2010) and histone modifications (Vogelauer, *et al*, 2002; Cimbora, *et al*, 2000) are all features of chromatin that regulate replication origin firing.

1.39 NUCLEAR LOCALIZATION AND NUCLEOSOME POSITIONING

Early studies of the activation of the late and early *S. cerevisiae* origins, ARS501 and ARS1 respectively, demonstrated that the different temporal programs were imposed by cis-acting surrounding regions of the origins (Ferguson and Fangman, 1992). Changing the chromosomal location of the ARS chromosome could alter the time of activation. ARS501, which is located in the proximity (27 kbp) of the telomere region on the right end of Chr. V, is an efficient origin per se (Ferguson, *et al*, 1991). Its late activation during S-phase is due to a telomeric effect (Raghuraman, *et al*, 1997) as demonstrated by its efficient and early activation in the absence of telomeric structures. The telomere position effect (TPE) is established at every cell cycle between mitosis and the START decision point (Raghuraman, *et al*, 1997). TPE is dependent on the Ku70 protein that, in association with Ku80p, contributes to a specialized telomere chromatin structure (reviewed in Weinreich, *et al*, 2004). Indeed deletion of Ku70p advances ARS501 timing and this effect is extended to all late telomere-proximal origins in yeast chromosomes (Cosgrove, *et al*, 2002). However, the TPE does not affect origins that are located more than a few tens of kilobases away (reviewed in Gilbert, 2001). Late origins located in internal chromosome regions are instead regulated by different mechanisms as demonstrated by studies of Chr. XIV, which contains a cluster of four origins spanning a region of 140 kbp (Friedman, *et al*, 1995). Late activation of these ARSs (1411-1414) relies on specific surrounding regions (Friedman, *et al*, 1996) that act as late determinants. These domains might nucleate specific chromatin structures, act as binding site for factors that impair replication initiation, or even target chromosome regions to specific nuclear compartments.

Studies of the Chr. XIV cluster and ARS501 have shown that late replicating regions are localized to the nuclear periphery, whereas early origins tend to be randomly localized (Heun, *et al*, 2001). The peripheral localization of late replication origins is neither stable neither required to maintain the temporal program. Therefore, the prerequisites for late origin activation are both the presence of late determinant chromatin structures and the localization to the nuclear periphery (Heun, *et al*, 2001). A recent study has shown that activation timing is not delayed when an early origin is tethered to the nuclear periphery (Ebrahimi, *et al*, 2010) indicating that the repositioning of chromosome regions toward the nuclear envelope is not the main determinant of the temporal program.

DNA replication and origin activation are influenced by the position of nucleosome surrounding ARSs. The first insight into this correlation came from studies on ARS1 (Venditti, *et al*, 1994). ARS1 activity is disrupted when a nuclesome is positioned over the ACS, while displacement of the nucleosomes surrounding the origin decreases its efficiency. Nucleosome mapping by indirect end labeling elucidated a central role for the ORC complex and Abf1p, a DNA binding protein, in this process (Lipford and Bell, 2001); both proteins are required to form the ARS1 nucleosome pattern specific . However studies of ARS307, which does not contain the Abf1 binding site B3 element, demonstrated that ORC is the main determinant of nucleosome positioning. ORC is required and sufficient to promote nucleosome positioning so that ARS1 will be asymmetrically positioned in a nucleosome free region (NFR) (Eaton, *et al*, 2010).

1.40 HISTONE MODIFICATIONS

Post-translational histone modifications can influence the accessibility of replication factors to origins and affect the timing of origin firing. Deletion of the deacetylase Rpd3p leades to an increase in histone acetylation and a general advancement of origin activation that is especially evident at late origins (Vogelauer, *et al*, 2002). The increase in histone acetylation levels results in earlier binding of Cdc45p to late origins. The earlier reruitment of Cdc45p was confirmed by tethering of the Gcn5p

histone acetylatransferase to the late origin ARS1412 (Vogelauer, *et al*, 2002; Espinosa, *et al*, 2010). Similar results have also been observed at the human β-globin replication origin (Cimbora, *et al*, 2000; Goren, *et al*, 2008). In *S. cerevisiae* Rpd3p is present in two complexes, Rpd3L (Large) and Rpd3S (Small), which differ in composition and function (Lee and Shilatifard, 2007). Genome-wide analysis of replication timing found that Rpd3L is responsible for the global changes in origin firing while Rpd3S has a role at only a subset of Rpd3L-regulated origins (Knott, *et al*, 2009); these disparities mey reflect the different roles of these proteins in regulating gene expression. Rpd3S and the NuA4 histone acetylase complex share the Eaf3 subunit that can bind methylated H3K36 residues which are involved in both transcription and replication. NuA4 and Rpd3S genetically interact with the FACT chromatin remodeling complex and compete for methylated H3K36 binding (Biswas, *et al*, 2008). A decrease in Rpd3S association with promotes NuA4 binding and stimulating FACT activity.

Moreover, FACT mutants are sensitive to agents that interfere with DNA replication demonstrating its role in the process; however, mutation of Set2p, a histone methyltransferase that methylates the H3K36 residue can suppress the sensitivity of FACT mutants. This genetic interaction provided the first evidence of a negative role for Set2p in DNA replication (Biswas, *et al*, 2008). Further studies have elucidated the role of Set2p in replication showing that Set2p is required for accelerated S-phase progression in the absence of Rpd3p (Pryde, *et al*, 2009). The deletion of Set2p delays binding of Cdc45p tp origins and eliminates the earlier association to late replication origins observed in *Δrpd3* cells. These data confirm the pivotal role of H3K36 post-translational modifications (acetylation and methylation) in DNA replication. ChIP analyses also demonstrated that the different states of H3K36 methylation influence origin firing. In particular, trimethylated H3K36 is enriched at late origins while monomethylated H3K36 is found predominantly at early origins (Pryde, *et al*, 2009). Thus, the balance of methylation and acetylation regulates replication timing through the recruitment of Cdc45p to origins. Recent mass spectrometry analysis of isolated origins-containing plasmids confirmed that acetylation of multiple lysine residues on H3 and H4 tails regulates origin firing

(Unnikrishnan, 2010). This study also found a correlation between H3K27 acetylation and H3K36 methylation. Another histone modification linked to DNA replication is represented by the acetylation of H3K56, which is mediated by Rtt109p. Indeed H3K56 acetylation has been shown to be involved in telomere localization (Hiraga, *et al*, 2008) and maintenance of replisome stability (Han, *et al*, 2007). Thus, the balance between different histone modifications may regulate origin firing through the recruitment of different factors.

1.41 DORMANT ORIGINS

Some replication origins are referred to as "dormant origins" because they are usually passively replicated. These ARSs are more likely to fire in the presence of replication stress (Santocanale, *et al*, 1999). In *S. cerevisiae*, the dormant origin ARS301 is not activated during an unperturbed S-phase as demonstrated by footprinting analysis whereby a direct transition from pre-RC to post-RC was observed. However, in cells harboring checkpoint mutations that are treated with HU, replication intermediates can be observed at ARS301; thus the origin can be activated in response to perturbed DNA replication. Indeed, the activation of dormant origins has an important role in maintaining genome stability in many organisms (Blow and Ge, 2009). Experiments in human cells in which the *MCM5* gene is knocked out indicated that the level of Mcm2-7 proteins influences dormant origin activation (Ge, *et al*, 2007); reduction of Mcm5p levels reduces the number of dormant origins that are activated.

On the other hand the excess of Mcm2-7 complexes activates these origins, which are normally inhibited by the S-phase checkpoint, in response to replication stress. Further studies confirmed that full licensing conditions are required to enable a reservoir of silent replication origins (Ibarra, *et al*, 2008); thus, a model explaining dormant origins activation has been proposed (Fig.19 A).

Under normal conditions an excess of Mcm2-7 complexes is required to license all the possible origins within a replicon. The firing of a replication origin sends a signal to inhibit activation of neighboring origins. However stalled forks will promote firing of silent origins that, in this condition, would have more time before being passively replicated (Fig.19 B) (Blow and Ge, 2008). Thus dormant origins are a backup system to maintain genome stability under stress conditions.

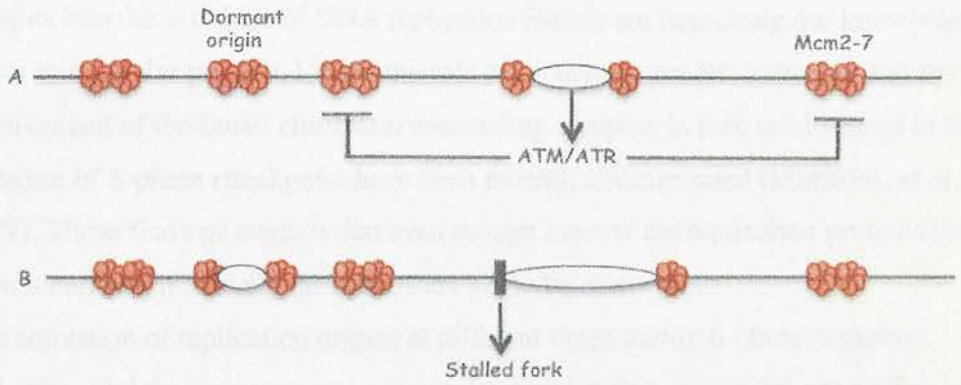


Fig.19 Model for dormant origin activation (adapted form Ibarra, *et al*, 2008). A) Under normal conditions origin activation sends a signal to inhibit firing of neighboring origins. B) Stalled forks induce the activation of dormant origins.

1.42 OVERVIEW OF PROTEINS INVOLVED IN DNA REPLICATION

DNA replication is an essential process that occurs through the coordinated actions of multiple factors. As described above, DNA replication involves various stages, each one characterized by the activity of specific replication factors. The proteins that take part in pre-RC assembly, such as Cdc6p, Cdt1p and the ORC and MCM complexes, are required in G1 phase to make origins competent for replication.

Transition to the RC is promoted by CDK and DDK complexes whose activities stimulate the recruitment of components of the RC, such as Cdc45p and GINS. Loading of the DNA polymerases is then required to initiate DNA synthesis. The entire genome must be correctly duplicated. Therefore, the proteins involved in fork stability and the S-phase checkpoint, such as Tof1p, Mrc1p and Rad53p, are required during DNA replication.

Redundant mechanisms have evolved to ensure that during DNA replication the stability and sequence of the genome are preserved. The complex process of genome duplication is fundamental for cell survival. Understanding the mechanisms of replication is vital to preventing diseases correlated replication defects. Recent insights into the activities of DNA replication factors are improving our knowledge about this cellular process. Indeed the role of Mcm9p in pre-RC assembly and the involvement of the Ino80 chromatin remodeling complex in fork stabilization in the presence of S-phase checkpoint have been recently demonstrated (Morrison, *et al*, 2007). These findings suggest that even though most of the replication proteins have been identified, it is likely that others are yet to be discovered.

The activation of replication origins at different times during S-phase is mainly determined by the chromosomal context. Late replication origins are generally located within heterochromatic regions and changes in chromatin structure due to posttranslational histone modifications can interfere with their activation. However, even if most of the proteins that regulate origin firing have been identified, the mechanisms which determine the early and late activation of replication origins remain to be clarified.

2.0 MATERIALS AND GENERAL METHODS

2.1 ANTIBODIES, ANTIBIOTICS AND OTHER DRUGS

Antibodies, and antibiotic were all purchased from SIGMA (unless otherwise specified). Protease inhibitors were purchased from Roche

2.2 CHEMICAL SUPPLIES

The salts were purchased from SIGMA or from Fisher Scientific unless otherwise specified; chemical supplies from different brands are indicated: Acrylamide 30% v/v (19:1 acrylamide/bis-acrylamide, Severn Biotech) and 40% v/v (19:1 acrylamide/bis-acrylamide, Fisher Chemical), ethanol and phenol-chloroform isoamyl alcohol 25:24:1 (VWR), TEMED (Tetramethylethylenediamine), dye for BSA assay, BBP (BromoPhenolBlue) and XC (Xilencianol) (BIORAD).

2.3 MOLECULAR WEIGHT MARKERS AND RESTRICTION ENZYMES

DNA and protein high molecular weight markers were purchased from Fermentas; restriction enzymes were purchased from Fermentas and New England BioLabs.

2.4 MEDIA AND PLATES

BactoAgar, Bactopeptone, Yeast Extract and Yeast Nitrogen Base (YNB) from BD. Aminoacids, Complete Synthetic Medium (CSM) minus LEU, CSM minus TRP from Formedium. Ammonium sulfate, X-Gal, Glucose, Raffinose and Galactose from SIGMA. IPTG from MELFORM.

2.5 OLIGONUCLEOTIDES

NAME	SEQUENCE
R-site-A-f	5′-TATCCCGGGCGTTGTAAAACGACGGCCAGTGAA-3′
R-site-A-r	5′-TACAGATCTGACCATGATTACGCCAAGCTTGCA-3′

R-site-B-f	5′-TACAGATCTCGTTGTAAAACGACGGCCAGTGAA-3′
R-site-B-r	5′-TACAGATCTCGTTGTAAAACGACGGCCAGTGAA-3′
LT2Rev1413	5′-AACTATATAATACACTCGGAATACAACTTTAAAA CCTTGAAAATAAGAATATGCCTAGCAGGTCGACCAC-3′
LT2Fwd1413	5′-AGCAAAGGAGTCTTGATGCCAAGTGCTTCCCTCC AAGATTTTTATGGCTAGAATTCGAGCTCGTTTAAAC-3′
LT1Rev1413	5′-CATCTGCAACTTTCAAAGTTTACCGTTTTCTCAAT ACTCTACTTGCTAGTGAATTCGAGCTCGTTTAAAC-3′
LT1Fwd1413	5′-AAATACCTTTTACAAAGTGGATACAGGTTGCCTAT CACTACCGCCATTTCCCGAGATCATATCACTGTGG-3′
ARS1413cLT2	5′-CCGTAGGTACAATGCTGATG-3′
ARS1413intC2	5′-TTGGGTTCAGAAGGAGACTAGGAG-3′
ARS1413intC1	5′-CTGTAGATTGCTCTTGGCGACAGA-3′
ARS1413cLT1	5′-GATTACGCGTCACAGCTACT-3′
TRP1intF	5′-CTGCAACATACTACTCAGTGCAGC-3′
KanMXintR2	5′-TCAGAAACAACTCTGGCGCATC-3′
ORC2FlagF	5′-ACACGTATGCGGAACTTGAAAAACTTCTGAAAACG TTTTAAATACTCTACGGATCCCCGGGTTAATTAA-3′
ORC2FlagR	5′CAAAGCTAGCAAGCCTAGTACTATTACAATTGTTCG TGATATGTATACATGAATTCGAGCTCGTTTAAAC-3′
ORC2upF	5′-GTCCTAAGCGTGGTACTCAAAG-3′
ORC2downR	5′-GAGGCACCGATTCTGATTCAAC-3′
ARS1413probeF	5′-AGGGTAGAGGACGATAGTGAAGGA-3′
ARS1413probeR	5′-GGTAGGATTGATATTTGGGGCG-3′
BBR2F	5′-GAGAAAGCTTAGTCCATTCGGC-3′

BBR2R	5′-CACCGATACGTACTTAAACTCT-3′
T7 primer	5′-ATACGACTCACTATAGGGCG-3′
Bar1upF	5′-ATGAGTCCTTAAGAAGGCCGTTG-3′
Bar1cIntR	5′-ATCGGCAGATCCGGTATCAAAC-3′
ARS603A	5′-CATAGGTTAAATAGGACGTTCTCC-3′
ARS603B	5′-CATTTCTCAATTTCGAGGC-3′
BAR1-Hyg F	5′-TTTTAGTTTATAGATAACGGCTCTTGCCGAATTCA TAGGCTGCACTCATTCGTACGCTGCAGGTCGAC-3′
BAR1-Hyg R	5′-TTGTTTTTTTATTATGCTTTCCATGTATTAAAAATG ACTATATATTTGATATCGATGAATTCGAGCTCG-3′
Hyg-IntR	5′-TCGTATGTGAATGCTGGTCG-3′

2.6 PCR

Taq Polymerase, 25 mM MgCl2 and 10 X PCR buffer (500 mM KCl, 100 mM Tris-HCl pH9, 1% v/v TritonX-100) were made in house (unless otherwise specified).

The dNTP stock solutions (100 mM) were purchased from Roche.

The PCR machine employed was purchased from PTC-100 MJ Research.

2.7 PLASMIDS

NAME	**DESCRIPTION**
pHM401	Plasmid derived from the insertion of the 58bp pSR1 region into the SalI site of the pUC19 plasmid (Araki, *et al*, 1992)
pHM153	Plasmid containing the Recombinase (R) gene under the Gal promoter (Matsuzaki, *et al*,1990)
pRINT	Plasmid derived from pHM153 by deletion of the EcoRI fragment (Matsuzaki, *et al*, 1990)

EasyVector	High copy-number vector that contains the T7 RNA polymerase promoter flanking a multiple cloning region within the coding region of the enzyme β–galactosidase (Promega)
EV/RsiteA	RsiteA fragment subcloned into EasyVector (this study)
EV/RsiteB	RsiteB fragment subcloned into EasyVector (this study)
pAF6a-KanMX6	Plasmid containing the *KanMX6* gene (Longtine, *et al*, 1998)
pAF6a-TRP1	Plasmid derived from pAF6a-KanMX6 by replacing the *KanMX6* module with the *S.cerevisiae TRP1* gene (Longtine, 1998)
pLT1/RsiteA	RsiteA fragment AvaI/BglII inserted into pAF6aKanMX6 AvaI/BglII (this study)
pLT1/RsiteB	RsiteB fragment AvaI/BglII inserted into pAF6aTRP1 AvaI/BglII (this study)
pGAL-VP16-ER	Plasmid containing the DNA-binding domain of GAL4, the activation domain of the viral protein 16 (VP16) and the hormone binding domain (HBD) of the human estrogen receptor (ER) (E.Schirmer)
FLAG-Nat	Plasmid containing the *CloNAT*-3X-FLAG sequence (S.Kagansky)
pAG32	Centromeric plasmid containing the sequence of the Hygromycin gene
pRS405	Integrative shuttle vector containing the *LEU2* sequence (Stratagene)

2. 8 YEAST STRAINS

NAME	GENOTYPE
W303-1a	*Mata ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100*
MMY021	*Mata, ade2-1, can1-100, his3-Δ200,leu2-3,112, trp1, and ura3-52, bar1::HIS3, ura3::LEU*
MMY033	*Mata, ade2-1, can1-100, his3-Δ200,leu2-3,112, trp1, and ura3-52, bar1::HIS3, Cdc45-Flag (KanMX), ura3::LEU2*

MVY51	*Mata, ade2-1, can1-100, his3-Δ200,leu2-3,112, trp1, and ura3-52, bar1::HIS3, Cdc45-Flag (KanMX), ura3::LEU,rpd3::TRP1*
MVY106	*Mata ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100* **pLT1/RsiteA**
MVY107	*Mata ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100* **pLT1/RsiteA pLT2/RsiteB**
MVY126	*Mata ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100* **pLT1/RsiteA pLT2/RsiteB** pHM153
MVY128	*Mata ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100* **pLT1/RsiteA pLT2/RsiteB** pHM153 pGAL-VP16-ER
MVY171	*Mata ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100* **pLT1/RsiteA pLT2/RsiteB** *leu2::* two copies pRINT
MVY172	*Mata ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100* **pLT1/RsiteA pLT2/RsiteB** *leu2::* one copy pRINT
MVY187	*Mata ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100, ARS1413::URA3, ORC2-3X-FLAG*
MVY200	*Mata ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100* **pLT1/RsiteA pLT2/RsiteB** *leu2::* two copies pRINT, *ORC2*-3X-FLAG
MVY201	*Mata ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100* **pLT1/RsiteA pLT2/RsiteB** *leu2::* two copies pRINT, *ORC2-3X-FLAG, BAR1::Hyg*
MVY202	*Mata ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100* **pLT1/RsiteA pLT2/RsiteB** *leu2::* two copies pRINT *BAR1::Hyg*

2.9 AGAROSE GEL ELECTROPHORESIS

Gels employed for the resolution of DNA samples were comprised of 1% v/v Agarose (SIGMA) and 1X TBE (0.9M Tris, 0.9 M Boric Acid, 20 mM EDTA); electrophoresis was performed in 1X TBE buffer.

2.10 BACTERIA TRANSFORMATION

Transformation of bacteria cells (Neo Blue and DH5α) was performed following the protocol described in Sambrook and Russell (Sambrook and Russell, 2001).

2.11 COLONY PCR

Cells were picked into sterile 0.5 ml PCR tubes (STARLAB), heated in a microwave for 1 min, transferred to ice and then resuspended in 25µl of reaction mix (1X PCR buffer, 2.5 mM $MgCl_2$, 200 µM dNTPs, 1 µM of each Primer, 2 U/reaction Taq Polymerase). The PCR cycles consisted of an initial denaturating step of 1min at 95°C, which was followed by 30 cycles of a denaturating step at 95°C for 30 sec, an annealing step at 55°C for 30 sec and an elongation step at 72°C for 2 min, and a final extension for 5 min at 72°C.

2.12 COMPETENT *E. coli* CELLS

The DH5α and the Neo Blue competent cells were prepared following the protocol described in Sambrook and Russell (Sambrook and Russell, 2001).

2.13 DNA EXTRACTION FROM AGAROSE GELS, PURIFICATION OF PCR PRODUCTS AND PLASMID EXTRACTION

Extraction of DNA fragments from agarose gels, purification of the PCR products used for probe labeling and yeast transformation and plasmid extraction from bacteria cells were all performed using the appropriate Qiagen Kit and following the manufacturer's instruction.

2.14 GENE DISRUPTION

Disruption of endogenous gene sequences was performed by yeast transformation using amplified DNA fragments designed to integrate into the genome by homologous recombination.

2.15 YEAST GENOMIC DNA PREPARATION

The pellet from approximately 25 OD_{600} of cells was resuspended in 200 µl of Lysis Buffer (10 mM Tris-HCL pH8, 1 mM EDTA, 10 mM NaCl, 1% w/v SDS, 2% v/v TRITON-X100) and 300 µl of glass beads (SodaLime BioSPec Products) were added. DNA was recovered by phenol extraction (phenol-chloroform isoamyl alcohol 25:24:1 WWR) and precipitated according to the protocol described by Sambrook and Russell (Sambrook and Russell, 2001).The pellet was resuspended in 50 µl of 1X TE pH8.

2.16 LIGATION

Ligation reactions were performed at 16°C over night using the DNA ligase T4 and the Ligation Buffer provided (300 mM Tris-HCl pH 7.8, 100 mM MgCl2, 100 mM DTT and 10 mM ATP). Ligation products were then used for bacterial transformations.

2.17 SEQUENCING

DNA (200 ng) to be sequenced was incubated with 2 µl of BigDye3.1 solution (Applied Biosystem) and 3.2 pmol of primer in a total volume of 10 µl. The DNA was amplified by PCR (25 cycles: denaturation at 96°C for 30 sec, annealing at 50°C for 15 sec, elongation 60°C for 4 min) and the amplification product was sent to the Gene Pool service (University of Edinburgh) in order to be resolved on a Service AB13730 sequencer. Sequence alignment was performed with the SeqMan program.

2.18 SOUTHERN BLOT AND HYBRIDIZATION

Southern Blot analyses and hybridization were performed following the protocol for by Sambrook and Russell (Sambrook and Russell, 2001). The DNA was transferred by capillary transfer to a nitrocellulose membrane (PROTRAN Nitrocellulose Transfer Membrane, Schleicher & Schuell). For hybridization the Nitrocellulose membrane was incubated with the purified radioactive probe. Probes were labeled with α-dATP- P^{32} by random priming using the DECAprime TM II kit (Ambion). To eliminate the unincorporated nucleotide the labeled probe was purified using S-50 sephadex columns (Amersham) following the manufacturer's instructions. The membrane was then exposed on a phosphorscreen (Amersham) with subsequent analysis on a STORM phosphorimager. For dehybridization the membrane was soaked in 500ml of a boiling dehybridization solution (0.1X SSC, 1% w/v SDS) ; loss of radioactive signals was confirmed by exposure of the membrane on a phosphorscreen with subsequent analysis on a STORM phosphorimager.

2.19 YEAST MEDIA AND PLATES

The media and plates for yeast were prepared according to the protocol by Sambrook and Russell (Sambrook and Russell, 2001). For SD plates the amino acid drop out mix were used as follow: CSM-Tryptophan at 740 mgl/l and CSM-Leucine at 690 mg/l.

For selection of Kanamicin and Hygromycin resistant cells resistant cells the G418 antibiotic (200 µg/ml) and the Hygromycin B antibiotic (Roche) (200 µg/ml) were added to the media after autoclaving respectively. The CloNAT antibiotic (100 µg/ml) (Werner BioAgents) was used for selection of CloNAT.

2.20 BACTERIA MEDIA AND PLATES

The media and plates for bacteria were prepared according to the protocol by Sambrook and Russell (Sambrook and Russell, 2001).For LB plates ampicillin was

added to a final concentration of 50 μg/ml. For detection of β-galactosidase 20 μl of 50 mg/ml X-Gal and 100 μl of 100mM IPTG were added to LB/Amp plates.

2.21 YEAST TRANSFORMATION

10 ml cultures were grown to exponential phase (OD_{600}= 0.6-0.7); cells were washed with 40 ml of H_2O, resuspended in 1ml of Solution 1 (1X TE, 100 mM LiAc) centrifuged for 5 min at 1200 g and resuspended in 250 μl of Solution 1. Cells (100 μl) were treated with 5 μl of ssDNA (10 mg/ml), 1 μg of DNA fragment (10 ng for plasmid transformation) and 700 μl of Solution 2 (1X TE, 100 mM LiAc, 40% v/v PEG) After incubation at 30°C for 1.5 hr and heat shock at 42°C for 15 min cells were plated on the appropriate selective media.

2.22 WESTERN BLOTTING

PROTEIN PREPARATION

Yeast cells (10 ml cultures) grown to exponential phase were centrifuged for 5 min at 3500 rpm at 4°C and resuspended in 100 μl of cold Lysis Buffer (100 mM KCl, 50 mM Tris-HCl pH7.5, 1% v/v TRITON X-100, 10% v/v Glycerol) containing a cocktail of 1X protease inhibitors (Complete-EDTA free, ROCHE) added just before use. The cell wall was mechanically disrupted using 100 μl glass beads (SodaLime BioSPec Products) and vortexing in a Rybolyser (Hybaid) for 20 sec (setting 4). 100 μl of cold Lysis Buffer with protease inhibitors were added to the samples and the protein fraction was recovered by centrifugation for 2 min at 16000 g at 4°C. The protein concentration was measured by Bradford assay (Bio-Rad Protein Assay).

BLOT AND HYBRIDIZATION

Protein samples were separated on SDS-polyacrylamide gels. Western blot was performed according to the protocol described by Sambrook and Russell (Sambrook and Russell, 2001) . The protein samples were transferred to the Millipore Immobilon P, 0.45 μm membrane. For protein detection the membrane was incubated with α-FLAG monoclonal antibody (1:2500) and with α-Mouse antibody

conjugated to horseradish peroxidase (1:5000) (α-Mouse IgG-HRP, NA931, Amersham supplied by GE healthcare) as first and second antibody. The membrane was then treated with the ECL plus solutions (GE- Healthcare) for 1min at Room Temperature (RT) and exposed on film (Konika Minolta).

2.23 MATERIAL AND METHODS SPECIFIC TO THIS STUDY

2.24 CONSTRUCTED PLASMIDS

EV/A and **EV/B** were constructed by insertion of RsiteA and RsiteB into the EasyVector. Plasmid pHM401 (Araki, *et al*, 1992) was amplified with two different pairs of primers (RsiteAf-RsiteAr and RsiteBf-RsiteBr) giving a 179 bp fragment (Fig.20 A).

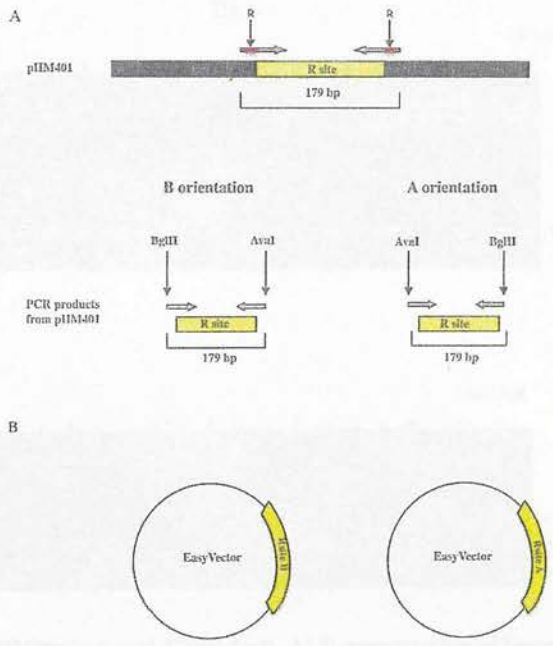


Fig.20 Strategy to construct vectors EV/A and EV/B. A) Schematic representation of the pHM401 region containing the sequence of the R site. The positions of the primers are reported. The R on the arrows indicates the site of the restriction enzymes. Below are details of the amplified recombination sequences 179 bp. The purple arrows represent the primers with the AvaI and BglII restriction sites; the two different orientations (A and B) obtained are shown. B) Representation of the EV/B and EV/A constructs resulting from subcloning R site A and R site B into the EasyVector.

These primers contains a 5' tail with restriction sites for the AvaI or BglII restriction enzymes (Fig.20 B) and will allow the production of the R sites in two different orientations, A and B, respectively. Correct insertion of EV/A and EV/B was confirmed by ApaI and SpeI digestion (Fig.21 A), which produces a 210bp fragment. The selected colonies for plasmids EV/A (panel A Fig. 21) and EV/B (panel B Fig.21) confirmed to harbor a correct insert (by T7 primer-directed sequence verification) were selected for further analysis.

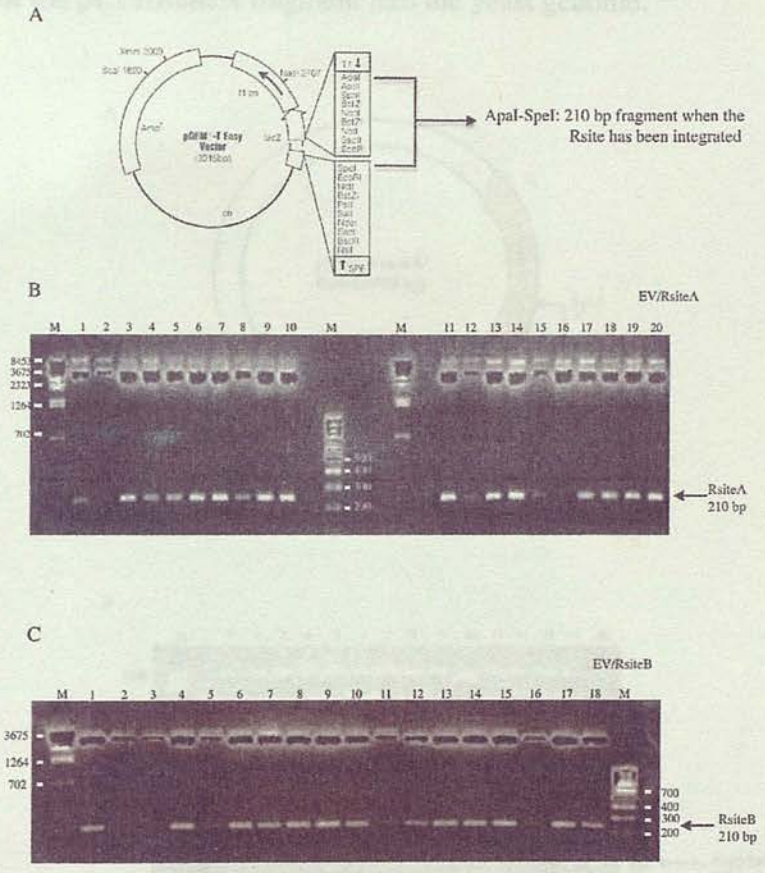


Fig.21 Construction of EV/RsiteA and EV/RsiteB. A) Representation of Easy Vector and restriction sites used to confirm the transformants. B-C) Miniprep preparations of the colonies obtained after transformation and resricted with ApaI and SpeI (210 bp). B) Lanes 1-20: transformants for the RsiteA orientation. C) Lanes 1-18: transformants for the RsiteB orientation. M: λ/Eco911 (left side of both gels) and Mass Ruler (100 bp ladder) DNA MW ladders. The gels demonstrate the successful construction of vectors EV/A and EV/B.

pLT1/RsiteA was derived from ligation of the RsiteA fragment (210 bp, AvaI-BglII) with the *KanMX6* gene within plasmid pAF6a-KanMX6 (3916 bp, AvaI-BglII). After ligation, products were transformed into Neo Blue bacterial cells and plated on LB/Amp plates. Colonies harboring the correct plasmid were confirmed by NdeI and SphI digestion (Fig.22 A). The samples loaded in lanes 2, 6, and 11 (Fig.22 B) are positive clones for the insertion of RsiteA into pFA6a-KanMX6. Colonies harboring the correct plasmid were confirmed via sequence analysis with the forward (RsiteA-f) and reverse (KanMXIntR) primers, and employed in ths subsequent integration of the pLT1/RsiteA fragment into the yeast genome.

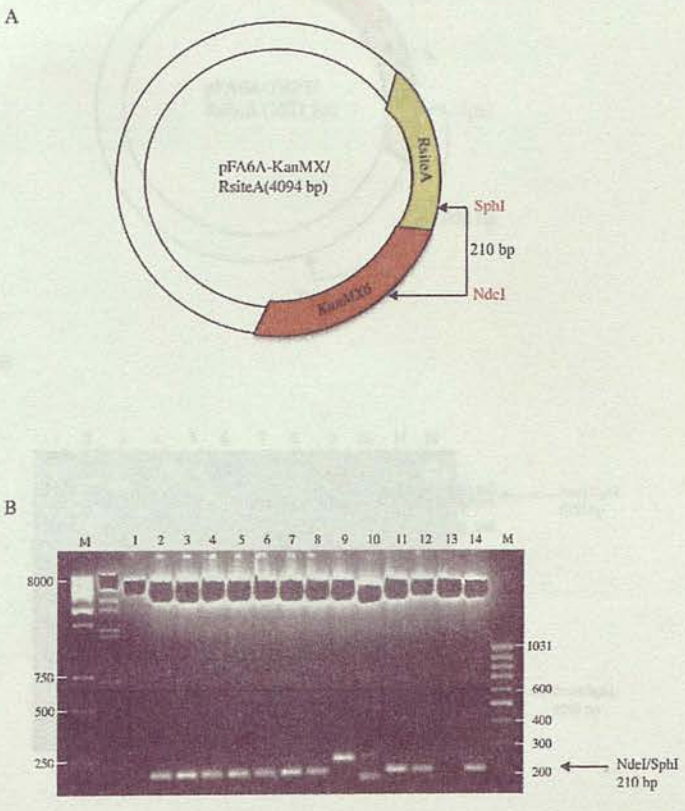


Fig.22 Cloning of the R site A into pFA6a-KanMX6. A) Schematic map of the plasmid pF6a-KanMX/RsiteA constructed by ligation of the RsiteA fragment with the *KanMX6* gene. The SphI and NdeI restriction sites are shown; the digestion product is a fragment of 210 bp. B) Lanes 1-14: miniprep preparations restricted with NdeI and SphI. A band of 210 bp is obtained if the insert is present inside the vector; the arrow indicates the digestion product. M: 1kbp plus (left side of the gel) and Mass Ruler DNA MW ladders. Positive clones were obtained for construction of plasmid pLT1/RsiteA.

pLT2/RsiteB was derived from ligation of the RsiteB fragment (210bp, AvaI-BglII) with the *TRP1* gene within plasmid pAF6a-TRP1 (3321bp, AvaI-BglII). After ligation, correct construction of pLT2/RsiteB was confirmed by PmeI and SphI restriction (Fig.23 panel A). One positive clone was obtained (lane 23 panel B). The correct integration of the RsiteB was confirmed upon sequence analysis using the TrpIntR primer (Fig.23 B). The clone had the exact sequence of the recombination site and therefore it was employed in the subsequent experiment.

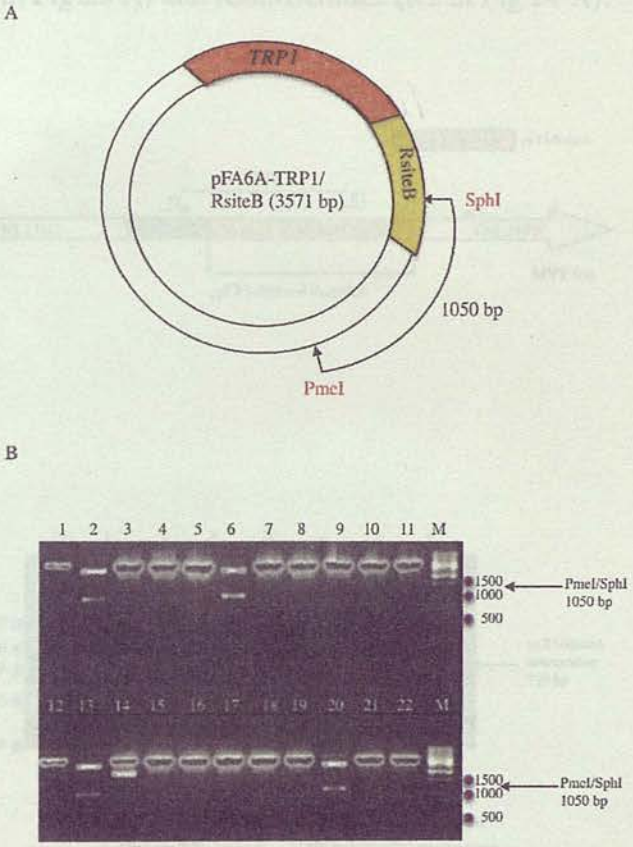


Fig.23 Cloning of the RsiteB into pFA6a-TRP1. A) Schematic map of plasmid pF6a-TRP1/RsiteB constructed by ligation of the RsiteB fragment with the *TRP1* gene. The SphI and PmeI restriction sites and their digestion product of 1050 bp are shown. B) Lanes 1-22: miniprep DNA preparations from the colonies obtained after transformation and restricted with PmeI and SphI; the arrow indicates the 1050 bp fragments. M: λ/Eco911 and 1 kbp plus DNA MW ladders. Four colonies contained RsiteB ligated to *TRP1*.

2.25 CONSTRUCTED YEAST STRAINS

MVY106, which contains the pLT1/RsiteA construct integrated at the right end of ARS1413, was created by transformation of W303-1a with the pLT1/RsiteA fragment. A 1.6 kbp RsiteA fragment (amplified via PCR using the primers LT1FwdARS1413 and LT1RevARS1413), was integrated between the 3′-end of ARS1413 and the 5′-end of *YNL210W* (Fig.24 A). Correct insertion was verified using colony PCR and the following primers: ARS1413cLT1 and ARS1413intC1 (F1 and R1 arrows in Fig.23 A) and KanMXIntR2 (R2 in Fig.24 A).

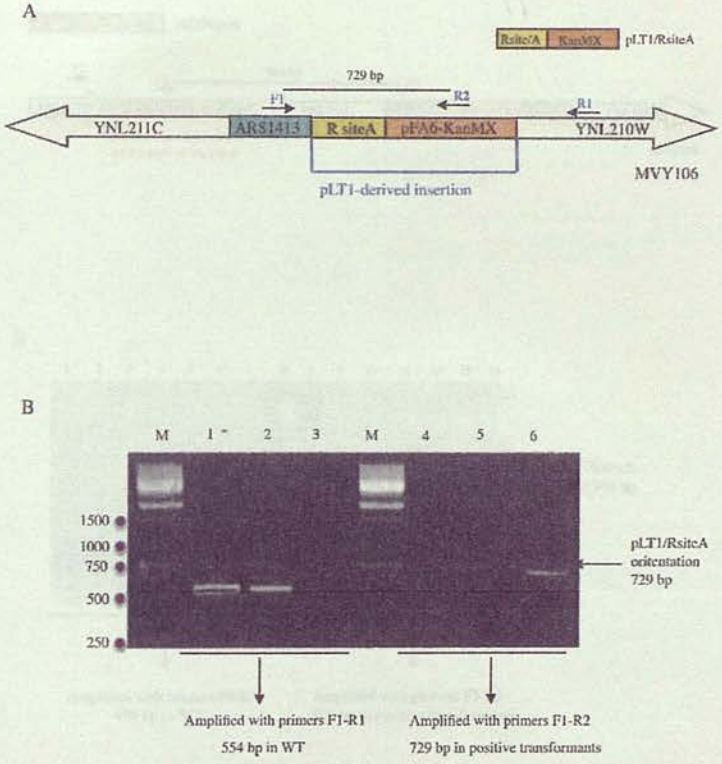


Fig.24 Integration of the pLT1/RsiteA fragment into the right side of ARS1413. A) Details of Chr. XIV after transformation with fragment pLT1/A. The arrows represent the locations of primers used for colony PCR; M: 1 kbp plus DNA MW ladder. B) PCR amplification to confirm the presence of the pLT1/RsiteA insert. Lanes 1-3: samples amplified with primers F1-R1 that yield a fragment of 554 bp in WT strain; lanes 4-6: same samples amplified with primers F1-R2 that yield a band of 729 bp if the pLT1/RsiteA fragment is inserted downstream of ARS1413. Samples 1 and 4 are from the WT untransformed strain (negative control). One positive clone (lanes 3 and 6) was obtained. No PCR amplification is observed in the lane 3 due to the size of the F1-R1 product (1804 bp).

MVY107 was derived by transformation of MVY106 with the pLT2/RsiteB fragment (1176 bp). The transforming fragment was prepared by amplification of plasmid pFA6a-TRP1/RsiteB with primers LT2FwdARS1413 and LT2RevARS1413. It integrates between the 3'-end of *YNL211C* and the 3'-end of the upstream *YNL212W* gene (Fig.25 panel A). Successful integration was confirmed by colony PCR with primers ARS1413cLT2 and ARS1413intC2 (F2 and R3 panel A Fig. 25) and TRPIntF (F3 Fig.25 panel A). One clone contained the pLT2/RsiteB transforming fragment upstream ARS1413 (lanes 1 and 8 Fig.25 panel B).

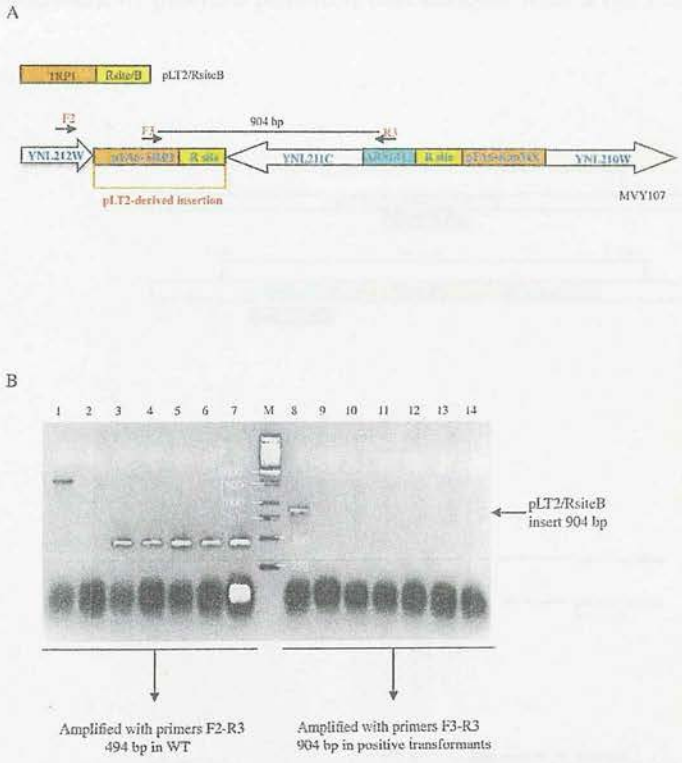


Fig.25 Integration of the pLT2/RsiteB fragment on the left side of ARS1413. A) Schematic representation of Chr. XIV after insertion of pLT1/RsiteA and pLT2/RsiteB fragments. The arrows indicate the primers used for colony PCR. B) M: 1kbp plus DNA MW ladder. Lanes 1-7: colonies amplified by PCR with primers F2-R3 that yield a band of 49 4bp in negative colonies and one of 1611 bp in positive colonies. Lanes 8-14: colonies amplified with primers F3-R3 giving a fragment (904 bp) only if the colonies contain the pLT2/RsiteB insert. Samples 7 and 14 are from the untransformed MVY106 strain (negative control). One clone (lanes 1 and 8) contained the pLT2/RsiteB fragment integrated into the left side of the ARS1413.

MVY126 was constructed by transformation of MVY107 with pHM153.

MVY128 derives from MVY126 transformed with pGAL-VP16-ER.

MVY171 and MVY172 contains one and two copies of pRINT (Matsuzaki, *et al*, 1988) inserted into the *LEU2* locus respectively. They were derived by transformation of MVY107 with pRINT (7 kbp), which has been previously linearized by ClaI that cuts once within pRINT (Matsuzaki, *et al*, 1988). The pRINT integration was confirmed by Southern blot using the *LEU2* probe (691 bp, prepared by EcoNI restriction of plasmid pSR405) that anneals within the *LEU2* locus (Fig. 26 panel A).

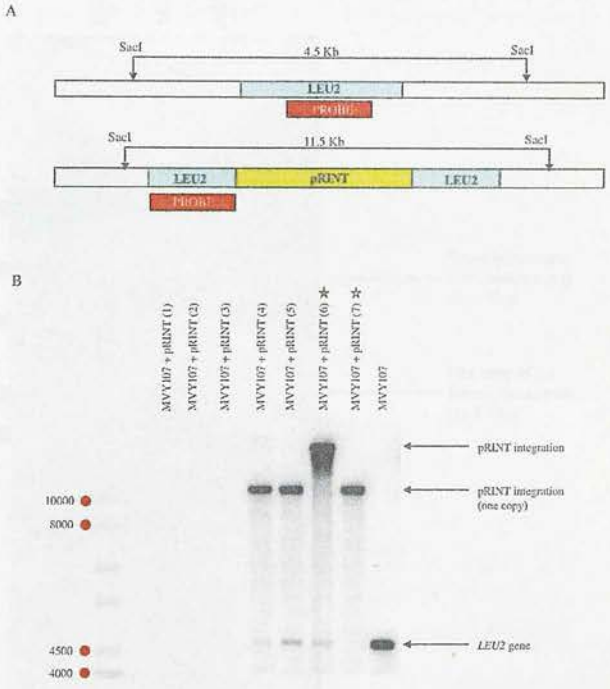


Fig.26 Integration of the R gene into the *S. cerevisiae* genome. A) Map of Chr. II where the *LEU2* gene is located. The integration of the pRINT fragment into the endogenous *LEU2* locus was confirmed by Southern analysis. DNA from transformants was restricted with SacI. The size of the fragments and the position of the probe are indicated on the map. B) Southern blot; M:1kbp plus DNA MW ladder. The arrows point to the different restriction products. Of the positive clones those loaded in lanes 6 and 7 (indicated with an asterisk) were chosen for future analyses. The strain loaded in lane 6 contains more than one copy of the R gene.

This analysis is also essential to determine the number of pRINT copies that are integrated into the *S. cerevisiae* genome. Restriction with SacI allows descrimination between WT and positive clones for pRINT integration. The positive clones (indicated with an asterisk in Fig.26 panel B) named MVY171 (lane 6 on the gel) and MVY172 (lane 7 on the gel) were selected for the following experiments. To better descrminate the exact R gene copy number in MVY171, the DNA preparation from MVY171, MVY172 and MVY107 were separated on a 0.5 % agarose gel (Fig.27). The LEU2 probe was used for the hybridization. This analysis demonstrated that MVY171 and MVY172 contain two copies (18.5 kbp) and one copy (11.5 kbp) of the pRINT fragment respectively (Fig.27).

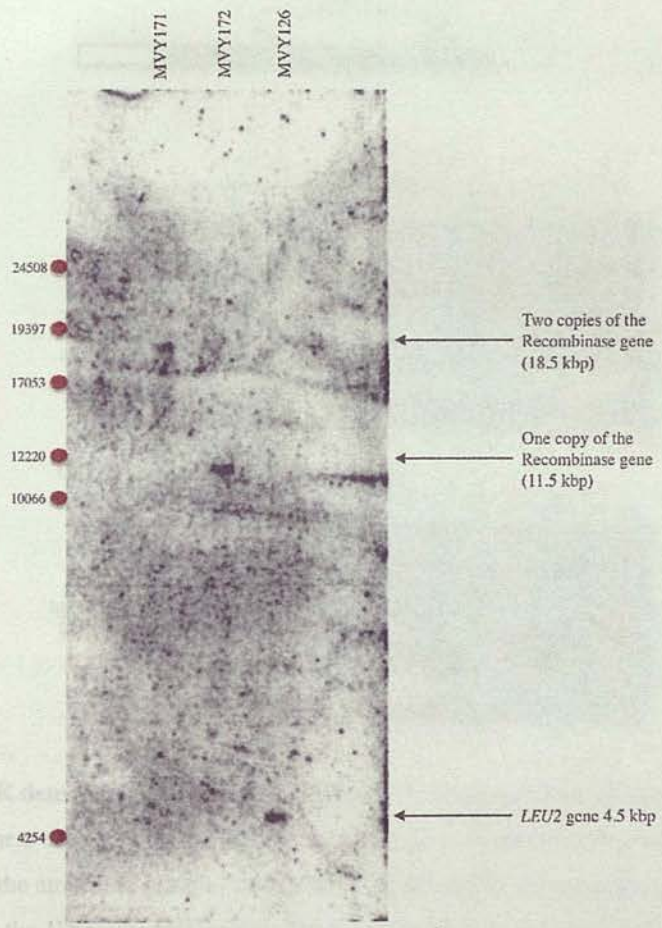


Fig. 27 Southern Blot to determine R gene copy number in MVY171. Samples were loaded on a 0.5% agarose gel; the membrane was hybridized with the LEU2 probe that anneals within the endogenous *LEU2* gene. The hybridization demonstrates that two copies of the pRINT fragment are present in MVY171, as indicated by the upper arrow.

Indeed the band observed in the lane where the MVY171 sample was loaded is of approximately 18 kbp, as expected if only two copies of pRINT have been integrated (Fig.27). Digestion of MVY172 genomi DNA with SacI produced a band of 11 kbp, indicating the presence of only one copy of pRINT (Fig.27).

MVY200 was constructed by MVY171 transformation with the 3X-FLAG-CloNAT fragment that integrates into the 3'-end of *ORC2*. Positive transformants were screened by colony PCR (Fig.28).

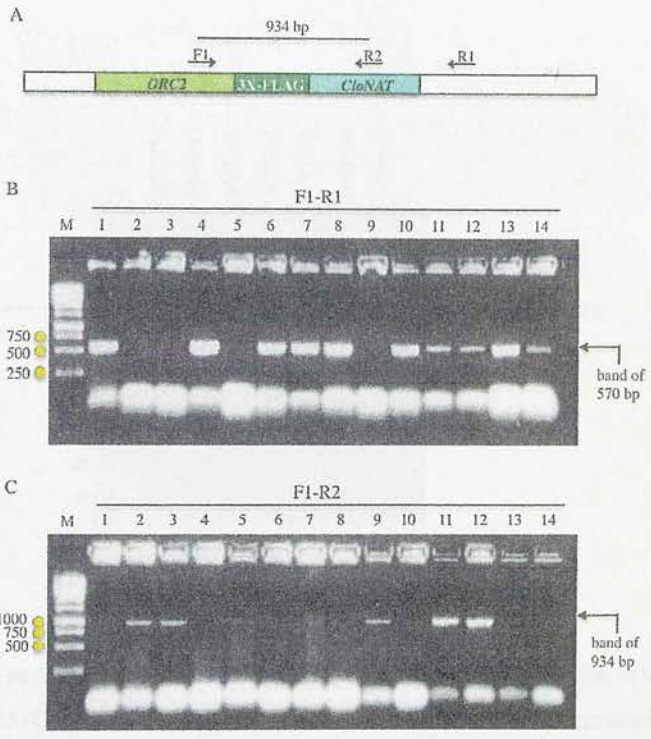


Fig.28 Colony PCR demonstrating the *ORC2* tagging. A) Schematic map of the *ORC2* locus; position of the primers used for colony PCR are indicated. B-C) Pictures of the ehtidium bromide stained gels where the amplified samples were loaded. FI-R1 and F1-R2 indicate the different primer combinations; M is the 1kbp plus MW ladder. The bands in panel B indicated by the arrow are the negative clones. The fragments visualized in panel C are the positive clones; of these only the samples loaded in lanes 2,3,5 and 9 were chosen for the following analyses.

The ORC2FLAGupf and ORC2FLAGdownr (F1 and R1 in Fig.28 A) primers yield an amplification product of 570 bp only if the transforming fragment has not been integrated into the genome. Amplification with NatIntR (R2 in the Fig.28 A) and ORC2FLAGupf produces a 934 bp fragment only in successfully transformed cells. The positive clones (lanes 2, 3, 5 and 9 panel B and C Fig.28) were named MVY197, MVY198, MVY199 and MVY200, following the loading order.
The verified strains were analyzed by western blotting to demonstrate the functional expression of Orc2-3X-FLAGp-(Fig. 29). As expected the α-FLAG antibody specifically recognized a species of the appropriate molecular mass in lysates derived from MVY197-200. The strain MVY200 was chosen for the subsequent analyses.

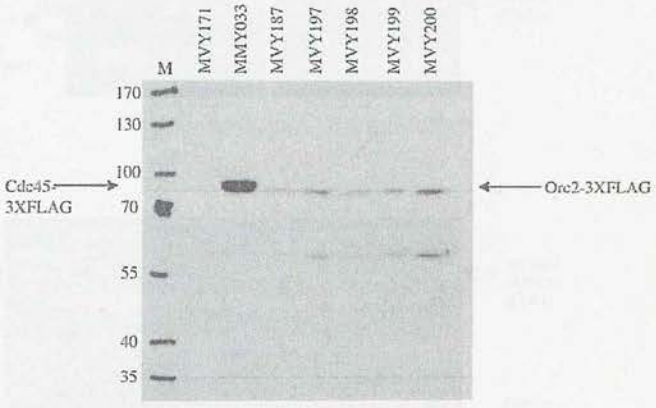


Fig.29 Western Blot to confirm expression of Orc2-3X-FLAGp in MVY197-200. MVY171 (untagged) and MMY033 (Cdc45-3X-FLAG) and MVY187 (Orc2-3X-FLAG) were used as negative and positive controls, respectively. The arrows indicate the Cdc45-3XFLAGp and Orc2-3XFLAGpp species detected by hybridization with αFLAG antibody. All the analyzed strains (MVY197-200) expressed Orc2-3X-FLAGp.

MVY201 and MVY202 were constructed by transformation, to disrupt *BAR1*, of MVY200 and MVY 171. The Hygromycin fragment for *BAR1* disruption PCR amplification of the Hygromycin sequence on pAG32 with Bar1-HygF and Bar1-

HygR oligonucleotides respectively; colonies were screened by colony PCR with Bar1upF and Bar1cIntrR (F1 and R1, Fig.30 A) and Hyg-IntR (R2, Fig.30 A). The sample loaded in the lanes 3 (and 9), named MVY201, was chosen for subsequent analyses. Transformation of MVY200 gave one positive clone (lane 10 panel C Fig.30), named MVY202.

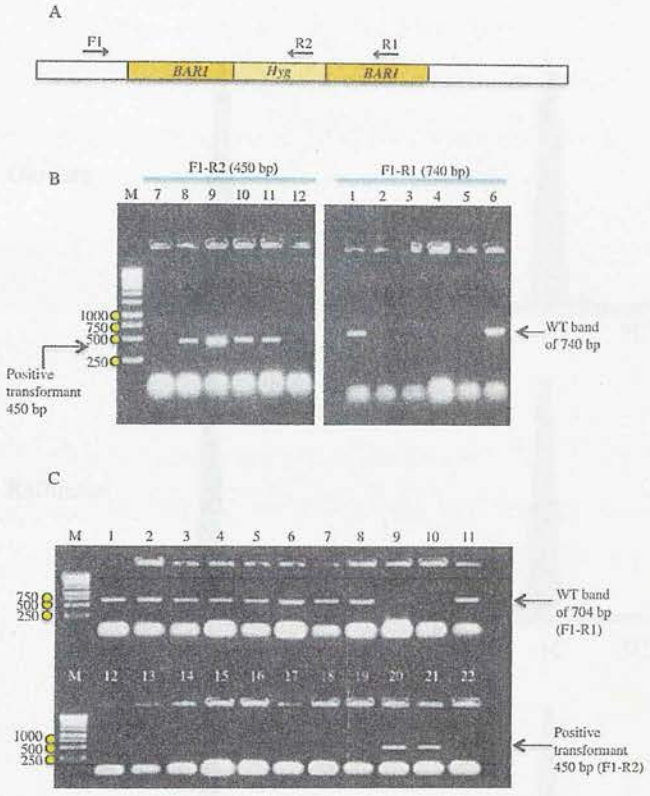


Fig.30 Disruption of the *BAR1* gene. A) Schematic representation of the *BAR1* locus disrupted with *Hyg*. The arrows indicate positions of the primers. B-C) *BAR1* disruption in MVY200 and MVY171 respectively; M is the 1kbp plus DNA MW. B) Lanes 6 and 12 indicate untransformed MVY200; lanes 1-5 and 7-11 are selected colonies amplified with F1-R1 and F1-R2 pairs of primers respectively. C) Lanes 11 and 22 indicate untransformed MVY171. The colonies amplified with F1-R1 and F1-R2 pairs of primers are loaded in lanes 1-10 and 12-21, respectively.

The arrest of cells in G1 is achieved by treatment of the cells with the mating pheromone α-factor (see paragraph 2.34 for details). It is known that disruption of the *BAR1* gene, which encodes for a protease that degrades the pheromone, increases

the sensitivity of the strain to α-factor treatment (Guthrie and Frink, 2002). Therefore to confirm disruption of *BAR1* a Fluorescence Activated Cell Sorting (FACS) analysis was performed on strains MVY201 and MVY202 arrested in G1 with α-factor (Fig.31).

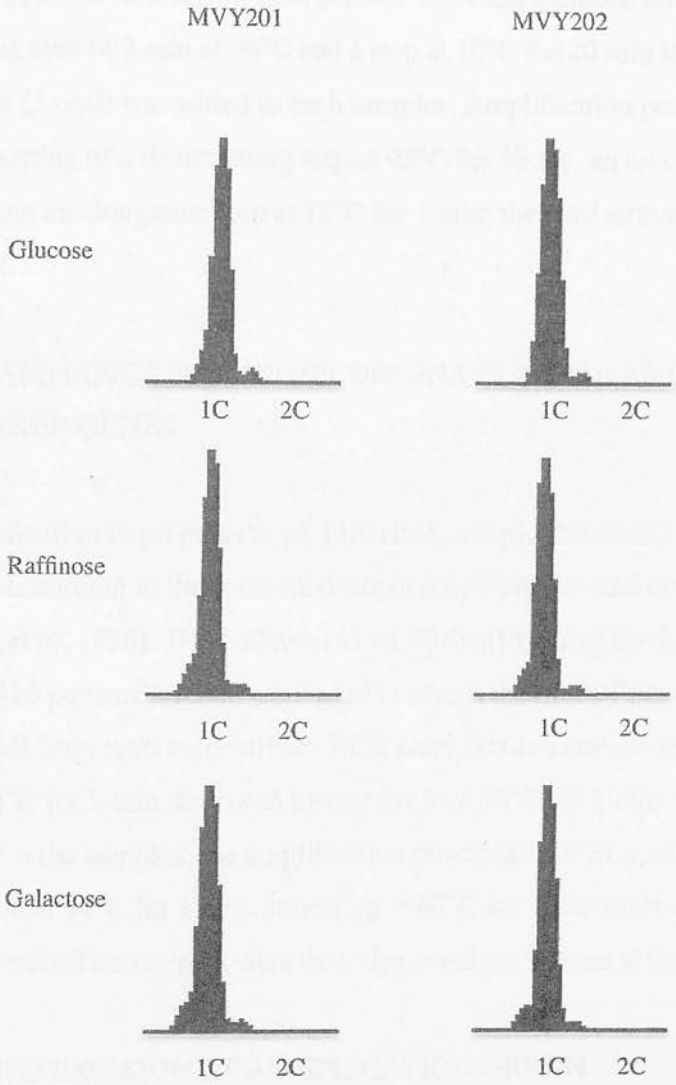


Fig.31. FACS analysis to confirm α-factor arrest after *BAR1* disruption. G1 arrest was performed on constructed MVY201 and MVY202 grown in Glucose, Raffinose and Galactose; FACS analysis of all the samples demonstrates their arrest in G1 after three hours treatment with α-factor.

2.26 PCR AMPLIFICATION OF RECOMBINATION SITE FRAGMENTS

The R siteA and RsiteB fragments were prepared by PCR amplification of pHM401 using the RsiteA F-RsiteA R and RsiteB F-RsiteB R primer combinations respectively. 100 ng of plasmid were mixed with 1X PCR buffer, 2.5 mM MgCl2, 1 mM dNTPs (20 mM), 5 μM of each primer. The PCR included an initial denaturating step of 2 min at 96°C and a step at 10°C for 20 min in which 1μl Pfu Polymerase (3 u/μl) was added to each samples. Amplification proceeded with 30 cycles consisting of a denaturating step at 95°C for 30 sec, an annealing step at 60°C for 1 min and an elongation step at 72°C for 1 min; the final extension step was for 5 min at 72°C.

2.27 PCR AMPLIFICATION OF RECOMBINATION SITE AND LIGATED MARKER GENES

PCR amplification to prepare the pLT1/RsiteA and pLT2/RsiteB fragments was performed according to the protocol described by Longtine and colleagues (Longtine, *et al*, 1998). The LT2Rev1413-LT2Fwd1413 and the LT1Rev1413-LT1Fwd1413 pairs of primers were used to obtain the pLT1/RsiteA and pLT2/RsiteB fragments respectively. PCR amplification consisted of: a denaturating cycle at 94°C for 2 min, followed by one cycle at 80°C for 2 min when the upper mix was added to the samples; the amplification proceeded for 31 cycles consisting of a denaturation at 94°C for 1 min, annealing at 60°C for 45 seconds and elongation at 68°C for 2 min. The samples were then elongated for 10 min at 68°C.

2.28 *IN VIVO* EXCISION OF A REPLICATION ORIGIN

For ARS1413 excision strains were grown to exponential phase and then half of the culture was treated to induce expression of the R gene, while the other half (uninduced) represented the negative control. Galactose was used to induce recombination in strains: MVY107, MVY126, MVY171, MVY172, MVY200,

MVY201; β- estradiol (SIGMA) was used to induce R gene expression in MVY128. Samples were collected at different times to follow the kinetics of excision. ARS1413 excision was confirmed by Southern blot analysis using radiolabelled probes directed against ARS1413 (340 bp, prepared via PCR amplification of genomic DNA from W303-1a, employing the primers ARS1413ProbeF and ARS1413ProbeR) and *TRP1* (1080 bp, prepared via digestion of pLT2/RsiteB with PmeI and SmaI).

2.29 PCR AMPLIFICATION FOR *ORC2* TAGGING

The 3X-FLAG-CloNAT fragment (1.5 kbp) was prepared by PCR amplification of the Flag-Nat plasmid with primers ORC2-FLAG R and ORC2-FLAG F. The PCR reaction mix was prepared as follows: pFLAG-Nat 10 ng, 1X PCR Buffer, 0.1 mg/ml BSA, 1.5 mM $MgCl_2$, 0.8 mM dNTPs, 0.5 mM each primer, 10 mM Tris-HCl pH8, 1μl of Taq/Vent (1:1 ratio); the final volume of the mix was 50 μl. The PCR cycles included an initial denaturating step at 94°C for 2 min, a second step at 10°C for 20 min to add the Taq polymerases, which was followed by 30 cycles of a denaturing step at 94°C for 1min, an annealing step at 55°C for 1min an elongating step at 72°C for 3min, and a final extension at 72°C for 20min. The 1.5 kbp fragment was purified with a PCR clean-up kit (QIAGEN).

2.30 PROTEIN TAGGING

The Orc2 protein was 3X-FLAG tagged by inserting the 3X-FLAG-CloNAT fragment at the 3′ end of the endogenous gene. After yeast transformation the transforming fragment integrated into the genome by homologous recombiantion. The correct integration was checked by colony PCR while the correct expression of the Orc2-3X-FLAGp was confirmed by Western blot (see description of MVY200 in paragraph 2.25)

2.31 CHROMATIN IMMUNOPRECIPITATION

50 ml of cells grown to OD_{600}= 1, were treated with 1% v/v Formaldehyde for 30 min at RT with regular shaking and then with 125 mM of Glycine (pH 2.5). After washing with 1X PBS the pellets were resuspended in 500 µl (for 50 $_{OD600}$ of cells) of FA buffer (50mM HEPES-KOH pH 7.6, 140mM NaCl, 1mM EDTA, 1% v/v TritonX-100) containing 1X protease inhibitor cocktail (Complete EDTA free ROCHE) and transferred to 1.5 ml siliconized eppendorf tubes (Fisher). Cells were mechanically broken by incubation in an eppendorf shaker for 50 minutes at 1500rpm at 4°C after one volume of glass beads was added to the samples. The recovered lysates were sonicated (3 times for 10sec with intervals of 30 sec), centrifuged for 15min at 4°C at 16000 g, transferred to Ultrafree-MC Millipore filter tubes and centrifuged twice as above.

PREPARATION OF IP SAMPLES

The crude lysates (100 µl) were diluted with 100 µl of FA-HS buffer (50mM HEPES-KOH pH 7.6, 500 mM NaCl, 1 mM EDTA, 1% v/v TritonX-100) containing 1X protease inhibitor cocktail and incubated over night at 4°C with regular shaking with 30 µl of αFLAG M2 agarose beads. The beads, recovered after centrifugation at 400 g for 1min at RT, were resuspended in 400 µl of FA-HS buffer (without protease inhibitors) and transferred to 0.45µm Ultrafree-MC Millipore filter tubes; the samples were rotated for 5min at RT and centrifuged for 1min at 2000rpm. The beads were then washed as follow: 3 times with 60 0µl of FA-HS buffer (10min-10min-5min) and once with 600 µl Deoxycholate Buffer (10mM Tris-HCl pH8, 0.25M LiCL, 0.5% v/v NP-40, 0.5% v/v sodium deoxycholate, 1 mM EDTA) (5 min). The protein-DNA complexes were then immunoprecipitated twice by incubating the beads with 100 µl of Deoxycholate buffer containing 0.3 mg/ml of FLAG peptide for 30 min at RT. The samples were treated with 20µl of 10X reverse Crosslink mix (100m Tris-HCl pH8, 40mM EDTA, 1M NaCl) over night at 65°C and then with 4µl Glycogene (20 mg/ml) (Applied Biosystem) and 3 µl ProteinaseK (20 mg/ml) (Applied Biosystem) for 3hours at 56°C. DNA was then recovered by

phenol extraction and precipitation; the air-dried pellets were resuspended in 50 µl of 1X TE pH8.

PREPARATION OF INPUT SAMPLES

20 µl of the crude lysates were treated with 80 µl of DNA Clean Up solution (1% w/v SDS, 1X TE pH8) over night at 65°C and then treated with 94 µl of 1X TE pH8, 1µl Glycogen (20mg/ml) and 5µl ProteinaseK (20mg/ml) for 3 hours at 56°C. The DNA samples were recovered by phenol extraction and precipitation; the dried pellet was resuspended in 700 µl of 1X TE pH8; 1µl of a 1:10 dilution was used for the PCR analysis.

2.32 PCR AMPLIFICATION OF ChIP SAMPLES

The PCR reaction mix was comprised of 1X PCR buffer, $MgCl_2$ 1.5 mM, dNTPS 0.4mM, 1.4 µM each primers (MWG), radioactive α-dATP-P^{32} 1 µCi/µl, and 0.025 U/µl Taq polymerase (Invitrogen). The PCR cycles comprised an initial denaturating step at 94°C for 1 min and 30 sec followed by 23 cycles including a denaturating step at 94°C for 30 sec, an annealing step at 55°C for 30 sec and an elongation step at 72°C for 45 sec and then a final elongation step at 72°C for 5 min.
Samples were loaded on a 7% v/v acrylammide (19:1 acrylamide:bis-acrylamide) gel and separated by electrophoresis for 45 min at 150V; the radioactive signal was acquired by analysis with the STORM phosphorimager.

2.33 PLASMID ISOLATION

Yeast cells were grown asynchronously in 3 L of YP 3% Raffinose and then Galactose was added to the culture to a final concentration of 2% v/v to the culture to induce the excision of ARS1413. After 5 hours of induction the cells, grown to late log phase (OD_{600} between 1.5-2), were collected and centrifuged for 5 min at 1100 rpm at 4°C. The cells were treated to separate the chromatin plasmid obtained after excision of ARS1413 from chromosome XIV from the rest of the genome. The ARS1413 plasmid was purified taking advantage of the technique published by Kim

and colleagues (Kim, *et al*, 2004). Modification according to the specific requirements of the experiment (described in paragraph 4.5), have been applied to the protocol. The main steps of the protocol are: formation of spheroplasts, preparation of nuclei and isolation of the chromatin plasmid.

PREPARATION OF SPHEROPLASTS

The pellet from 1 litre of culture was thawed in 50ml of Spheroplasting Medium (SM) (1M Sorbitol, 50 mM Tris-HCl pH8, 20 mM β-mercaptoethanol) and warmed at 30°C for 15 min with regular swirling. Formation of spheroplasts was induced by treatment of each sample with 20 ml of SM, in which 0.022 gr of lytic enzyme (Zymolyase-100T ImmunO Lytic Enzyme 20000 U/gr) had been dissolved. Samples were incubated at 30°C for up to 1 hour until formation of spheroplasts was completed; the time course of spheroplasting was followed by measuring the OD_{600} of an aliquot (40 µl) of cells in 1% SDS and it was considered completed when the OD_{600} value was reduced to less than 10% of the starting value. Cells were then transferred to two polycarbonate high speed tubes, spun in a JA-25 rotor at 6416 g for 5 min at 4°C and washed with 25 ml of cold ST buffer (1M Sorbitol, 50mM Tris-HCl pH8) to remove the lytic enzyme.

PREPARATION OF NUCLEI

The spheroplast pellets were resuspended on ice in 5ml of F-Buffer (18% Ficoll-400, 40 mM Potassium Phosphate pH 6.5, 1 mM EDTA) with reducing agents and protease inhibitors (5mM 2β-mercaptoethanol, 0.1 mM AE-BSF, 5µg/ml Leupeptin, 15 µg/ml Pepsatin A) added just before use. Samples were diluted with other 15ml of F-buffer and layered carefully on top of 15 ml of FG-buffer (7% Ficoll-400, 40 mM Potassium Phosphate pH 6.5, 1 mM EDTA, 20% Glycerol); the gradients were then spun at 22300 g for 1hour at 4°C in a JA-25 rotor. The supernatant was decanted and the pellet represents the crude nuclei preparation.

ISOLATION OF PLASMID FROM NUCLEI

Nuclei were resuspended in 4ml of Nuclear Lysis Buffer (50 mM Tris-HCl pH 8, 5 mM EDTA) with 5mM 2β-mercaptoethanol and protease inhibitors as above and 40

μl of RNaseA 100 mg/ml (Qiagen; DNAse free) and then incubated on ice for 30min. Samples were spun at 11400 g for 5 min at 4°C in a JA-25 rotor to eliminate the nuclear debris and the recovered supernatant contained the chromatin plasmid. A sucrose cushion of 850 μl, (30% sucrose, 40mM Tris adjusted to pH 7.9 with acetic acid, 2 mM EDTA) with 5 mM 2β-mercaptoethanol and protease inhibitors as above, was placed in Beckman Sw55 polyallomer tubes; the lysate was applied slowly on the top of the cushion and samples spun in a Beckman Sw55Ti rotor for 4.5 hours at 367000 g at 4° C; the high speed centrifugation deposits and concentrates the chromatin plasmid into the cushion. The supernatant was removed and the sucrose cushion used to resuspend any pelleted material; both the supernatant and sucrose samples were then transferred to cold falcon tubes and stored at -80°C.
An aliquot (500 μl) of the sucrose cushion and the supernatant was treated with the Qiagen kit for plasmid isolation.

2.34 MASS SPECTROMETRY ANALYSIS OF THE ISOLATED CHROMATIN PLASMID

Sample preparation and mass spectrometry analysis were performed by Thierry Le Bihan (CSBE University of Edinburgh). The sucrose cushion recovered after plasmid isolation (approximately 5 ml) was incubated for 30 min with 1ml Acetonitrile (Fisher), 50 μl 1M Ammonium Bicarbonate, 100 μl 200 mM DTT; peptide digestion was performed over night at Room Temeprature (RT) by incubation with 100 μl 500 mM Iodoacetamide and 10 μg of Trypsin (Sequencing grade modified porcine trypsin (Promega)). Samples were cleaned on Seppak Columns (H_2O Corp) according to the manufacturer's instructions; the extract was diluted to 300 μl with H_2O and protease inhibitors (Roche complete EDTA free) were added to the sample. The sample was treated with 125 μl 8M Urea, 25 μl 1M ammonium bicarbonate, 25 μl 200 mM DTT and incubated for 30 min at Room Temperature; 25 μl 500 mM Iodoacetamide and 20 μg of Trypsin were added to the sample and digestion proceeded over night at RT. The samples were then stored at -80°C in aliquots. Capillary Picotip columns (10cm X 360 μm OD X 75 μm ID) with a 15 μm tip

opening fitted with a borosilicate frit (New Objective) and Pursuit C18 5 μm (Variant) were used for sample preparation. For HPLC and Mass Spectrometry analyses a binary HPLC system coupled to a hybrid LTQ-Orbitrap XL instrument were used; spectra with less than ten fragments were discarded. Protein identification was performed using MASCOT Version 2.2 against the SGD yeast database.

2.35 ARREST OF CELLS WITH α-FACTOR AND RELEASE INTO S-PHASE

An over night cell culture was diluted into YPD to a density of OD_{600}= 0.15; when the culture reached a concentration of approximately OD_{60}= 0.4-0.45 α-factor was added to the culture to a final concentration of 1μg/ml and cells were incubated at 30°C for approximately 3 hours. Cell cycle arrest was confirmed by microscope analysis. Release into S-phase was performed by pelletting and resuspending the cells in fresh YPD media containing 50 μg/ml Pronase.

2.36 ARREST OF CELLS WITH HU

Cells grown to exponential phase in YPD were washed twice with sterile distilled H_2O and resuspende in prewarmed YPD containing 200 mM HU; cultures were then incubated at 30°C for two hours to induce S-phase arrest due to checkpoint activation. Arrest of cells in S-phase was confirmed by FACS analysis.

2.37 FACS ANALYSIS

2.5 ml of cells arrested in S-phase with HU, or in G1 with α-factor and then released into S-phase were centrifuged (Beckman centrifuge) for 4 min at 4500 g at 4°C; the pellet was resuspended in 1.5 ml of H_2O and 3 ml of ethanol and then incubated over night at 4°C with regular agitation. Cells were centrifuged for 4 min at 4500 g at 4°C, washed twice with 5 ml of 50 mM Na citrate (Zou, *et al*, 1997) and then once for 30 sec on medium frequency (Bioruptor sonicator). The samples were centrifuged for 4 min at 1500 rpm at 4°C, resuspended in 1ml 50 mM Na citrate containing 1

μg/μl RNaseA and incubated at 37°C for 5 hours. After over night incubation at 4°C cells were resuspend in 1 ml 50 mM Na citrate, centrifuged and then 250 μl aliquots transferred to 1.5 ml eppendorf tubes. The samples were incubated over night at RT with regular agitation after 250 μl 50 mM Na citrate containing 2 μM Sytox Green (Invitrogen) were added. Cells were sonicated twice for 10 seconds, 200 μl aliquots transferred to 5ml round bottom falcon tubes and 1 ml 50 mM NaCitrate containing 1 μM Sytox Green added. The samples were analyzed using the FACS machine (Becton Dickenson FACS Analyser).

2.38 CHROMATIN FRACTIONATION

The chromatin fractionation for the Cdc45p immunoprecipitation was performed according to the protocol described in Sheu and Stillman, 2006 (Fig.32)

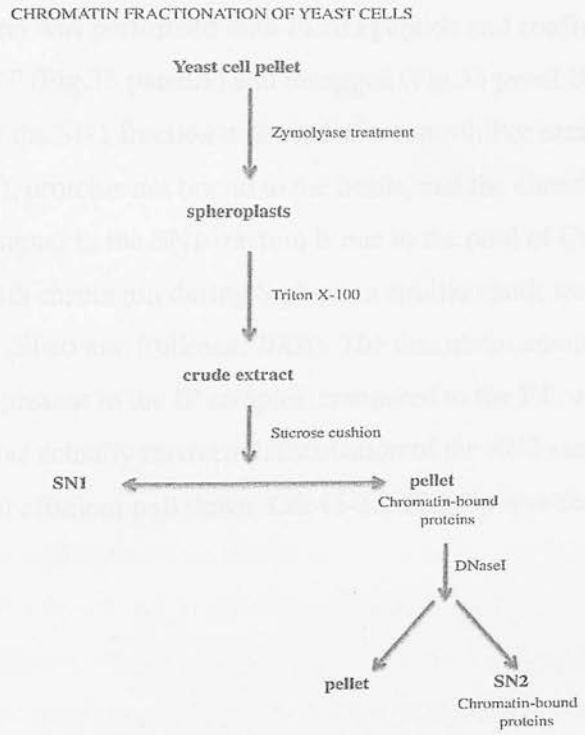


Fig.32 Outline of the chromatin fractionation experiment (adapted from Sheu and Stillman, 2006). Crude extract is recovered by TritonX-100 incubation of yeast spheroplasts. The SN1 fraction, containing proteins not bound to chromatin is separated through a sucrose cushion centrifugation. The recovered pellet is then incubated with DNaseI promoting release of chromatin bound proteins (SN2).

2.39 Cdc45-3X-FLAGp IMMUNOPRECIPITATION WITH M2-αFLAG-AGAROSE BEADS

Cdc45p immunoprecipitation was performed on the protein bound fraction (SN2) recovered from HU arrested cells according to the protocol described by Sheu and Stillman, 2006. Preliminary experiments were performed to optimize the Cdc45-3X-FLAGp immunoprecipitation and elutions steps (2.40 and 2.41 paragraphs respectively) in order to recover enough material for the mass spectrometry analysis.

2.40 OPTIMISING CONDITIONS FOR THE IMMUNOPRECIPITATION STEP

To find the amount of M2-αFLAG-agarose beads that allowed a selective immnunoprecipitation of Cdc45-3X-FLAGp, the SN2 fractions (500 µl) were incubated with different volumes of beads: 10 µl, 20 µl, 40 µl and 80 µl. Cdc45-3X-FLAGp elution was performed with FLAG peptide and confirmed by western blotting in WT (Fig.33 panelA) and untagged (Fig.33 panel B) strains.
An aliquot of the SN1 fraction was used as a control. For each sample the Flow Through (FT), proteins not bound to the beads, and the eluted proteins (IP) were loaded. The signal in the SN1 fraction is due to the pool of Cdc45p that is not associated with chromatin during S-phase; a similar result was also obtained by Sheu and Stillman (Sheu and Stillman, 2006). The maximum amount of Cdc45-3X-FLAGp was present in the IP samples, compared to the FT, indicating that most of the protein was actually recovered. Incubation of the SN2 sample with 10µl of beads gave the most efficient pull down. Cdc45-3X-FLAGp was detected only in the WT samples.

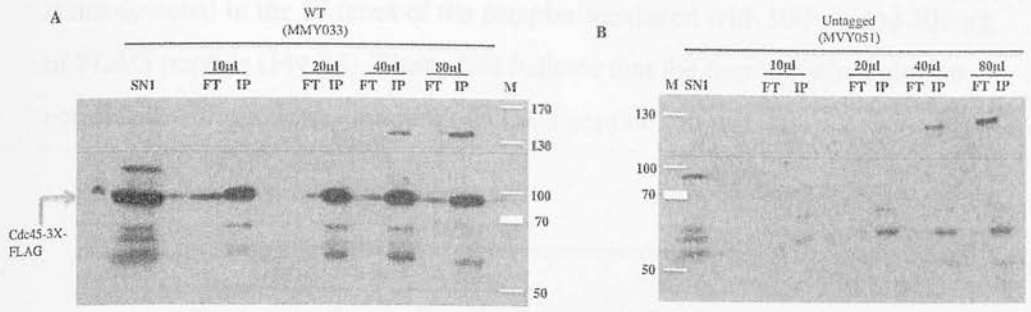


Fig.33 Western Blot to optimize conditions for Cdc45p immunoprecipitation. WT (panel A) and untagged (panel B) samples prepared from cells arrested in S-phase with HU and incubated with different amounts of M2-αFLAG-agarose beads; Cdc45-3X-FLAGp (74 kDa) was detected by hybridization with αFLAG antibody. SN1: proteins not bound to chromatin. FT (Flow Through): proteins not bound to the beads. IP: eluted proteins. Samples were grouped according to the different amounts of beads used: 10 µl, 20 µl, 40 µl and 80 µl. Cdc45-3X-FLAG is only detected in the WT sample. The best pull down is observed in the sample incubated with 10µl of beads.

2.41 OPTIMISING CONDITIONS FOR CDC45-3X-FLAGp ELUTION

In order to optimise the condition for Cdc45-3-FLAGp immunoprecipitation, experiments were performed to adjust the amount of FLAG peptide used for the elution step. Thus the SN2 fraction from MMY033 and MVY51 was incubated with M2-αFLAG-agarose beads according to the condition previously determined (paragraph 2.39): 10 µl of beads for 0.5 ml of lysate. The samples were then incubated with different amount of FLAG peptide (25 mg/ml), 50 µg, 100 µg and 400 µg, for 30 minutes at RT. After centrifugation for 2 minutes at 2400 g, the eluted sample was transferred to a new eppendof tube and a second elution was performed. The two eluted samples were then pooled and analyzed by Western blotting (Fig.34). Samples are grouped according to the amount of FLAG peptide: 50 µg,100 µg and 400 µg. For each group FT, IP and B (proteins that remained bound to the beads) samples have been loaded. Immunoprecipitation of Cdc45-3X-FLAGp occurred only in the sample treated with 50 µg of FLAG peptide (Fig.34). The Cdc45-3X-FLAGp

is not detected in the IP lanes of the samples incubated with 100 µg and 400 µg of FLAG peptide (Fig.34). These data indicate that the more efficient elution correlated with the lower amount of FLAG peptide (50 µg).

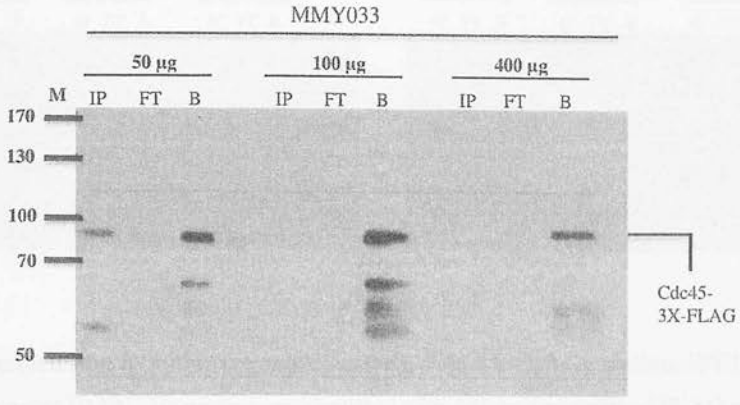


Fig.34 Western blot to set up the FLAG peptide amounts for Cdc45-3X-FLAGp elution. WT samples treated with HU and incubated with different amounts of FLAG peptide: 50 µg,100 µg and 400 µg. Cdc45-3X-FLAGp (80 kDa) from cells was detected by hybridization with αFLAG antibody. For each group FT, IP and B samples have been loaded. Sizes of protein ladder (M) bands are indicated on the side of the gel (kDa); the arrow points to the Cdc45-3X-FLAGp band. A signal is present in the B lane of each samples; a faint band for Cdc45-3X-FLAGp can only be detected in the IP of the sample incubated with 50 µg of FLAG peptide. These results demonstrate that high amounts of FLAG peptide impaired Cdc45-3X-FLAGp pull down.

Considering the results described above (Fig.34) a second elution was performed by incubating the lysate with 5 µg and 10 µg of FLAG peptide; as a control an aliquot of lysate was not treated with peptide. Samples from MMY033 and MMY021 were loaded on separate gels, panel A and B (Fig.35) respectively. No proteins could be detected in the IP lane from the lysate that was not incubated with the peptide. The most efficient elution was obtained by treating the lysate with 10µg of peptide. No Cdc45-3X-FLAGp band was detected in any lanes of the untagged strain.

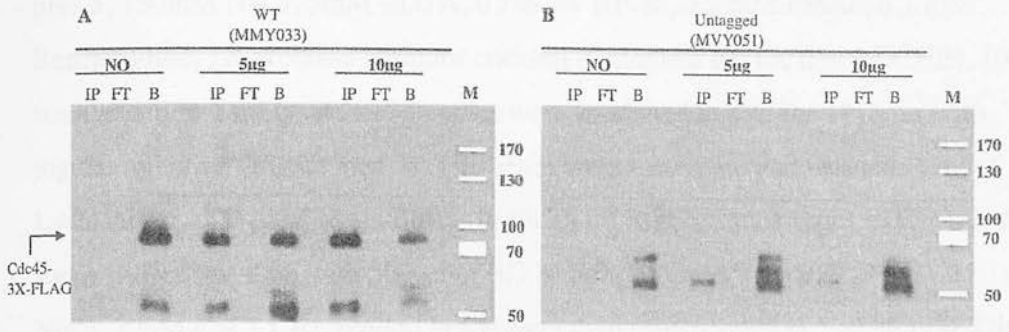


Fig.35 Western Blot to optimize conditions for Cdc45-3X-FLAGp elution. WT (panel A) and untagged (panel B) SN2 fractions were incubated with different amounts of FLAG peptide. Cdc45-3X-FLAGp was detected by hybridization with αFLAG antibody. Samples are grouped according to the amount of FLAG peptide used: NO (no FLAG peptide), 5 µg and 10µg of peptide. For each group FT, IP and B samples were loaded. Cdc45-3X-FLAGp is only detected in the WT samples, as indicated by the arrow (panel A). No signal is detected in the untagged strain or in the WT strain that was not incubated with FLAG peptide (NO) demonstrating the specificity of the elution for Cdc45-3X-FLAGp The best result was obtained by treatment with 10 µg of peptide.

2.42 Cdc45-3X-FLAGp IMMUNOPRECIPITATION WITH PROTEIN G DYNABEADS AND TRYPSIN DIGESTION

The experiment was performed on WT, *Δrpd3* and untagged strains arrested in HU for two hours (Fig.36 step 1). Another WT culture was arrested in G1 with α-factor (cell cycle arrest was confirmed by microscopy) and then released into S-phase (Fig.36 step 1). Aliquots of cells were collected from the four samples to demonstrate, by FACS analysis (Fig.69), that the cells were effectively arrested in S-phase and released into S-phase from G1 arrest (Fig.36 step 2). The SN2 fraction (Fig36. step 3) was transferred to 0.45 µm Ultrafree-MC Millipore filter tubes, centrifuged at 4°C for 15 min at 16000 g and stored over night at 4°C. Protein G Dynabeads (Invitrogen), 4 µl for each 1 ml of lysate, were washed with 1 ml 1X

PBS, resuspended in 0.5 ml 1X PBS containing 8 µl/1ml lysate of α-FLAG antibody and incubated at 4°C with regular agitation. The affinity beads were then washed with 1ml 1X PBS and then with 1ml Lysis Buffer (50mM HEPES-NaOH pH7.5, 150mM NaCl, 5mM EDTA, 0.1% v/v NP-40, 0.2 mM PMSF, 0.2 mM Benzamidine, 1X protease inhibitor cocktail (Complete EDTA free ROCHE). 10 µl were added to 1 ml lysate and samples were incubated at 4°C for $1^{1/2}$ hour with regular agitation (Fig36. step 4). The beads were recovered and washed: 1 ml of Lysis buffer, 500 µl of Lysis Buffer (with 2 mM $MgCl_2$ and without EDTA), 1 ml of Lysis Buffer and then with 500 µl of BD buffer (200 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM $CaCl_2$). Recovered beads were then resuspended in 5 µl 20 mM Tris-HCl pH 8.0; 2.5 µl of Trypsin solution (100 ng/µl trypsin powder resuspended in 20 mM Tris-HCl pH 8.0 (Trypsin, proteomic grade, Sigma)) were added and samples were incubated at 37°C for 4 hours with regular agitation (Fig36. step 5). The recovered supernatant was incubated over night at 37°C after 2.5 µl of Trypsin solution were added. The supernatant was transferred to new eppendorfs and treated with 50 µl of 0.1% v/vTFA/H_2O to acidify peptides (Fig36. step 6). The protein samples were cleaned-up and concentrated using stage tips, which consist of chomatographic beads immobilized in a Teflon meshwork (Rappsilber, *et al*, 2003). Stage tips were prepared by cutting C-18 filters of the appropriate size and placing them into 200 µl tips, 1mm from the tip-end. Filters were moistened with 5 µl methanol and equilibrated with 10 µl 0.1% v/v TFA/2% v/v acetonitrile H_2O. Samples were loaded onto the filters, which were then washed with 20 µl 0.1% v/v TFA/2% v/v acetonitrile H_2O. The samples were and analyzed by mass spectrometry using a LTQ-Orbitrap machine in collaboration with Juri Rappsilber. Peptide identification was performed using the SGD yeast database.

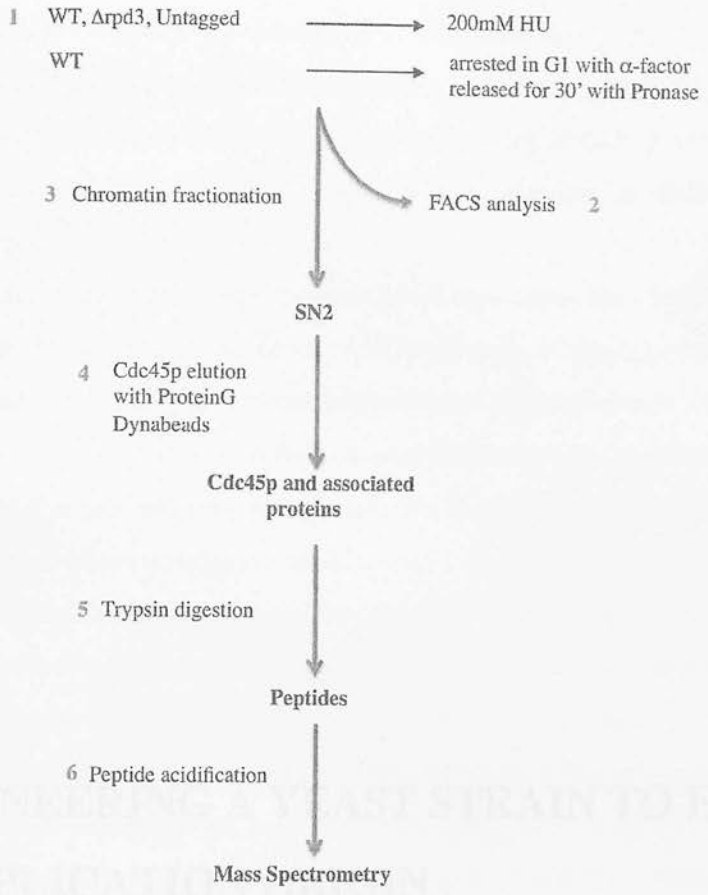


Fig.36 Schematic representation of the adapted immunoprecipitation experiment. The three strains are grown in different conditions as indicated (step1). FACS analysis, on the cultures, confirms that proteins are recovered from cells in S-phase (step2). The SN2 fraction is recovered using the chromatin fractionation protocol (step3); Cdc45p is immunoprecipitated with αFLAG antibody and ProteinG Dynabeads (step4); proteins sample are digested with Trypsin (step5). The recovered peptides are acidified and analyzed by mass spectrometry.

2.43 SILVER STAINING

Silver staining was performed with the PluOne Silver Staining Kit, Protein (GE Healthcare) and according to the manufacturers' instruction.

3 ENGINEERING A YEAST STRAIN TO EXCISE A REPLICATION ORIGIN

3.1 AIM OF THE PROJECT

DNA replication is a tightly regulated process to ensure the genome is completely and accurately duplicated before cell division so that each daughter cell will receive an entire chromosomal complement. Considering the importance of DNA replication and its implication in cell viability it is essential to understand the mechanisms that regulate this process.

During the past years our knowledge about DNA replication has improved due to the identification of new proteins involved in different steps of this process. Despite these discoveries it is likely that many others, which might associate with replication origins during the assembly of pre-RCs or once they have fired, have escaped identification. The proteins that regulate late origin activation or inhibition in response to the S-phase checkpoint are also under investigation.

Each phase of DNA replication is strictly controlled by the action of multiple factors that are required to guarantee that genome duplication occurs as a series of coordinated events and only once per cell cycle. Moreover it has been demonstrated that the chromatin context has a role in determining the timing of origin activation. In fact, both post-transcriptional histone modifications and changes in nucleosome localization, due to chromatin remodeling complexes, can influence the efficiency of replication origins as well as their time of firing. Therefore, due to the presence of these redundant mechanisms, the identification of new proteins through genetic means has become particularly difficult.

It can be appreciated how important it is to identify the proteins that are involved in the DNA replication process and also to understand how origin firing is affected by the chromosomal context. Therefore the aim of my project was to develop an unbiased assay to identify, *in vivo*, the proteome of a eukaryotic replication origin using *S. cerevisiae* as a model organism. For this purpose I took advantage of a recombination system present in the yeast, *Zygosaccharomyces rouxii*, which has already been used to study the replication timing program in budding yeast (Raghuraman, *et al.*, 1997). The idea of my project was, instead, to excise *in vivo* a replication origin from its chromosomal context. The obtained chromatin circle

containing the replication origin of interest would then be isolated from the rest of the genome and analyzed by mass spectrometry to identify the proteins associated with the origin.

This strategy represents a new biological tool to identify novel factors, which are difficult to discover with genetic approaches, involved in DNA replication. Therefore it will have a pivotal role in the improvement of our knowledge regarding one of the most important processes for cell survival.

3.2 THE pSR1 SITE-SPECIFIC RECOMBINATION SYSTEM

My project was focused on the study of DNA replication in the budding yeast *S. cerevisiae* and in particular on the discovery of new proteins required in this cellular process. To increase our knowledge about the factors involved in genome duplication I developed an unbiased assay to study, *in vivo*, the proteomic of eukaryotic replication origins in *S. cerevisiae*. The first part of the strategy, which was chosen to accomplish this task, consisted of the excision of an origin of replication from its chromosomal context. For this purpose I decided to take advantage of the recombination system present in the yeast *Zygosaccharomyces rouxii*. This organism contains a 6 kbp plasmid, called pSR1 (Fig.37), which has a molecular architecture similar to that of the 2 micron (2μ) plasmid generally present in many strains of *S. cerevisiae* (Matsuzaki, *et al*, 1988). Despite this structural similarity, pSR1 does not share sequence homology with the 2μ plasmid, nor with any other region of the budding yeast genome and it can replicate well in host organisms, such as the budding yeast *S. cerevisiae*. The architecture of pSR1 consists of two pairs of recombination sites (IRs) (Matsuzaki, *et al*, 1988) and three different open reading frames, R, S and P; the gene indicated with the R encodes for a Recombinase protein. This plasmid therefore encodes for a site-specific recombination system. Studies in the *S. cerevisiae* host have demonstrated that a 58 bp region, conserved in both recombination sites, was delineated as the essential region for intramolecular recombination of pSR1 (Matsuzaki, *et al*, 1988). Each 58 bp region contains 12 bp recombination sites (Recombination site: R site), separated by 7 bp, which are specifically recognized by the Recombinase (R) protein. This recombination system has two important hallmarks: it relies only on the R protein, which can catalyze recombination in the absence of an energy supply (Araki, *et al*, 1992) and it is completely independent from the Rad52p recombination system present in the host organism (Matsuzaki, *et al*, 1988). Further studies in budding yeast have demonstrated that site-specific recombination of pSR1 represents a powerful system in recombinant DNA technology (Matsuzaki, *et al*, 1990).

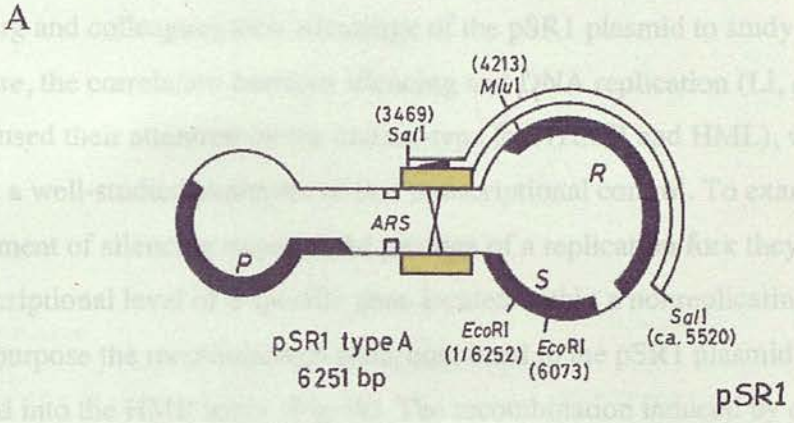


Fig.37 Structure of plasmid pSR1 present in the yeast *Z.rouxii* (Matsuzaki, *et al*, 1988). The yellow blocks indicates the region where the 12 bp recombination sites are located. The site specific recombination depends on the Recombinase gene encoded by the R gene,

Indeed chromosome rearrangements can be obtained by insertion of two RSs in different genomic regions. Moreover the R protein system yields both deletion and inversion according to whether the R sites are in the same or in the opposite orientations (Araki, *et al,* 1992).

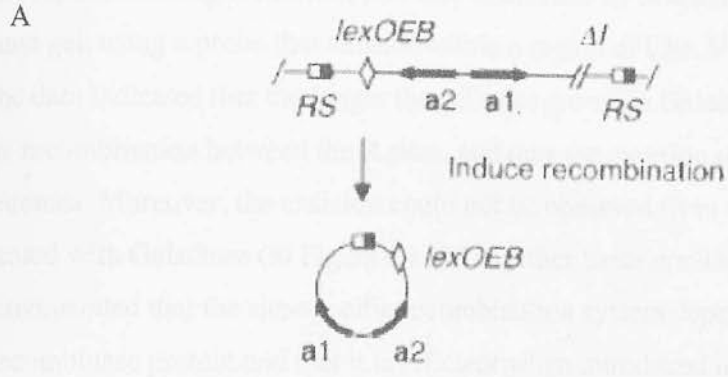
3.3 PREVIOUS STUDIES USING THE pSR1 SITE-SPECIFIC RECOMBINATION SYSTEM

Because of these features the pSR1 system has previously been used by other groups to study different cellular processes.

Gartenberg and colleagues took advantage of the pSR1 plasmid to study, in *S. cerevisiase*, the correlation between silencing and DNA replication (Li, *et al*, 2001). They focused their attention on the mating-type loci (HMR and HML), which represent a well-studied examples of this transcriptional control. To examine if the establishment of silencing required the passage of a replication fork they analyzed the transcriptional level of a specific gene located within a nonreplicating DNA ring. For this purpose the recombination sites, contained in the pSR1 plasmid, were integrated into the HMR locus (Fig.38). The recombination induced by expression of the R gene promoted the excision of the genomic region comprised between the recombination sites as a DNA ring (Fig.38) (Li, *et al*, 2001).

A similar approach has been used to analyze the temporal program of a eukaryotic replication origin in budding yeast (Raghuraman, *et al*, 1997). In this study Raghuraman and colleagues have integrated the pSR1 recombination sites into the flanking region of the subtelomeric late origin ARS501, localized on Chr. V (Fig.38 B). The engineered yeast strain was grown in Galactose to induce the expression of the R gene that is under the Gal promoter; cells were collected hourly during the induction. Upon induction of recombination a 30 kbp cassette, comprised of the sequence between the R sites, was excised obtaining a chromosomal circle.

To confirm that the excision had occurred the chromosomes were separated on a Clamped Homogeneous Electric Field (CHEF) gel in order to visualize shortening of Chr. V.



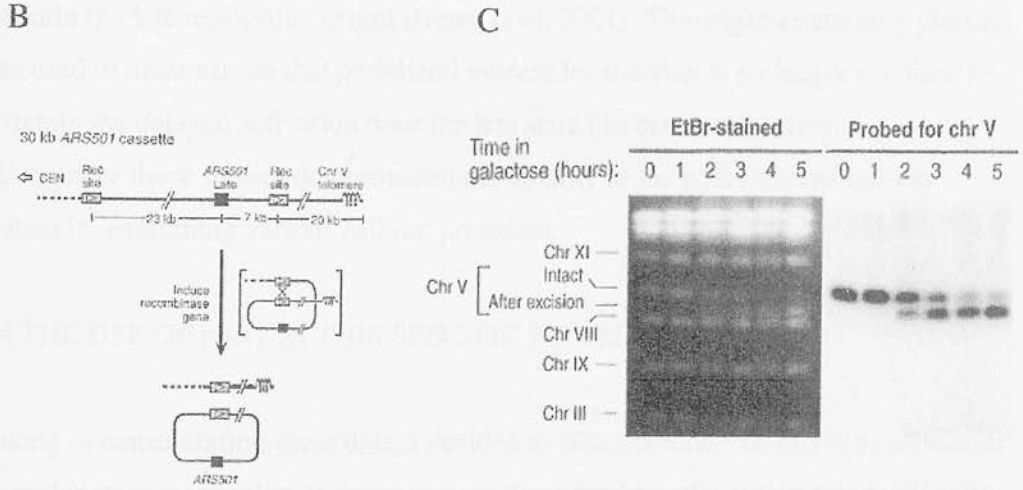


Fig.38 Previous studies demonstrating the efficiency of the site-specific recombination system (Li, *et al*, 2001; Raghuraman, *et al*, 1997). A)The same recombination system was used by Gartenberg and colleagues to excise the mating-type locus HML in order to study the relationship between DNA replication and silencing (Li, *et al*, 2001). The figure shows the DNA ring obtained upon recombination between the recombination sites. **B-C**) Raghuraman group experiments. (Raghuraman, *et al*, 1997). B: Map of the engineered ChrV showing where the recombination sites have been integrated. C: Pictures of the Ethidium Bromide stained agarose gel and the corresponding Southern blot performed to confirm the excision of ARS305.

On the picture of the ethidium bromide stained gel (Fig.38 C) the appearance of a lower chromosomal band could be observed. It corresponds to the shorter size of Chr. V that could be obtained only from the excision of the 30 kbp.
The replication origin excision was also confirmed by Southern blot analysis, on the same gel, using a probe that anneals within a region of Chr. V outside the cassette. The data indicated that the longer the cells are grown in Galactose the more efficient the recombination between the R sites, and thus the excision of the replication origin, becomes. Moreover, the excision could not be observed from cells that were not treated with Galactose (t0 Fig.38 C). All together these previous studies demonstrated that the site-specific recombination system depends exclusively on its Recombinase protein and that it is efficient when introduced into a host organism.
Similarly to the strategy adopted by Raghuraman and colleagues, the recombination sites were integrated into the flanking regions of ARS501 to excise a 30 kbp cassete harborin the late replication origin (Heun, *et al,* 2001). The origin-containing plasmid was used to demonstrate that peripheral nuclear localization is no longer required to maintain the delayed activation once the late state has been established.
All together these studies demonstrated the validity of the pSR1 recombination system in examining various cellular processes.

3.4 THE USE OF pSR1 IN THIS SPECIFIC PROJECT

Taking in consideration these data, I decided to take advantage of this intramolecular recombination mechanism in order to study the proteome of a eukaryotic replication origin. The idea of the presented project was to introduce the recombination sequences from the pSR1 plasmid on either side of a chosen replication origin (Fig.39). In this way induction of recombination would lead to the excision of the genomic region contained between the recombination sites; the resulting chromatin plasmid contains the chosen ARS.
In contrast to the previously described experiments the recombination sites will be integrated in close proximity of the replication origin. In this way only a small genomic region will be excised (approximately 1.3 kbp).

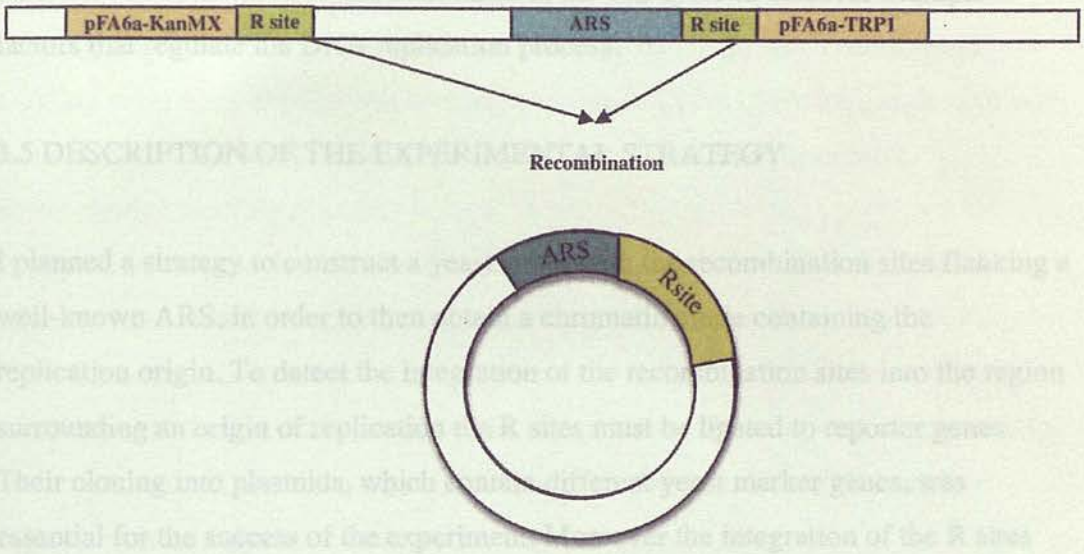


Fig.39 Representation of a late origin flanked by the Recombination Sequences (R site). Each R site is ligated to a reporter gene for selection of its insertion into the genome. The plasmid is the ARS1413 episome produced upon recombination between the recombination sites.

This is required to minimize the presence of external genomic region in order to concentrate the analysis exclusively on the origin of replication. Indeed different requirements had to be accomplished in order to obtain a chromatin plasmid containing only the chosen ARS.

Subsequently the excised episomal plasmid would be isolated from the rest of the genome and then analyzed by mass spectrometry to identify the proteins associated with the replication origin.

Thus, the high efficient recombination system will be combined with a more deep analysis to study the proteome of a eukaryotic replication origin. Moreover it must be considered that the R gene is under the control of the Gal promoter. Therefore, the excision can be induced at different times in the cell cycle to discover multiple factors that regulate the DNA replication process.

3.5 DESCRIPTION OF THE EXPERIMENTAL STRATEGY

I planned a strategy to construct a yeast strain with the recombination sites flanking a well-known ARS, in order to then obtain a chromatin circle containing the replication origin. To detect the integration of the recombination sites into the region surrounding an origin of replication the R sites must be ligated to reporter genes. Their cloning into plasmids, which contain different yeast marker genes, was essential for the success of the experiment. Moreover the integration of the R sites into the genome must satisfy two crucial requirements. Firstly the recombination sites must be inserted in the same orientation in order to promote the excision, and not the inversion, of the replication origin. The second condition is that the produced chromatin plasmid must contain only the replication origin, without the marker genes. To accomplish these necessities the sequences of the recombination sites were amplified in such a way as to obtain two different orientations, named A and B. This system will then allow integration of the recombination sites in the same orientation and at the same time exclusion of the ligated reporter genes from the ARS chromatin plasmid (Fig.39). The different steps required to engineer a yeast strains with the recombination sites in the correct orientation are schematically reported in Fig.40.

The initial step consisted of the amplification, with two different pairs of primers, of the R sites from plasmid pHM401, which derives from pSR1 and contains the R sites, by PCR (Fig.40.1). The primers have been designed in order to obtain two different amplification products named: RsiteA and RsiteB. The DNA fragments obtained were then subcloned into the high copy-number vector, Easy Vector, producing plasmids EV/A and EV/B (Fig. 40.2).

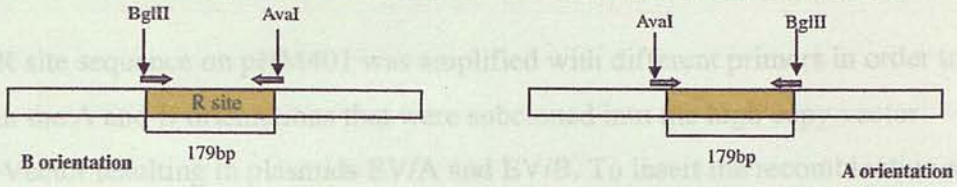
The recombination sites were excised, by restriction digestion, from these constructs and inserted into two plasmids containing different yeast-marker genes to produce constructs pLT1/RsiteA and pLT2/RsiteB (Fig.40.3). These plasmids were then used as templates to amplify DNA fragments containing the R site and the selection marker. The amplified regions will share homology with the genomic sequence, flanking the replication origin, where they have to be integrated. Therefore the budding yeast *S. cerevisiae* was then transformed with the DNA fragments obtained from pLT1/RsiteA and pLT2/RsiteB to insert the R sites into the correct chromosomal loci (Fig.40.4).

It must to be noticed that the constructed pLT1/RsiteA and pLT2/RsiteB plasmids represent two important genetic tools (Mariotti, *et al*, manuscript in preparation). They have been constructed in order to direct the integration of the R site into the yeast genome. Indeed their amplification with primers that share homology with the yeast genome will allow the obtained fragments to be introduce in the required regions by homologous recombination.

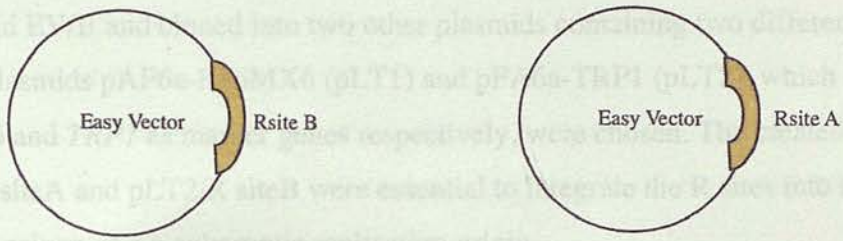
Therefore, by designing the appropriate primers these plasmids can be used to integrate the Rsite in any genomic region of the budding yeast *S.cerevisiae*. Taking in consideration the high efficiency of the site-specific recombination system pLT1/RsiteA and pLT2/RsiteB can be used to excise any sequence of interest within the yeast genome.

Moreover the construction of these plasmids will ensure that the Rsites will always be integrated in the same orientation and both the marker genes will be excluded from the chromatin plasmid upon induction of recombination.

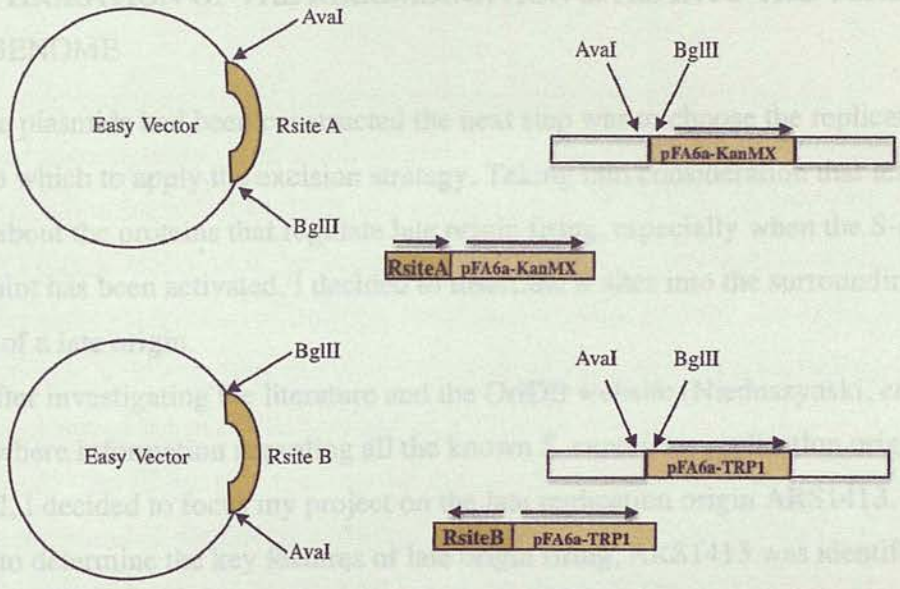
1. Amplification of R site on pHM401 with different primers to have the A orientation and the B orientation



2. Integration od Riste A and Rsite B into the Easy Vector



3. Restriction and ligation of the R sites with the marker genes



4. Yeast transformation to insert the constructs to both sides of the late origin

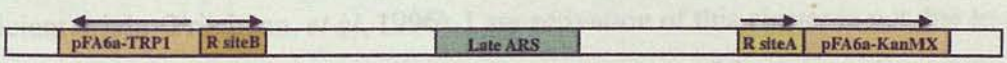


Fig.40 Schematic representation of the designed strategy. The technique used to construct a yeast strain with recombination sites flanking a late replication origin. The different steps have been described in the text.

3.6 RESULTS

3.7 CONSTRUCTION OF PLASMIDS pLT1/RsiteA and pLT2/RsiteB

The R site sequence on pHM401 was amplified with different primers in order to obtain the A and B orientations that were subcloned into the high copy vector EasyVector resulting in plasmids EV/A and EV/B. To insert the recombination sites into the *S.cerevisiae* genome the RsiteA and RsiteB regions were then excised from EV/A and EV/B and cloned into two other plasmids containing two different reporter genes. Plasmids pAF6a-KanMX6 (pLT1) and pFA6a-TRP1 (pLT2), which have *KanMX6* and *TRP1* as marker genes respectively, were chosen. The created plasmids pLT1/R siteA and pLT2/R siteB were essential to integrate the R sites into the flanking reigon of a a eukaryotic replication origin.

3.8 INTEGRATION OF THE RECOMBINATION SITES INTO THE YEAST GENOME

Once the plasmids had been constructed the next step was to choose the replication origin to which to apply the excision strategy. Taking into consideration that less is known about the proteins that regulate late origin firing, especially when the S-phase checkpoint has been activated, I decided to insert the R sites into the surrounding regions of a late origin.

Thus, after investigating the literature and the OriDB website (Nieduszynski, *et al*, 2007), where information regarding all the known *S. cerevisiae* replication origins is reported, I decided to focus my project on the late replication origin ARS1413. In studies to determine the key features of late origin firing, ARS1413 was identified as one of the four origins that form a cluster on Chr. XIV (Friedman, *et al*, 1995). Even though it is activated late during the S-phase (T.rep 28.4 min) ARS1413 is an efficient origin (Friedman, *et al*, 1996). Late activation of this cluster is not due to a telomeric effect but depends on the chromosomal context of the surrounding regions that can act as late determinants (Friedman, *et al*, 1996). Indeed construction of plasmids in which ARS1413 was cloned with differing amounts of its flanking

sequences demonstrated that the origin alone fires early in S-phase while association with flanking chromosomal sequences determined its late activation. Therefore, ARS1413 represents a well characterized late replication origin.
Moreover considering that the pLT1/RsiteA and pLT2/RsiteB constructs will be integrated into the genome, it is essential that ARS1413 surrounding regions do not encode for essential genes, ensuring that cell survival is not compromised. Indeed *YNL211C* and *YNL212W*, located upstream of ARS1413, are genes that encode for proteins of unknown function; both null mutants are viable. *YNL210W*, situated at the 3′-end of the origin, synthesizes a protein with an RNA-binding motif required for RNA splicing of meiosis genes.
Deletion of this gene does not interfere with cell viability and considering that the experiment will be performed in mitotic cells, insertion of the R site should not compromise any essential function.

3.9 SOUTHERN BLOT ANALYSIS OF W303-1a, MVY106 AND MVY107

In order to further confirm the insertion of the recombination sites into the *S. cerevisiae* genome and to evaluate how many copies of each fragment had been integrated, a Southern blot analysis was performed on MVY106 and MVY107; the initial WT strain W303-1a is the negative control for the experiment. In order to demonstrate the integration of both the constructs into the genome, the DNA samples were restricted with enzymes that give fragments of different sizes, according to the presence or not of the pLT1/RsiteA and pLT2/RsiteB inserts.
A schematic map of the genomic regions surrounding ARS1413 in the strains used with the positions of the EcoNI and PvuI restriction sites is reported in Fig.41 (panels A-C). Digestion of the WT strain with EcoNI produces a fragment of 1.9 kbp (panel A of Fig.41. Instead the presence of the pLT1/RsiteA fragment introduces a new restriction site for EcoNI resulting in a 2.5 kbp fragment (Fig.41 B). The PvuI enzyme was used to detect the integration of the pLT2/RsiteB construct on the left side of ARS1413. This enzyme cuts both inside the replication origin and in *YNL212C* producing a 3.6 kbp fragment in a WT strain (Fig.41 A).

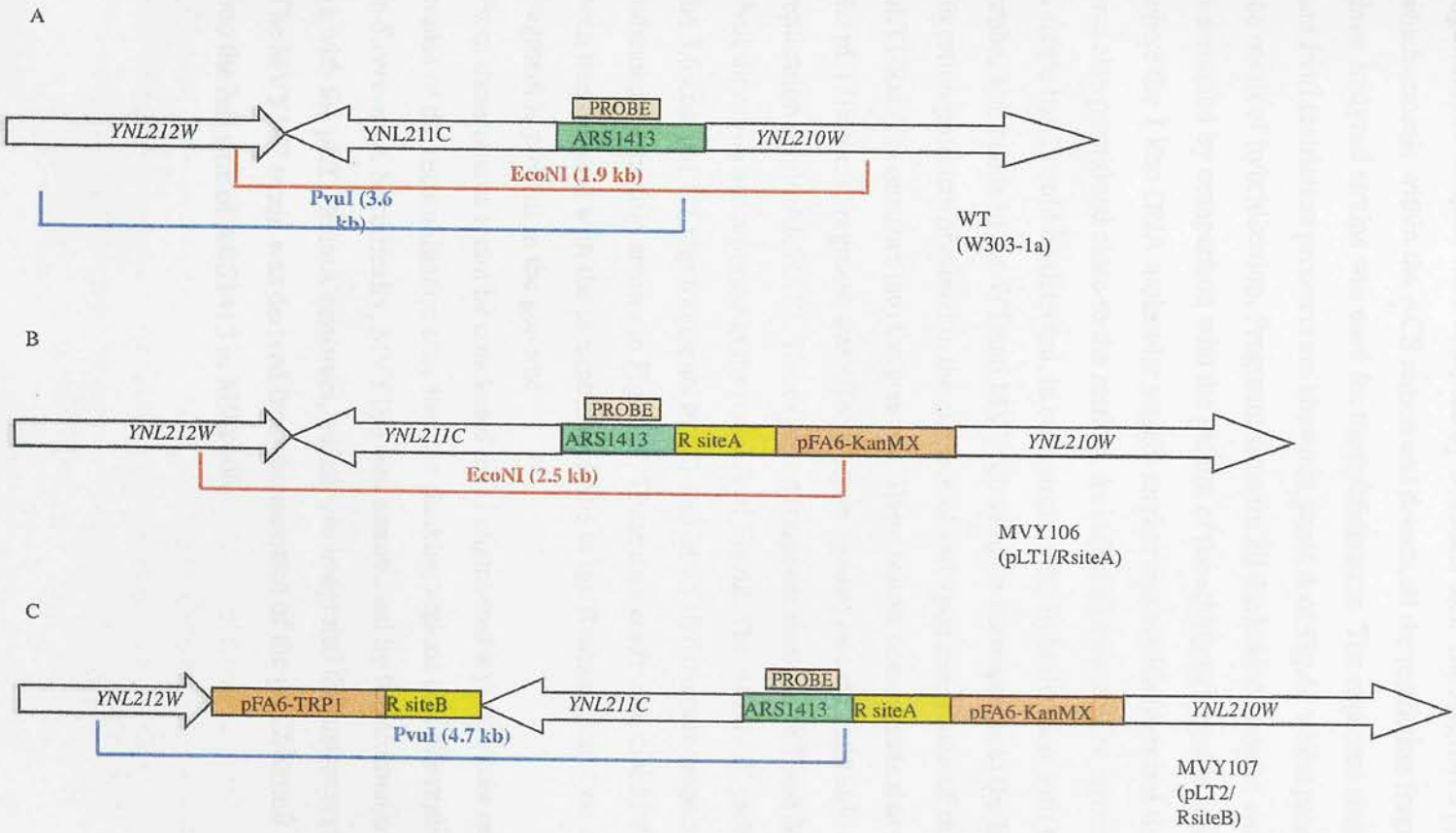


Fig.41 Schematic maps of the ARS1413 regions in the strains used to confirm the integration of the R sites. The sizes of EcoNI and PvuI digestion products are reported. A) Map of the untransformed WT, W303-1a strain. B) MVY106 derived from the integration of the pLT1/RsiteA fragment into the right side of ARS1413. C) Map of MVY107 derived from transformation of MVY106 with the pLT2/RsiteB fragment.

However if the recombination sites, ligated to the *TRP1* gene, are integrated upstream of ARS1413 the digestion produces a 4.7 kbp fragment (Fig.41 C). The differences in the bands obtained upon restriction with Eco NI and PvuI allows the discrimination between the WT and the engineered MVY106 and MVY107. After digestion the samples were analyzed by Southern blot. The ARS1413 probe (Fig.41), which anneals within the ACS region and detects all the restriction fragments in the three analyzed strains was used for the hybridization. The expected sizes of EcoNI and PvuI restriction products are shown in panel A of Fig.42 whilst panel B shows the result of hybridization. Fragment sizes for all the loaded samples were determined by comparison with the picture of the ethidium bromide stained gel where the 1 kbp DNA molecular weight marker was visible (data not shown); a ruler was also positioned close to the marker. As indicated by the thick arrows (Fig.42 B) a single band could be detected, in each strain, after hybridization with the ARS1413 probe. The bands in the WT and MVY106 samples corresponds to the EcoNI digestion products obtained in the absence of and upon integration of the pLT1/RsiteA construct into the genome. These results demonstrate that one copy of the pLT1/RsiteA fragment was effectively integrated into the right side of the late replication origin ARS1413. The expected fragment sizes for WT and MVY107 after PvuI digestion are reported in the panel A of Fig.42. The ARS1413 probe detected the 3.6 kbp and 4.7 kbp fragments in WT and MVY107 samples respectively, as indicated by the thin arrows in Fig.42 B. These data confirmed that MVY107 had been transformed with the recombination site in the B orientation. One copy of the fragment is present in the genome.

From these results it can be concluded that I engineered a yeast strain integrating two copies of the recombination sites into the flanking regions of a late replication origin in *S.cerevisae*. Specifically, MVY106 was constructed by transformation of W303-1a with the pLT1/RsiteA construct, which was integrated downstream of ARS1413. The MVY107 strain was derived from the insertion of the pLT2/RsiteB fragment into the left side of ARS1413 in MVY106.

A

EcoNI → WT=1900 bp

→ MVY106 = 2500 bp (pLT1/RsiteA)

PvuI → WT=3600 bp

→ MVY107 = 4700 bp (WT/pLT2RsiteB)

B

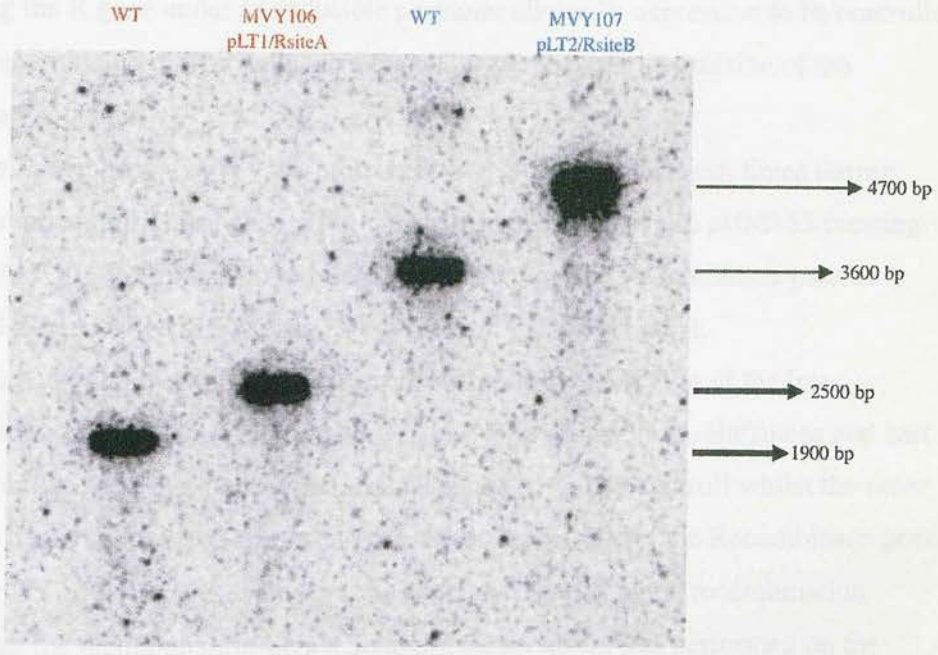


Fig.42 Southern Analysis to confirm the integration of both the R sites into the flanking regions of ARS1413. A) Sizes of the fragments obtained upon digestion with EcoNI and PvuI. B) Picture of the membrane hybridized with the ARS1413 probe. The different colours of the strains correspond to the enzyme used for the DNA digestion; the arrows on the right side of the gel indicate the sizes of the detected bands

3.10 ARS1413 EXCISION IN MVY126

Recombination between the recombination sites will induce the excision of ARS1413, producing a shorter Chr. XIV and a chromatin plasmid consisting of the origin, the R sites A and B and the comprised *YNL210C* gene. As described before this site-specific recombination system is completely independent from the one present in *S. cerevisiae*. Indeed it relies exclusively on its specific Recombinase protein, which is encoded by the R gene, contained in plasmid pSR1.

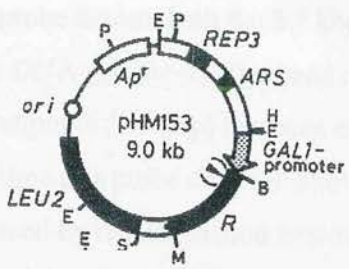
In studies to demonstrate the power of the pSR1site-specific recombination system, Matsuzaki and colleagues constructed the plasmid pHM153 (Matsuzaki, *et al*, 1988). This is a high copy number vector containing the 2 micron plasmid replication origin and the *LEU2* selective marker (Fig.43 A). Moreover it contains the sequence of the R gene, which is under the GAL promoter.

Therefore I took advantage of pHM153 to introduce the R gene into strain MVY107. Having the R gene under an inducible promoter allows its expression to be controlled in order to decide in which phase of the cell cycle to promote excision of the replication origin.

This will enable us to study the proteomes of the origin at different times during DNA replication. Thus MVY107 was transformed with plasmid pHM153 creating strain MVY126. Expression of the gene encoding for the Recombinase protein should induce ARS1413 excision from its chromosomal context.

A schematic graph of the experiment planned to induce excision of the late replication origin is shown in Fig.43 B. Cells were grown in 3% Raffinose and half of the culture was maintained in the same media (negative control) whilst the other half is treated with 2% Galactose to induce the expression of the Recombinase gene. To confirm if ARS1413 had been excised from Chr. XIV upon recombination between the recombination sites a Southern blot analysis was performed on the prepared samples.

A



B

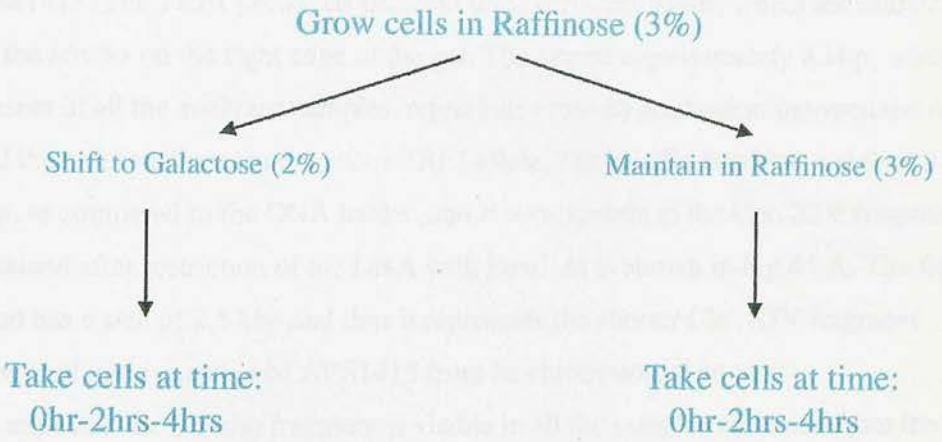
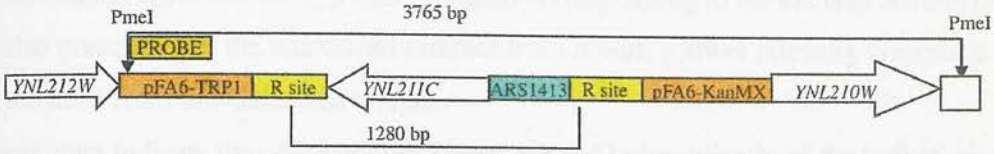


Fig.43 Induction experiment. A) Map of the pHM153 plasmid; *LEU2* is the selective marker gene. The R indicates the sequence of the Recombinase gene; as shown in the map the gene is under the control of the Gal promoter. B) Schematic diagram of the experiment designed to induce ARS1413 excision. Cells are grown in 3% Raffinose and then half of the culture is treated with 2% Galactose to promote the expression of the R gene. Samples are collected at different times to follow the kinetics of the excision.

A diagram of the Chr. XIV region where ARS1413 is located and the results obtained with the Southern blot analysis are reported in panels A and B, of Fig.44 respectively. The TRP1 probe detects both the 3.7 kbp Chr. XIV fragment after PmeI digestion of the genomic DNA and the smaller band of 2.5 kbp, which will only be present if the ARS1413 segment (1.3 kbp) has been excised from the chromosome (Fig.44 A). At the same time this probe does not allow the visualization of the chromatin plasmid produced by recombination between the recombination sites. The picture of the hybridized membrane is shown in Fig.44 B. To compare the kinetics of the excision, samples from the induced (Galactose) and uninduced (Raffinose) cultures were collected at the same time points. The first three samples loaded on the gel represent the digested DNA extracted from the cells grown in 3% Raffinose; the t0, t2 and t4 labels indicate the time points (hours) that cells were collected. The second set of samples derives from the cells that were treated with 2% Galactose in order to induce the expression of the Recombinase gene and thus the excision of ARS1413. The TRP1 probe has detected three different bands, which are indicated by the arrows on the right edge of the gel. The one of approximately 8 kbp, which is present in all the analyzed samples, represents cross-hybridization between the probe and the non functional endogenous *TRP1* allele. The middle band has a size of 3.7 kbp, as compared to the DNA ladder, and it corresponds to the Chr. XIV fragment obtained after restriction of the DNA with PmeI, as is shown in Fig.44 A. The third band has a size of 2.5 kbp and thus it represents the shorter Chr. XIV fragment produced upon excision of ARS1413 from its chromosomal context.

As expected the 3.7 kbp fragment is visible in all the samples collected from the cells grown in Raffinose; the same band is also visualized in the induced samples. However in the samples derived from the cells grown in Galactose a decrease in the intensity of the 3.7 kbp fragment is observed (Fig.44 B). This is due to the induction of recombination and the appearance of the 2.5 kbp band as a result of ARS1413 excision. Indeed the increase in the signal of the lower band is proportional to the reduction of the intensity of the upper band. This data indicates that ARS1413 has been excised from its chromosomal location and that with Gal induction there is an increase in the recombination event.

A



No excision: PmeI digestion product ⟶ 3765 bp

ARS1413 excision: PmeI digestion product ⟶ 2485 bp

B

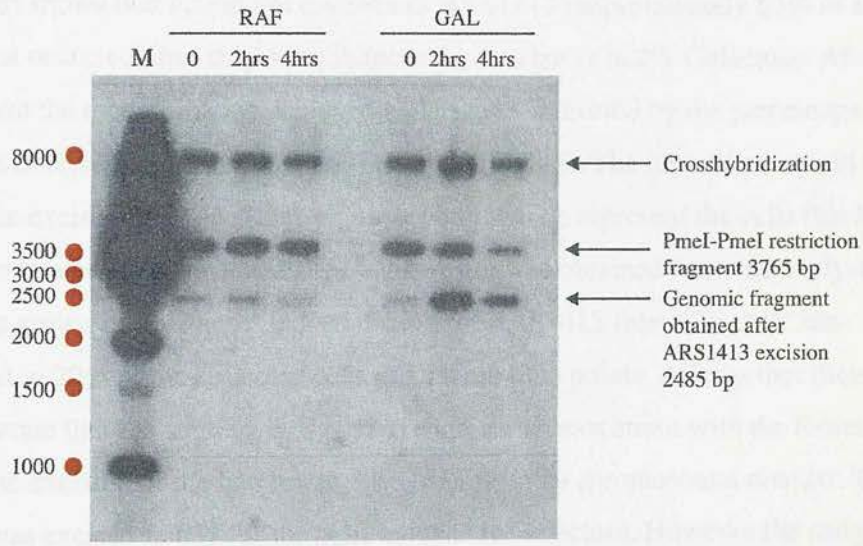


Fig.44 Excision of the late replication origin ARS1413 in MVY126. A) Schematic map of the Chr. XIV region; ARS1413 and the PmeI restriction sites are shown. PmeI digestion in the absence of recombination or after ARS1413 excision produces fragments of 3765 bp and 2485 bp respectively. The TRP1 probe (orange box) anneals within the inserted *TRP1* gene and detects both the digestion products. B) Picture of the Southern blot analysis performed to confirm ARS1413 excision. M: 1 kbp plus MW ladder. Samples from cultures grown in Raffinose (Raf) and Galactose (Gal) were loaded; the arrows indicate bands detected by the probe. The origin has been excised from its chromosomal context. However the recombination also took place in the uninduced samples.

However the band indicating that excision of the replication origin has occurred was also detected in the sample taken prior to Gal induction (t0 time point, Fig.44 B). A similar result is also observed in the cells derived from the Raffinose culture. Indeed, apart from the 3.7 kbp band, a signal corresponding to the excised ARS1413 is also present in all the uninduced samples even if with a lower intensity compared to the cells from the Galactose culture.

These data indicate that ARS1413 has been excised independently of the induction of the Recombinase gene. To measure the efficiency of the recombination I analyzed the percentage of cells in which the excision had occurred; for this purpose the intensities of the bands on the membrane were quantified using the OptiQuant Image program. The values for both the 3.7 kbp and 2.5 kbp fragments in each sample were summed and then normalized to the value of the lower band; the resulting values are reported as a percentage in the histogram graph in Fig.45.

The chart shows that maximum excision of ARS1413 (approximately 65% of the cells) has occurred after they were induced for two hours in 2% Galactose. After this time point the recombination achieved a plateau as indicated by the percentage of excision obtained after four hours of induction (Fig.45). The replication origin has also been excised in 20% of the "t0" time point, which represent the cells that have not been treated with Galactose. The same result was obtained from the analysis of the cells grown in Raffinose. Indeed the loss of ARS1413 from Chr. XIV has occurred in 20% of the collected cells at all three time points. All together these data demonstrate that the strategy designed to engineer a yeast strain with the Rsites has led to the excision of the late origin ARS1413 from its chromosomal context. The origin was excised in 65% of the cells induced in Galactose. However the samples from the Raffinose culture, where the Recombinase gene should be completely repressed, showed excision in approximately 20% of the cells, at all three time points. This is a high value expecially when compared to the one obtained in the induced cells (65%). All together these data indicate that the recombination had to be improved in order to reduce the background and increase the percentage of cells in which the excision occurs for the following mass spectrometry analysis.

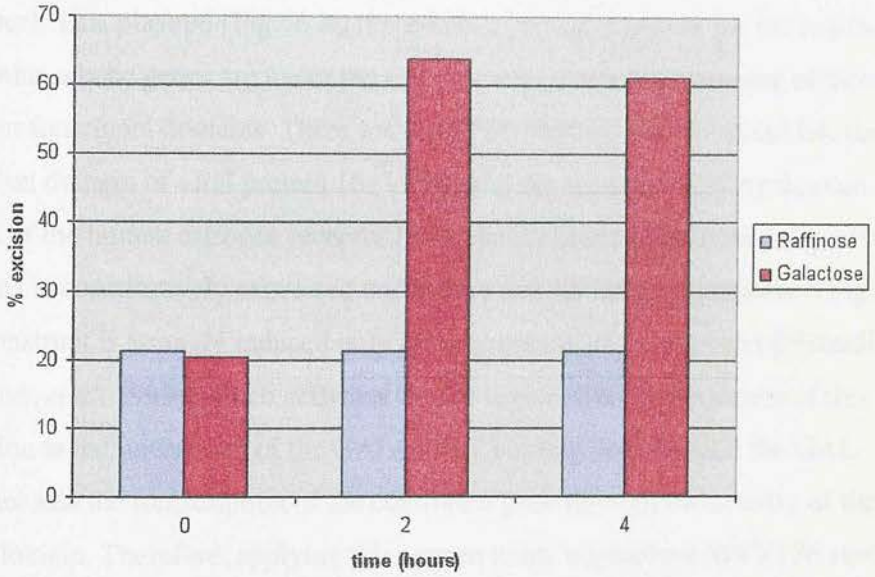


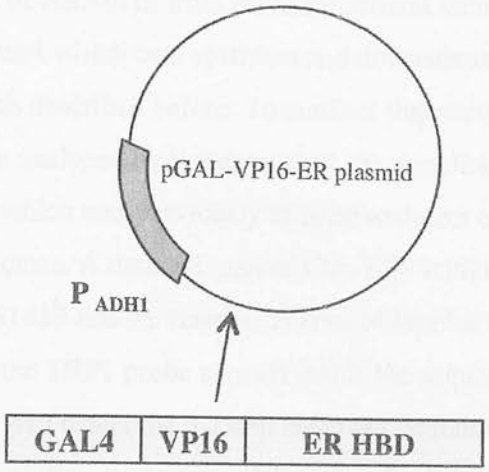
Fig.45 Quantification of the ARS1413 excision in MVY126. The intensities of the bands detected by the TRP1 probe were measured and plotted as a percentage of cells in which the excision has occurred. Excision at each time point is reported for both growth conditions. In the induced cells (Gal) the excision of the replication origin occurred in 60% of the cells after two hours growth in Galactose. However recombination also took place in 20% of the cells grown in Raffinose (uninduced).

3.11 INDUCTION OF R GENE EXPRESSION THROUGH PLASMID pGAL-VP16-ER

The recombination system is strictly dependent on the specific Recombinase protein encoded by the R gene, which was introduced by transformation of MVY107 with plasmid pHM153. The expression of the gene requires the presence of Galactose because it is under the control of the Gal promoter.

The excision of ARS1413 observed in the Raffinose sample could be due to the R gene not being completely repressed in this growth media; as a consequence the low transcription could be sufficient to induce a low rate of recombination. For this reason I decided to take advantage of a different system to control the expression of the R gene, which relies on plasmid pGAL-VP16-ER (kindly donated by E. Schirmer). This plasmid (Fig.46 A) represents a powerful system for the regulation of proteins whose genes are under the Gal promoter due to the presence of three different functional domains. These are the DNA-binding domain of GAL4, the activation domain of viral protein 16 (VP16) and the hormone binding domain (HBD) of the human estrogen receptor (ER); the plasmid region containing all these elements is constitutively expressed under the yeast *ADH1* gene promoter (Fig.46). This construct is strongly induced only in the presence of the estrogen βEstradiol (Louvion, *et al*, 1993), which activates the ER region. The consequence of this activation is the interaction of the GAL4 DNA-binding domain with the GAL promoter and the transcription of the controlled gene through the activity of the VP16 domain. Therefore, applying this system to the engineered MVY126 strain will allow cells to be grown in Glucose. Subsequent treatment with β-Estradiol will then promote the interaction of the Gal binding domain with the promoter of the R gene and the synthesis of the Recombinase protein. Plasmid pGAL-ER-VP16 was introduced into MVY126 by yeast transformation creating strain MVY128.

A



B

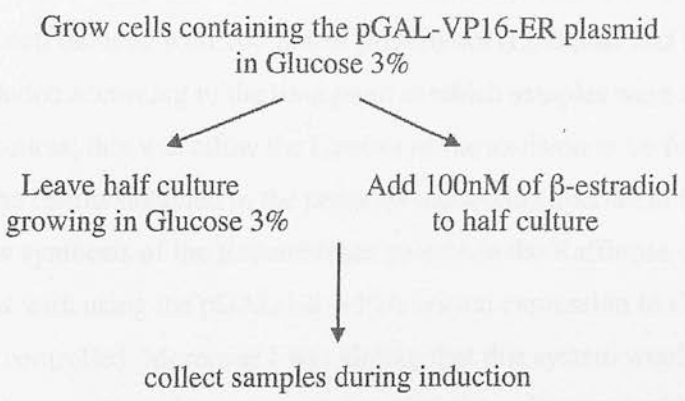


Fig.46 Induction of the excision using the pGAL-VP16-ER system. A) Schematic representation of the plasmid used. The cassette containing the GAL4 binding domain, the VP16 region and the binding domain for the human hormone, β-estradiol, are expressed under the promoter of the *ADH1* gene. B) Outline of the induction experiment. The transformed MVY128 strain is grown in Glucose; half of the culture is then treated with 100 nM of β-estradiol. Samples are collected at different times during the induction.

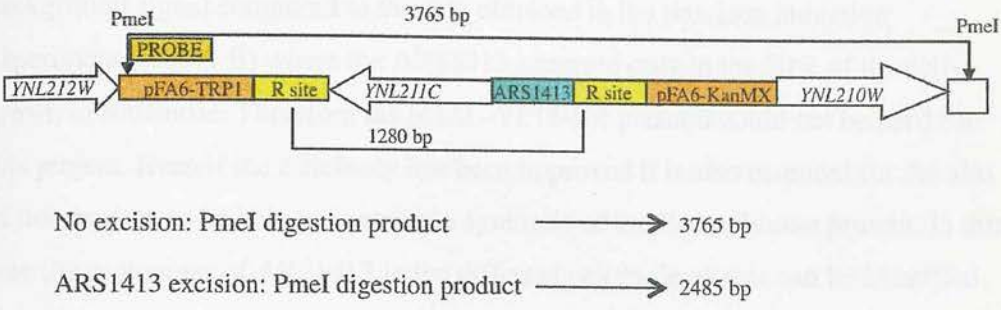
This new method of R gene induction was used to test the efficiency of ARS1413 excision. A schematic representation of the experiment designed to promote the hormone-inducible expression of the R gene is reported in Fig.46 B. To detect the excision of ARS1413 from its chromosomal locus the DNA samples were restricted with PmeI which cuts upstream and downstream of the region of interest on Chr. XIV, as described before. To confirm that recombination had occurred samples were analyzed by Southern blot; the membrane was hybridized with the TRP1 probe, which was previously utilized to detect excision in the strain induced with 2% Galactose. A detailed map of Chr. XIV with the positions of the replication origin ARS1413 and the restriction sites of PmeI is reported in Fig.47 A. As shown on the map the TRP1 probe anneals inside the sequence of the integrated *TRP1* gene and thus it will detect the 3.7 kbp fragment containing the replication origin produced by PmeI digestion of Chr. XIV (Fig.47 A). If ARS1413 has been excised from its chromosomal context PmeI digestion will produce a fragment of 2.5 kbp, which will also be detected by the same probe. The picture of the hybridized membrane is shown in panel B of Fig.47. All the DNA samples collected from both cultures, at the different time points, were loaded; samples were divided according to the growth condition of the cells. The first group represents MVY128 grown in YPD whilst the second set of samples derives from the cells collected from the same strain that has instead been induced with 100 nM of β-Estradiol (Lindquist and Kim, 1996). The DNA was loaded according to the time point at which samples were taken during the time course; this will allow the kinetics of the excision to be followed. Supposing that the results obtained in the previous induction experiment (Fig.44 B) were due to a low synthesis of the Recombinase protein in the Raffinose culture, I would expect that with using the pGAL-ER-VP16 system expression of the R gene would be tightly controlled. Moreover I was aiming that this system would not only eliminate the background but also improved the efficiency of the recombination. Therefore in the uninduced MVY128 strain the R gene should be completely repressed and no recombination should occur. At the same time treatment of the other half of the MVY128 culture with the hormone should specifically promote ARS1413 excision.

Actually contradictory results were obtained after hybridization with the TRP1 probe; in fact the excision was improved in both the cultures analyzed.

In the samples treated with 100nM of β-Estradiol (Fig.47 B) only one band could be detected by the probe. As indicated by the arrow this band represented the 2.5 kbp fragment obtained by PmeI digestion of Chr. XIV lacking the replication origin. Considering that the signal for the 3.7 kbp fragment could not be visualized it can be concluded that the excision of ARS1413 has occurred in 100% of the induced cells. However the fragment indicating ARS1413 excision was also present in the sample collected before induction (t0 time point). This data demonstrates that the replication origin was excised even before activation of the pGAL-ER-VP16 plasmid with βEstradiol.

A similar hybridization pattern is observed in the negative control represented by cells grown only in 2% Glucose (Fig.47 B). Indeed in all the samples the TRP1 probe has only detected the lower band, as indicated by the arrow. Thus the recombination also occurred in the samples that were not treated with βEstradiol. In addition the intensities of the signals obtained for these samples are comparable to those obtained for the induced samples.

A



B

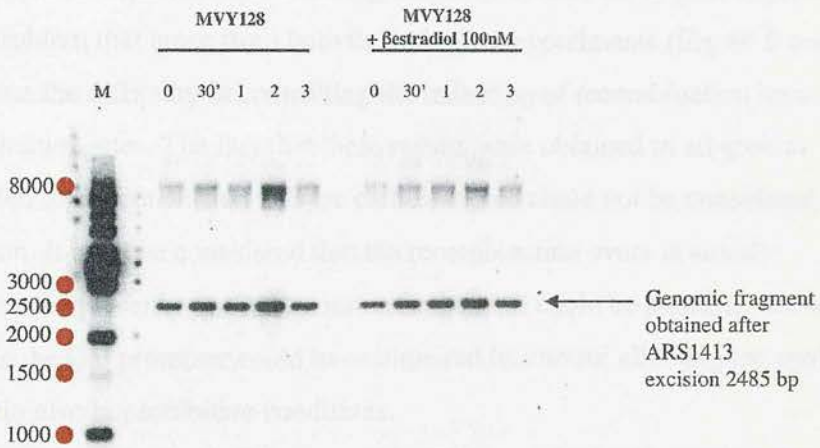


Fig.47 Induction of ARS1413 excision in MVY128. A) Representation of the Chr. XIV region containing ARS1413. The digestion products obtained upon restriction of the DNA preparations with PmeI are reported. The TRP1 probe used for the hybridization is represented by the orange box. B) Picture of the membrane after hybridization with the TRP1 probe. M: 1 kbp MW ladder. The MVY128 indicates the samples collected from the culture grown in Glucose. The other samples (MVY128 +β-estradiol) derived from the culture treated with the human hormone. The arrow points to the 2458 bp band detected by the probe. The results demonstrate that the excision occurred in 100% of the cells independent of the hormone treatment.

Taking into consideration all the data it can be concluded that the efficiency of the recombination system has been improved because the origin is excised in approximately 100% of the cells. However there is also an increase of the background signal compared to the data obtained in the previous induction experiment (Fig.44 B) where the ARS1413 occurred only in the 20% of the cells grown in Raffinose. Therefore the pGAL-VP16-ER plasmid could not be used for this project. Even if the efficiency has been improved it is also essential for the aim of this project, to be able to control the synthesis of the Recombinase protein. In this way the proteomes of ARS1413 in the different cell cycle phases can be identified.

3.12 ARS1413 EXCISION IN THE CONSTRUCTED MVY107, MVY126 AND MVY128

Different hypotheses could be proposed to explain why the pGAL-ER-VP16 system did not function as expected, thus allowing a tight control of the R gene expression. The main problem that arose from both the induction experiments (Fig.44 B and Fig.47 B) was the difficulty in controlling the induction of recombination between the recombination sites. The fact that these results were obtained in all growth conditions led to the conclusion that the carbon source could not be considered the main solution. It must be considered that the recombination event is strictly dependent on the presence of the Recombinase gene. It could be possible that some mutations in the Gal promoter could have impaired its control allowing the synthesis of the protein also in prohibitive conditions.

However a defect at the level of the recombination event by itself could not be excluded. It might be that in MVY107 the excision is not independent from the host organism. In fact this hypothesis could also explain the detection of ARS1413 excision in all the conditions used.

Therefore, I decided to discriminate if the difficulty in controlling the recombination event was due to a problem in the promoter of the R gene, or to the fact that the adopted system is not dependent on the presence of its specific Recombinase protein.

For this purpose the efficiency of ARS1413 excision was compared between strains MVY107, MVY126 and MVY128; the genotype of the three strains is summarized in the table reported in panel A of Fig.48.

The MVY126 strain contains the recombination sites and plasmid pHM153, which contains the sequence of the R gene under the Gal promoter; instead the MVY128 strain, obtained by transformation of MVY126 with pGAL-VP16-ER, contains both plasmids. The MVY126 and MVY128 strains were derived from strain MVY107, containing the R sites flanking ARS1413, which was also used to confirm if the excision occurs even in the absence of the Recombinase gene.

All the three strains were grown in different conditions as is schematically reported in panel B of Fig.48 and the induction of recombination was performed as in the previous experiments.

Strain MVY107 was grown exclusively in Glucose. It does not contain the R gene sequence and therefore, considering that the site specific recombination system relies only on the Recombinase protein (Matsuzaki, *et al*, 1988; Raghuraman, *et al*, 1997), no excision should be detected in this strain. The presence of a signal indicating that the recombination has occurred would instead demonstrate that the system is not independent from the one present in *S. cerevisiae*.

Strains MVY126 and MVY128 were grown in Glucose as well. However a portion of the MVY126 culture was induced with Galactose to promote expression of the R gene. Therefore loss of the origin from Chr. XIV should take place only in the cells treated with Galactose whilst the R gene should be completely repressed in the cells left growing in Glucose.

The MVY128 culture was not treated with βEstradiol to confirm the results obtained previously for the uninduced samples; the profile of this strain should correspond to the one of strain MVY126 grown in Glucose.

A comparison of the results obtained in these strains will allow to understand what could impair the control of the induction of recombination.

A

Yeast strain / Plasmid	MVY107	MVY126	MVY128
pHM153	-	+	+
pGAL-VP16-ER	-	-	+

B

MVY107 ⟶ **Glucose**

MVY126 ⟶ **Glucose and Galactose 2%**

MVY128 ⟶ **Glucose**

Fig.48 Description of the strains and of the experiment performed to determine what is causing the loss of control of the R gene expression. A) Table indicating which plasmids are contained in the strains tested for the excision. B) Schematic representation of the media used for the induction experiment. MVY107, which contains only the R sites, was grown in Glucose; no excision should be detected. MVY126 and MVY128 were both grown in Glucose. MVY126 was also treated with Galactose to induce the expression of the R gene.

The DNA recovered from all the described samples was restricted with PmeI and the digestion products were detected by hybridization with the TRP1 probe that allows visualization of the 3.7 kbp and 2.5 kbp fragments. The results of the Southern blot are reported in Fig.49 B. The samples from the three strains are divided in three groups according to the time in which they were collected: thus the first group (t0) represent the samples collected before the start of the induction, whilst the other two indicate the time, after shifting MVY126 to Galactose, at which the cells were taken: 2hours and 4 hours. In each group the samples were loaded in the following order: MVY107, MVY126 Glucose (negative control), MVY126 Galactose (positive control) and MVY128.

In the strain that does not have the Recombinase gene, MVY107, the only visible band at all the three time points (Fig.49 B first lane in each group) is the 3.7 kbp fragment, which is obtained only if ARS1413 has not been excised. Therefore these data demonstrate that the recombination system is independent from the host organism and relies only on its specific Recombinase protein. In the samples derived from MVY126 grown in 2% Glucose two bands could be visualized of which the higher one corresponds to the fragment obtained when no excision has occurred. Instead the lower fragment indicates the 2.5 kbp fragment detectable only when ARS1413 has been excised. A gene under the Gal promoter should be completely repressed when grown in media containing Glucose, therefore the Recombinase gene should not be activated in this condition. Nonetheless in MVY126 grown in Glucose 2% the band indicating the excision of the origin was present at all the analyzed time points. These results could indicate, considering also the data obtained in the study of MVY107, that even in the presence of Glucose the R gene is derepressed.

A similar profile is observed in MVY126 induced for two hours in Galactose. In the samples collected at 2 hours and 4 hours after the shift to Galactose both the 3.7 kbp and 2.5 kbp bands were detected by the TRP1 probe. From the intensities of the two bands, in all the samples, it could be deduced that the excision occurred in all the treated cells. Recombination between the recombination sites has also taken place before the induction in Galactose, as evidenced by the bands present at the t0 time point.

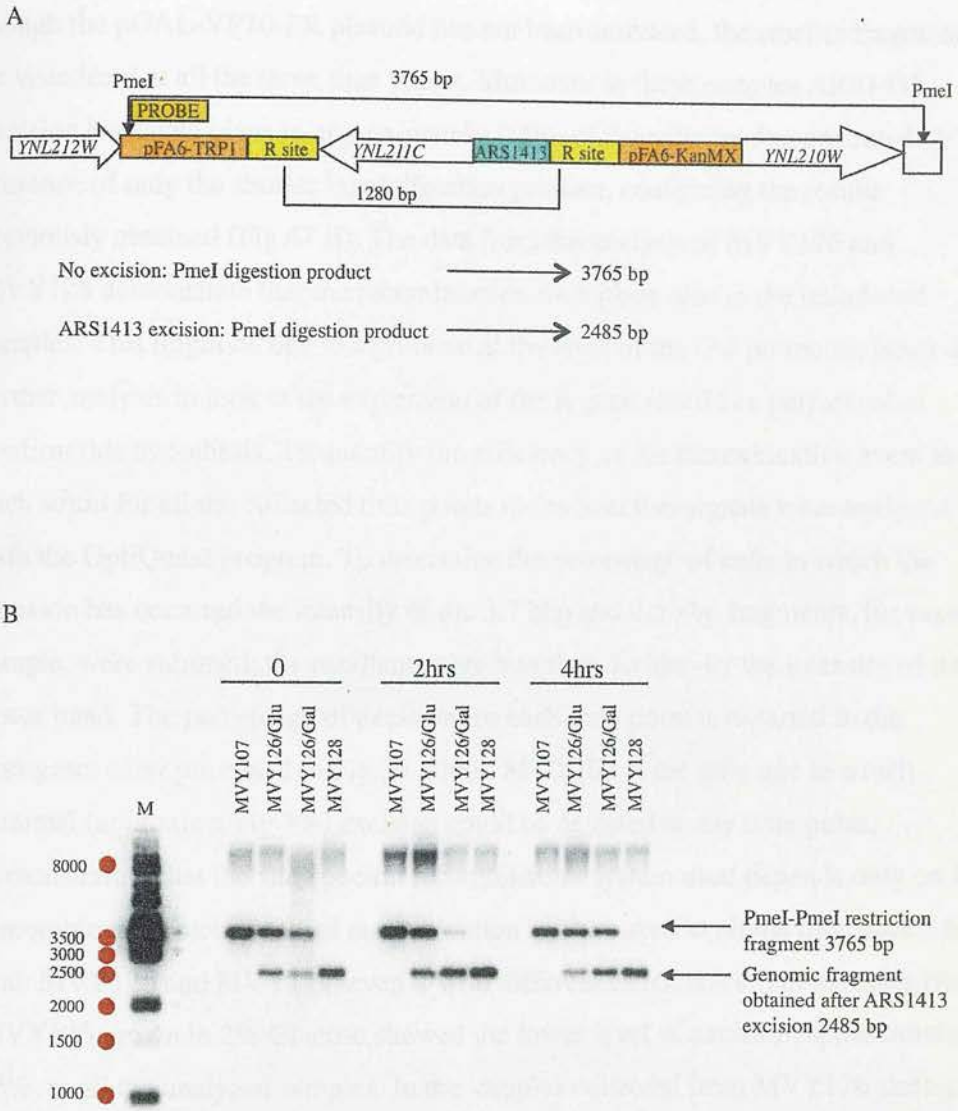


Fig.49 Excision of ARS1413 in MVY107, MVY126 and MVY128. A) Schematic graph of the Chr. XIV region containing ARS1413; PmeI sites and position of the TRP1 probe are reported on the map. Different PmeI digestion fragments are produced according to the presence or not of the origin on the chromosome. B) Picture of the membrane after hybridization. M: 1 kbp plus MW ladder. Samples are divided according to the time point at which they were collected; the arrows indicate the bands detected by the probe. No excision could be observed in MVY107 demonstrating that the system is independent from the host organism. The recombination occurred in all the other strains confirming that the difficulty in controlling the R gene expression is due to a defect in the Gal promoter.

These data confirmed that the R gene promoter is not controlled by the different carbon sources used.The last sample loaded in each group is MVY128, which was grown only in YPD to prevent induction of the Recombinase gene expression. Even though the pGAL-VP16-ER plasmid has not been activated, the smaller fragment can be visualized at all the three time points. Moreover in these samples ARS1413 excision has taken place in approximately 100% of the cells, as demonstrated by the presence of only the shorter PmeI digestion product, confirming the results previously obtained (Fig.47 B). The data from the analysis of MVY126 and MVY128 demonstrate that the recombination took place also in the uninduced samples. This might be due to a problem at the level of the Gal promoter; however further analysis to look at the expression of the R gene should be performed to confirm this hypothesis. To quantify the efficiency of the recombination event in each strain for all the collected time points the radioactive signals were analyzed with the OptiQuant program. To determine the percentage of cells in which the excision has occurred the intensity of the 3.7 kbp and 2.5 kbp fragments, for each sample, were summed; the resultant value was then divided by the intensity of the lower band. The percentage of excision for each time point is reported in the histogram chart presented in Fig.50. Strain MVY107 is the only one in which minimal (approximately 3%) excision could be detected at any time point, demonstrating that the site-specific recombination system used depends only on its Recombinase protein. Instead recombination has occurred at all the time points for both MVY126 and MVY128, even if with different efficiencies. The samples from MVY126 grown in 2% Glucose showed the lower level of excision, approximately 35%, in all the analyzed samples. In the samples collected from MVY126 shifted to Galactose the percentage of recombination reached 80% after 2 hours of induction. However the efficiency of recombination seemed to decrease with the proceeding of the induction as demonstrated by 50% excision obtained after 4 hours of growth in Galactose. Thus in MVY126 the percentage of excision is higher in the Galactose samples than in the corresponding Glucose samples.

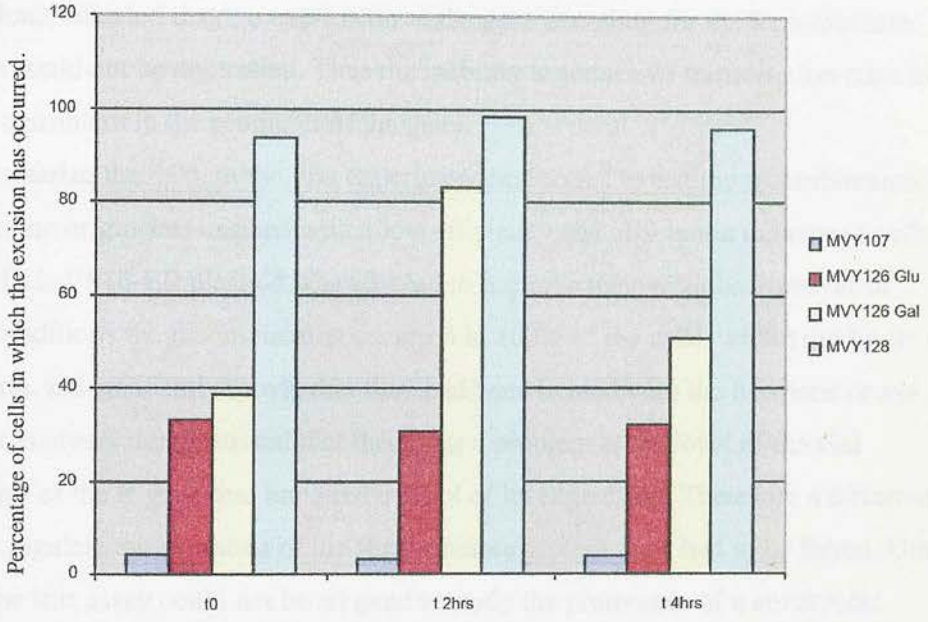


Fig.50 Quantification of the excision in MVY107, MVY126 and MVY128. The intensities of the bands visualized on the Southern shown in the Fig.36 B were quantified and the values obtained are presented as a percentage in the histogram graph. Practically no excision could be observed in MVY107. The recombination took place in the other strains independently of the growth conditions. Maximum excision is observed in MVY128, which was not treated with the human hormone.

The maximum of excision was observed in strain MVY128 where the recombination occurred in 100% of the collected cells. The higher efficiency of ARS1413 excision in this strain could be due to the presence of the pGAL-VP16-ER system.

From these results it can be concluded that the recombination system used is independent from the host organism as demonstrated by the data obtained in MVY107; in fact in the absence of the R gene no excision was detected. At the same time the percentage of excision in MVY126 and MVY128, for any analyzed growth condition, indicated that the expression of the gene encoding for the Recombinase protein could not be controlled. Thus the inability to induce its transcription must be due to a problem in the promoter of the gene.

To summarize the data, in the first experiment performed to test the recombination system, the origin was excised with a low efficiency and also in the uninduced cells. The pGAL-VP16-ER plasmid was adopted to improve the excision. However in these conditions the recombination occurred in 100% of the cells, within the limits of detection, independently of whether they had been treated with the hormone or not. Further analyses demonstrated that there was a problem at the level of the Gal promoter of the R gene that impaired control of its expression. Therefore a different way to regulate the synthesis of the Recombinase protein must had to be found. Until that time this assay could not be adopted to study the proteomes of a eukaryotic replication origin.

3.13 COMPLETE EXCISION OF THE LATE REPLICATION ORIGIN ARS1413 IN THE INDUCED CELLS

The aim of the project was to induce the recombination during different phases of the DNA replication to discovery the proteins involved in each step of genome duplication. Therefore it was essential not only to increase the efficiency of the excision, to then recover enough chromatin plasmid for the mass spectrometry analysis, but also to be able to control the expression of the R gene. A new strategy to control the expression of the Recombinase protein synthesis was adopted. It

consisted of the integration of the gene encoding for the Recombinase protein into the *S. cerevisiae* genome: specifically the R gene was introduced into the endogenous *LEU2* locus of MVY107. Two colonies were then selected; MVY171 and MVY172, which contain two copies and one copy of the R gene respectively.

3.14 EFFECT OF R GENE COPY NUMBER ON ARS1413 EXCISION

Then newly constructed strains MVY171 and MVY172 were then tested for excision of the late replication origin ARS1413; the induction experiment was performed in parallel in both strains to compare the results. Assuming that the insertion of the R gene into the genome has allowed control of its expression, the experiment was designed in the same way as the one described in the first paragraph of this chapter (Fig.43 B). The cells were grown in 2% Galactose to promote the synthesis of the Recombinase protein; samples were collected to follow the kinetics of the excision. To determine if both integration and copy number of R gene into yeast genome have improved the recombination ARS1413 excision was confirmed by Southern blot analysis in MVY171 and MVY172 using the TRP1 probe for the hybridization.

3.15 ARS1413 EXCISION IN MVY172

The picture of the hybridized membrane for strain MVY172 is shown in panel A of Fig.51. The samples collected from the cells grown in Raffinose were loaded in the first part of the gel whilst the following lanes contain the DNA preparations derived from the cultures that were shifted to Galactose. The numbers on the top of each lane indicate the time point (hour) at which the cells were collected from the cultures. In all the samples from the uninduced culture (3% Raffinose) only one band, indicated by the thin arrow, could be visualized. Its size corresponds to the 3.7 kbp fragment obtained after PmeI digestion of Chr. XIV that contains ARS1413. These results demonstrate that the R gene has not been expressed in the cells grown in 3% Raffinose because no recombination event could be detected. However a different digestion profile was observed in the samples derived from the induced culture. In

the lanes where the DNA preparation from the cells grown in Galactose were loaded (1-5 Fig.51), two bands are observed as indicated by the arrows. The upper fragment represents the 3.7 kbp fragment whilst the lower one the 2.5 kbp band, which is obtained only if ARS1413 has been excised. The intensity of the upper band decreases with increasing induction and it completely disappears at the last time point. Contemporary to the reduction in the upper band intensity, an increase in the signal of the lower fragment can be observed. The 2.5 kbp band can be observed exclusively in the samples collected from the cells grown in Galactose. These results indicate that the recombination has only occurred in the induced samples, as also demonstrated by the absence of the fragment obtained upon ARS1413 excision in the samples collected from the Raffinose culture. Therefore integration of the R gene into the genome has resulted in the control of its expression and of the recombination event. Moreover an improvement in the efficiency of the excision could also be observed.

In order to evaluate the percentage of cells in which the excision had occurred the intensities of the radioactive signals were analyzed with the OptiQuant Image program. As before the values resulting from the sum of the intensities of the 3.7 kbp and 2.5 kbp bands were divided by the intensity of the corresponding lower fragment; the results obtained, expressed as a percentage, are reported in the histogram graph in panel B of Fig.51.

The analysis confirms that the R gene is repressed in the samples grown in Raffinose because essentially no excision (only 3% in the samples collected after 3 hours) of ARS1413 could be observed. Instead treatment of cells with Galactose has induced the synthesis of the Recombinase protein. Indeed excision of the origin has been detected in these samples.

Moreover an increase in the percentage of cells in which the excision has occurred was observed over the course of the induction reaching approximately 95% 5 hours after shifting to Galactose. Therefore integration of the R gene into the genome has increased the percentage of excision and, at the same time, it has completely eliminated the background observed in all the previous experiments

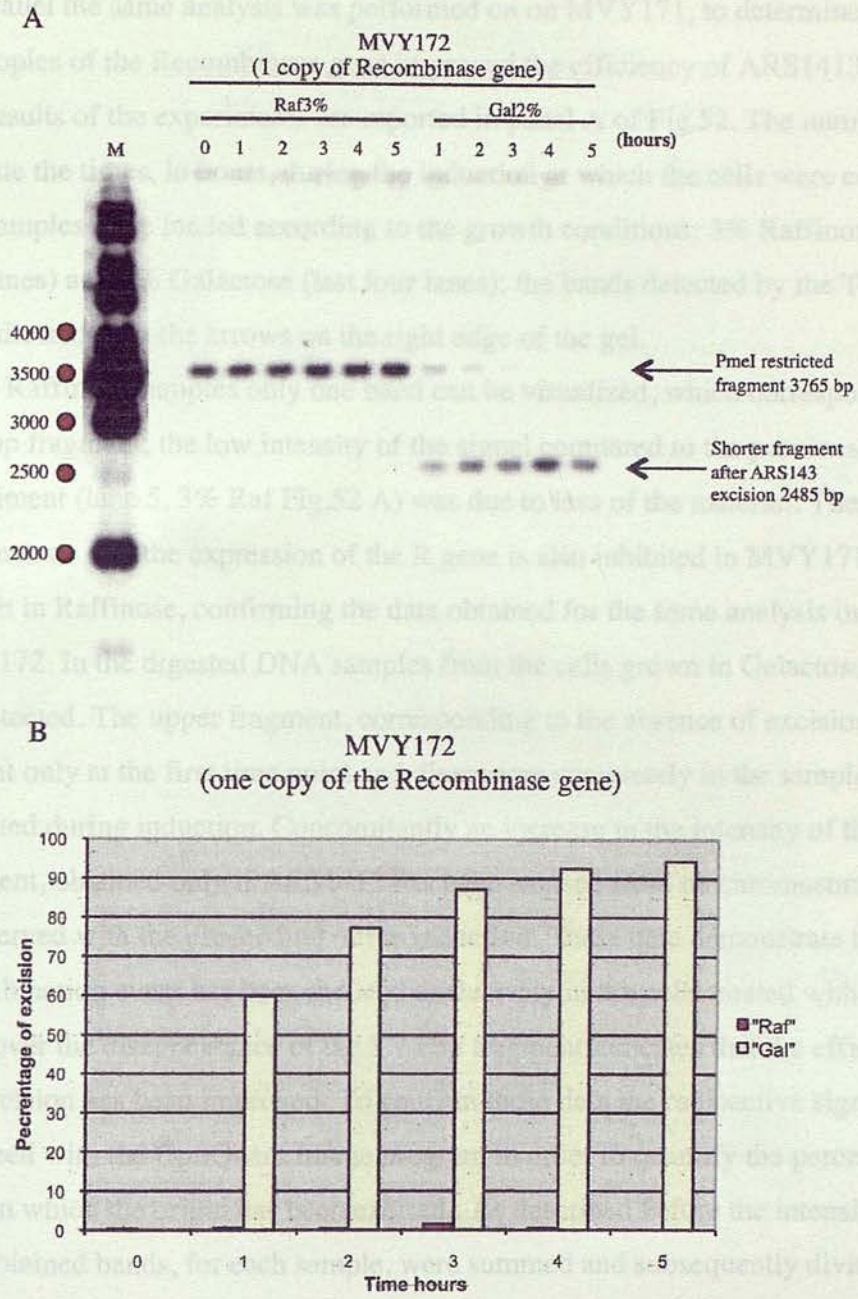


Fig.51 Excision of ARS1413 in MVY172. A) Southern blot analysis on samples collected from both Raffinose and Galactose cultures; the membrane was hybridized with the TRP1 probe. The arrows indicate the band detected by the probe; the 1 kbp molecular weight marker was loaded in the lane labeled with M. B) Quantification of the percentage of cells in which recombination has occurred. ARS1413 has been excised in 90% of the cells grown 5 hours in Galactose, with no excision in the cells grown in Raffinose.

3.16 ARS1413 EXCISION IN MVY171

In parallel the same analysis was performed on on MVY171, to determine if having two copies of the Recombinase gene improved the efficiency of ARS1413 excision. The results of the experiments are reported in panel A of Fig.52. The numbers 0-5 indicate the times, in hours, during the induction at which the cells were collected. The samples were loaded according to the growth conditions: 3% Raffinose (first five lanes) and 2% Galactose (last four lanes); the bands detected by the TRP1 probe are indicated with the arrows on the right edge of the gel.

In the Raffinose samples only one band can be visualized, which corresponds to the 3.7 kbp fragment; the low intensity of the signal compared to the previous experiment (lane 5, 3% Raf Fig.52 A) was due to loss of the material. These results demonstrate that the expression of the R gene is also inhibited in MVY171 by growth in Raffinose, confirming the data obtained for the same analysis on MVY172. In the digested DNA samples from the cells grown in Galactose two bands are detected. The upper fragment, corresponding to the absence of excision, is present only at the first time point and disappears completely in the samples collected during induction. Concomitantly an increase in the intensity of the shorter fragment, obtained only if ARS1413 has been excised from its chromosomal context, is observed with the proceeding of the induction. These data demonstrate that the recombination event has been induced exclusively in the cells treated with Galactose. Moreover the disappearance of the 3.7 kbp fragment indicates that the efficiency of the excision has been improved. To confirm these data the radioactive signals were analyzed with the OptiQuant Image program in order to quantify the percentage of cells in which the origin has been excised,. As described before the intensities of the two obtained bands, for each sample, were summed and subsequently divided by the value of the 2.5 kbp fragment.

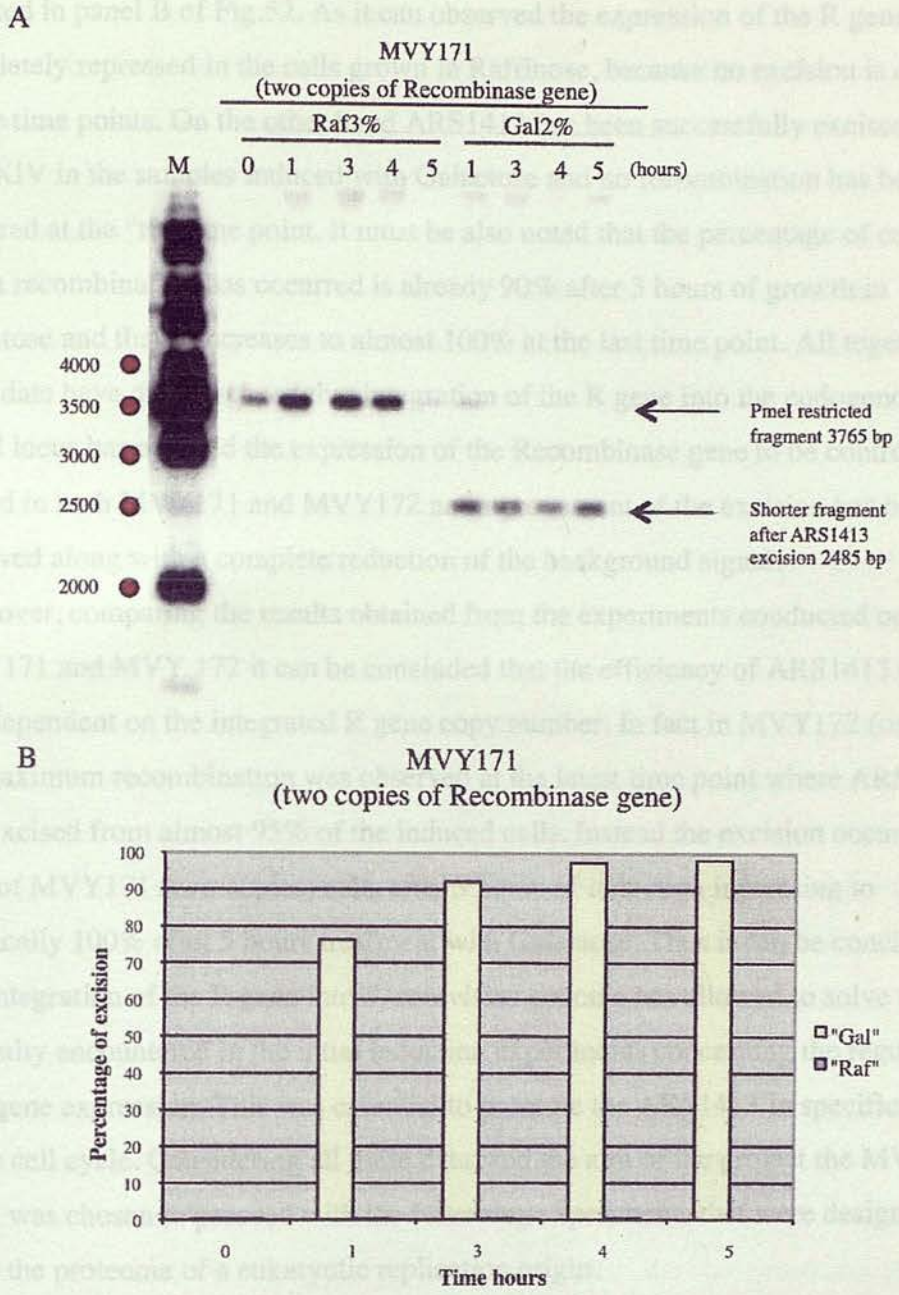


Fig.52 Excision of ARS1413 in MVY171. A) Southern blot analysis on samples collected from Raffinose and Galactose cultures; the membrane was hybridized with the TRP1 probe. M is the 1 kbp plus DNA MW ladder; the arrows indicate the band detected by the probe. B) Quantification of the percentage of cells in which recombination has occurred. ARS1413 has been excised in almost the 100% of cells grown 5 hours in Galactose. The efficiency of the excision is influenced by the R gene copy number.

The resulting values are plotted, as a percentage, in the histogram chart that is reported in panel B of Fig.52. As it can observed the expression of the R gene was completely repressed in the cells grown in Raffinose, because no excision is detected at any time points. On the other hand ARS1413 has been successfully excised from Chr. XIV in the samples induced with Galactose and no recombination has been occurred at the "t0" time point. It must be also noted that the percentage of cells in which recombination has occurred is already 90% after 3 hours of growth in Galactose and that it increases to almost 100% at the last time point. All together these data have demonstrated that integration of the R gene into the endogenous *LEU2* locus has allowed the expression of the Recombinase gene to be controlled. Indeed in both MVY171 and MVY172 an improvement of the excision has been observed along with a complete reduction of the background signal.
Moreover, comparing the results obtained from the experiments conducted on MVY171 and MVY 172 it can be concluded that the efficiency of ARS1413 excision was dependent on the integrated R gene copy number. In fact in MVY172 (one copy) the maximum recombination was observed at the latest time point where ARS1413 was excised from almost 95% of the induced cells. Instead the excision occurred in 90% of MVY171 (two copies) cells after 3 hours of induction increasing to practically 100% after 5 hours treatment with Galactose. Thus it can be concluded that integration of the R gene into *S. cerevisiae* genome has allowed to solve the difficulty encountered in the intial induction experiments concerning the regulation of R gene expression. This was essential to promote the ARS1413 in specific phases of the cell cycle. Considering all these data, and the aim of the project the MVY171 strain was chosen to proceed with the following experiments that were designed to study the proteome of a eukaryotic replication origin.

3.17 THE ARS1413 PLASMID IS MAINTAINED INSIDE THE CELLS DURING GALACTOSE INDUCTION

Upon induction of the Recombinase gene ARS1413 is excised from its chromosomal context producing a chromatin plasmid. This episome contains the replication origin,

the R sites (A and B) and the *YNL210C* gene located between RsiteB and ARS1413. However it does not contain any centromeric region and this could result in loss of the chromatin plasmid during cell division. Thus before proceeding with the second part of the strategy, which consists of isolation of the plasmid from the rest of the genome, it was essential to confirm if the excised plasmid was maintained inside the cells. For this purpose a Southern blot analysis was performed on MVY171. To induce ARS1413 excision MVY171 cells were collected after 5 hours of induction, as this was the time point at which the recombination reached its maximum efficiency. The recovered DNA samples were restricted with PmeI and the ARS1413 probe that anneals within the ACS region of the late replication origin was used for the hybridization. Indeed this probe can detect, in addition to the 3.7 kbp fragment, the 1.3 kbp excised chromatin plasmid instead of the shorter PmeI digestion product (2.5 kbp). A schematic map of the region of interest on Chr. XIV is shown in Fig.53 A. The PmeI restriction sites and the region where the ARS1413 probe (green box) anneals are also reported. The results of the hybridization are reported in panel B of Fig.53; the label "5 hrs" represents the time point at which the samples were collected from both the uniduced and induced cultures. As indicated by the upper arrow, after 5 hours of growth in Raffinose the only band detectable by the probe was the 3.7 kbp fragment. No other signal could be seen confirming that no recombination had occurred in the absence of induction. This data confirms the previous results (Fig.53) demonstrating that the expression of the Recombinase gene can be completely controlled. In the adjacent lane where the DNA extracted from the cells induced in Galactose was loaded, two bands were detected; the sizes of these signals indicate that they both correspond to the ARS1413 episome (1.3 kbp). The presence of a doublet was explained by considering that the chromatin plasmid, being circular, does not migrate as a linear fragment. The signals might represent the two forms, circular and nicked, of the excised chromatin plasmid. However to confirm that the doublets observed in the Southern blot (Fig.53) represent different forms of the excised chromatin plasmid a DNA digestion with enzymes that cut within the ARS1413 episome should be performed.

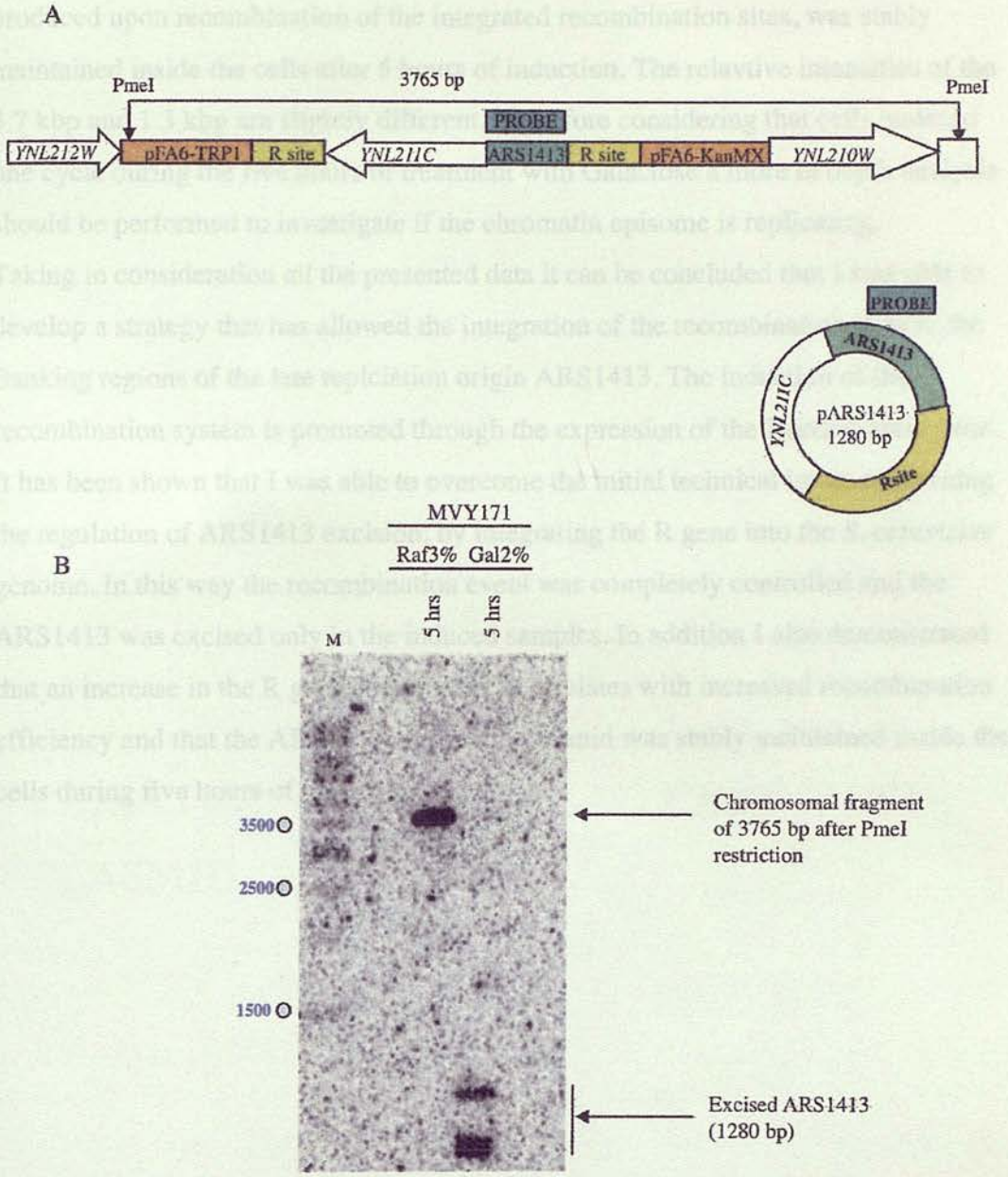


Fig.53 Demonstration that ARS1413 plasmid is maintained inside the cells during the induction. A) Map of the Chr. XIV region. The membrane was hybridized with the ARS1413 probe which detects both the 3765 bp fragment and the 1280 bp ARS1413 chromatin plasmid. B) Southern blot. Samples were collected from the Raffinose and Galactose cultures after five hours of induction; the arrows indicate the bands detected by the probe whilst the M represent the 1 kbp MW ladder. The results demonstrate that ARS1413 is maintained inside the cells during the five hours of induction.

It must be noticed that this experiment has not only reproduced the results of the induction reported in Fig.53 but it has also demonstrate that the ARS1413 plasmid, produced upon recombination of the integrated recombination sites, was stably maintained inside the cells after 5 hours of induction. The relavtive intensities of the 3.7 kbp and 1.3 kbp are slightly different. Therefore considering that cells undergo one cycle during the five hours of treatment with Galactose a more in depth analysis should be performed to investigate if the chromatin episome is replicating.

Taking in consideration all the presented data it can be concluded that I was able to develop a strategy that has allowed the integration of the recombination sites in the flanking regions of the late replciation origin ARS1413. The induction of the recombination system is promoted through the expression of the Recombinase gene. It has been shown that I was able to overcome the initial technical issue, concerning the regulation of ARS1413 excision, by integrating the R gene into the *S. cerevisiae* genome. In this way the recombination event was completely controlled and the ARS1413 was excised only in the induced samples. In addition I also demonstrated that an increase in the R gene copy number correlates with increased recombination efficiency and that the ARS1413 chromatin plasmid was stably maintained inside the cells during five hours of induction.

4 ISOLATION OF THE ARS1413 CHROMATIN PLASMID

The aim of this project was to develop an unbiased assay to identify proteins involved in DNA replication through the *in vivo* excision of a eukaryotic replication origin from its chromosomal context. The excised chromatin plasmid would then be isolated to analyze the proteins associated with the origin by mass spectrometry analysis.

In the previous chapter I have demonstrated the successful excision of the ARS1413 and the maintenance of the chromatin plasmid inside the cell during the induction. Therefore in this chapter the analyses performed to confirm if proteins were still bound to ARS1413 after recombination had occurred and the pivotal experiment to purify the excised origin are fully described.

4.1 CONSTRUCTION OF THE MVY201 AND MVY202 YEAST STRAINS

Thus once the recombination system had been functionally optimized in order to obtain a controlled and efficient excision of ARS1413 I proceeded with the isolation of the excised chromatin plasmid from the rest of the genome. To confirm that replication proteins were still bound to the ARS1413 origin, once it was excised, a Chromatin Immunoprecipitation (ChIP) assay was performed before proceeding with isolation of the 1.3 kbp chromatin plasmid. The binding of a protein that is involved in DNA replication and that it is known to directly interact with the replication origin was analyzed. It has been demonstrated that ORC complex subunits interact with specific sequences within the ACS region of the replication origins (Rao and Stillman, 1995). Moreover it is known that the Orc2 protein is the only one that remains bound to the replication origin throughout the replication process. Indeed after origin firing, while the Replication Complex moves along the chromosome with the replication fork, Orc2p remains associated with the origin. Therefore the binding of Orc2p to the excised ARS1413 was studied. A positive result for the association of the protein with the chromatin plasmid would support the idea to identify proteins involved in DNA replication by the isolation of an origin from the rest of the genome. For this analysis strain MVY200 that contains the 3X-FLAG tag integrated

at the 3'-end of the *ORC2* gene was constructed. MVY201 and MVY202 derive from disruption of the *BAR1* gene in strains MVY200 and MVY171 respectively.

4.2 CHROMATIN IMMUNOPRECIPITATION OF Orc2-3X-FLAGp

As mentioned at the beginning of this chapter I decided to study, by ChIP assay, the presence of the Orc2 protein on the excised ARS1413 to demonstrate that the excision of the replication origin did not interfere with the binding of the protein to the DNA. The previously constructed MVY201 and MVY202 strains were used for this analysis. The binding of the protein to the replication origin was assayed after the origin had been excised and also when no recombination had occurred, as a control. Thus the samples were collected from both the strains grown in Raffinose and in Galactose; induction in Galactose was performed for 3 hours. However to be sure that the carbon sources used did not interfere with the association of the proteins with the replication origin, the same strains grown in Glucose, were also analyzed.

4.3 ARS1413 EXCISION IN ChIP SAMPLES

To demonstrate that the result for the ChIP assay in the induced strains was due to the Orc2 protein bound to the excised ARS1413 an aliquot of cells was taken from each culture to confirm the efficiency of the recombination event. Therefore before proceeding with the ChIP analysis, the excision of the late replication origin was confirmed. The extracted DNA was restricted with PmeI and analyzed by Southern blot using the TRP1 probe for the hybridization. The map of Chr. XIV region, with the positions of the probe and the PmeI restriction sites is shown in Panel A of Fig.54. The PmeI digestion produces fragments of 3.7 kbp and 2.5 kbp depending on if ARS1413 had been excised from the chromosome or not.
The results of hybridization of the membrane with the TRP1 probe are reported in Fig.54 B. The samples collected from MVY201 and MVY202 were divided according to the growth media and loaded in the following order: Glucose, Raffinose and Galactose.

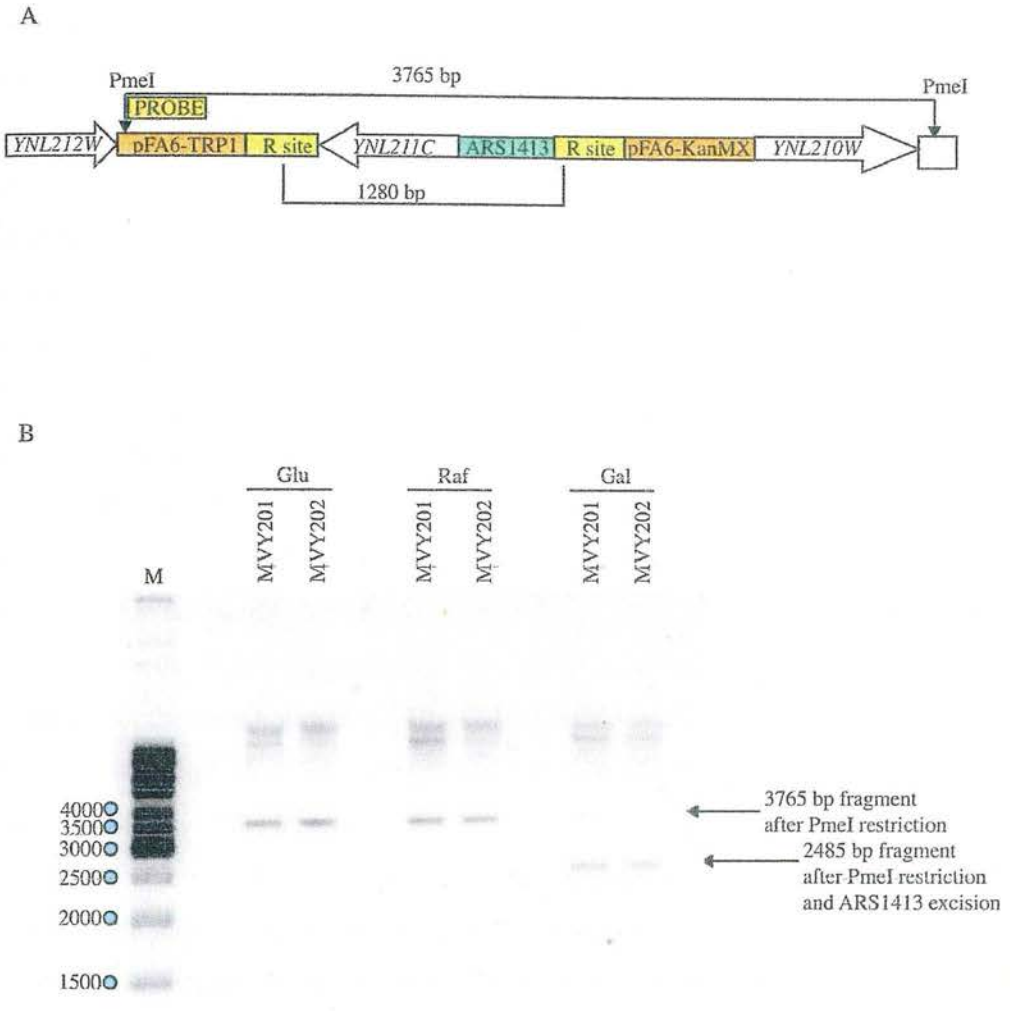


Fig.54 ARS1413 excision in ChIP samples. A) Schematic diagram of Chr. XIV where ARS1413 is located. Position of the TRP1 probe and size of PmeI digestion fragments are reported. B) Southern Blot on the DNA preparations from MVY201 and MVY202 grown in Glucose (Glu), Raffinose (Raf) and Galactose (Gal). M indicates the 1 kbp plus DNA MW ladder; some of the ladder sizes are shown. In the Glu and Raf samples only the 3765 bp band is detected (upper arrow). ARS1413 excision occurred only in the Gal samples, as indicated by the lower arrow (2485 bp); a faint 3765 bp band is also detected.

As indicated by the arrows (Fig.54 B) on the right edge of the picture only the fragment produced by the absence of ARS1413 excision is detected in MVY201 and MVY202 grown in Glucose and Raffinose.

This demonstrates that recombination between the inverted repeats has not occurred in these samples.

In the samples collected from the induced culture (Gal) the 2.5 kbp band produced only if the ARS1413 has been excised from its chromosomal context is identified; in the same samples a faint 3.7 kbp band was also detected by the TRP1 probe.

To quantify the percentage of cells in which the replication origin has been excised the intensities of each band were analyzed with the OptiQuant Image program. The histogram graph in Fig.55 shows the percentage of cells, for each strain, in which excision of ARS1413 has occurred.

In MVY201 and MVY202 grown in Glucose and Raffinose ARS1413 has not been excised, as also indicated by the hybridization results. Instead in the same strains induced in Galactose for three hours recombination has occurred in approximately 95% of the cells. This is in agreement with the presence of the faint 3.7 kbp band, visualized in the membrane (Fig.54B). Moreover these data confirm the results previously shown (Fig. 53 panel B) concerning ARS1413 excision in MVY171 in 3% Raffinose and 2% Galactose.

In conclusion it was demonstrated that in the samples that will be analyzed by ChIP the recombination has occurred only in the cells grown in Galactose. In addition ARS1413 was excised in close to 100% of the induced cells. Therefore any signal for Orc2p association with ARS1413, in the induced strain, would be indicative of the excised plasmid and not of the chromosomal ARS.

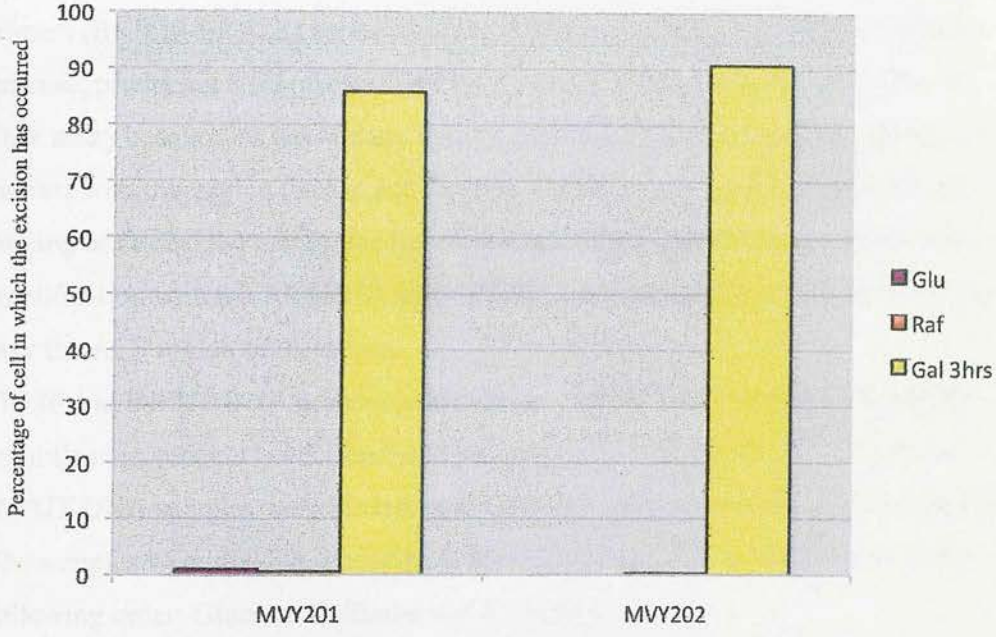


Fig.55 Quantification of ARS1413 excision in ChIP samples. Values are reported as a percentage of cells in which ARS1413 was excised. Recombination occurred in approximately 90% of the cells induced in Galactose for three hours (yellow bars).

4.4 Orc2-3X-FLAGp ChIP ON THE ARS1413 PLASMID

For the ChIP analysis proteins were crosslinked to DNA with formaldehyde; samples were then treated with α-FLAG antibody to recover the DNA regions associated with Orc2p (IP). An aliquot of each sample was not incubated with the α-FLAG antibody and thus it represents the whole DNA (INPUT).

The IP and INPUT samples were then analyzed by PCR using three different pairs of primers. The ARS1413 primers (newURA3F-ARS1413B) amplify a fragment of 381 bp over the ACS region of the origin of interest. The data obtained for amplification over ARS1413 was compared to a negative and a positive region to demonstrate the effective binding of the Orc2-3X-FLAG protein to the replication origin. The BRR2 primers (BRR2F-BRR2R) anneal over the *BRR2* gene, which encodes for a RNA helicase, producing a fragment of 240 bp. This gene is the negative control for the ChIP assay because, as has already been demonstrated, the Orc2 protein should not interact with this region (Bolon and Bielinsky, 2006). As a positive region for the binding of Orc2-3X-FLAGp, the late replication origin ARS603 was chosen. DNA amplification with the ARS603A and ARS603B primers produces a band of 161 bp over the ACS region of the origin.

The IP and INPUT DNA preparations were amplified by radioactive PCR and the amplification products were separated by electrophoresis (Fig.56 A). The IP and INPUT DNA samples were loaded on the left and right sides of the gel, respectively. The samples were divided according to the different growth conditions and in the following order: Glucose, Raffinose and Galactose.

The two samples in each of these groups are MVY201, which has the Orc2 protein tagged with the 3X-FLAG tag, and the untagged MVY202 strain. The upper (380 bp) and lower (161 bp) bands detected in each sample indicate the amplification products obtained with the ARS1413 and ARS603 pairs of primers respectively. The signals from MVY202 represent non specific background and show no enrichment of the ARS DNA compared to the negative control region (Fig.56 A, left panel).

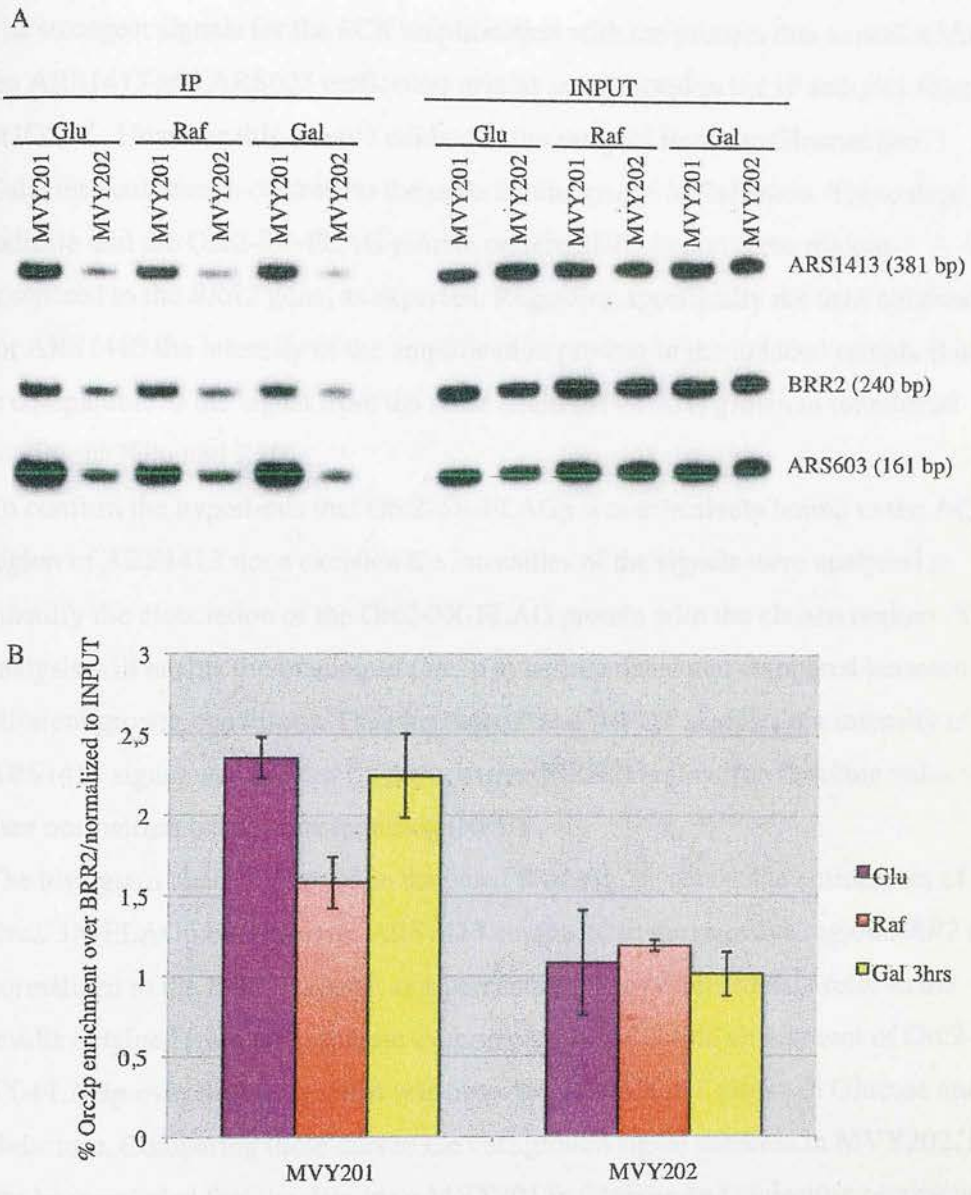


Fig.56 Orc2-3X-FLAGp ChIP. A) Radioactive PCR amplification of IP and INPUT samples from MVY201 and MVY202 grown in Glucose (Glu), Raffinose (Raf) and Galactose (Gal). The amplification products correspond to ARS1413, *BRR2* gene and ARS603, as indicated. B) Histogram chart showing the Orc2-3X-FLAGp enrichment at ARS1413 relative to *BRR2*, normalised to the INPUT. This data indicated that Orc2-3X-FLAGp was bound to the excised ARS1413.

The differences observed in the band intensities between the tagged (MVY201) and untagged (MVY202) strains demonstrate that the antibody specifically recognizes the FLAG epitope on the N-terminus of the Orc2 protein.

The strongest signals for the PCR amplification with the primers that anneal within the ARS1413 and ARS603 replication origins are detected in the IP samples from MVY201. However this is more evident in the samples from the Glucose and Galactose cultures in contrast to the same strains grown in Raffinose. These data indicate that the Orc2-3X-FLAG protein preferentially bound these regions compared to the *BRR2* gene, as expected. Regarding specifically the data obtained for ARS1413 the intensity of the amplification product in the induced sample (Gal) is comparable to the signal from the same strain (MVY201) grown in uninduced conditions (Glu and Raf).

To confirm the hypothesis that Orc2-3X-FLAGp was effectively bound to the ACS region of ARS1413 upon excision the intensities of the signals were analyzed to quantify the association of the Orc2-3X-FLAG protein with the chosen regions. This analysis will enable the binding of Orc2p to be calculated and compared between the different growth conditions. Thus for both IP and INPUT samples the intensity of ARS1413 signal was divided by the negative (BRR2) region; the resulting value was then normalized to the corresponding INPUT.

The histogram chart, presented in the panel B of Fig.56, shows the enrichment of Orc2-3X-FLAGp binding over ARS1413 compared to the negative region *BRR2* and normalized to the INPUT signal, as a percentage. The presented data refer to the results obtained from two separate experiments. A 2-2.5 fold enrichment of Orc2-3X-FLAGp over the background was observed in MVY201 grown in Glucose and Galactose. Comparing these data to the background signal detected in MVY202, it can be concluded that the data from MVY201 in Glucose and Galactose represent a real association of Orc2-3X-FLAGp with the studied late replication origin.

Moreover, the interaction between Orc2p and the ACS region of the replication origin of interest was the same independent of whether the ARS1413 had been excised (Gal samples Fig.56, panel B) or was still in its chromosomal context (Glu samples Fig.56, panel B). Only in the Raffinose sample the percentage of Orc2-3X-

FLAGp binding over ARS1413 was sligthly lower compared to the Glucose and Galactose sample. Therefore it has been demonstrated that Orc2-3X-FLAGp remains bound to ARS1413 when the origin was contained in the chromatin plasmid.

A graph showing Orc2p enrichment at ARS603, the positive control for ChIP analysis, is reported in Fig.57. IP values divided by the *BRR2* signal were normalized to the corresponding INPUT. The data, which refer to the mean of two separate experiments, validate the ChIP analysis. Indeed association of Orc2-3X-FLAGp is observed over ARS603 in the tagged strains compared to the untagged samples. The enrichment of Orc2-3X-FLAGp obtained in MVY201 grown in Glucose (purple bars Fig.57) is comparable to the one obtained in the induced sample (yellow bars Fig.57). Comparing the results reported in Fig. 55 and Fig.56 Orc2-3X-FLAGp shows a higher enrichment for ARS603 than ARS1413. This observation might simply reflect a difference in the interaction of Orc2p over the two different replication origins. Indeed a variation in ORC binding among origins of replication has been observed (Xu, *et al*, 2006).

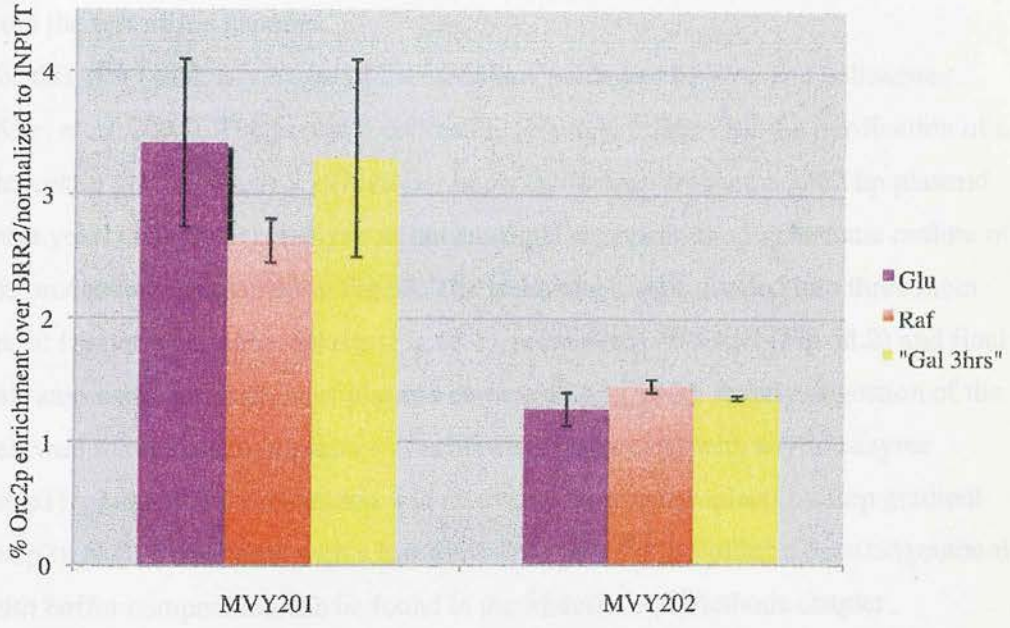


Fig.57Orc2-3X-FLAGp enrichment over ARS603. Histogram chart showing the enrichment of Orc2-3X-FLAGp at ARS603 in MVY201 and MVY202, in the three different growth conditions.

4.5 ISOLATION OF THE ARS1413 CHROMATIN PLASMID

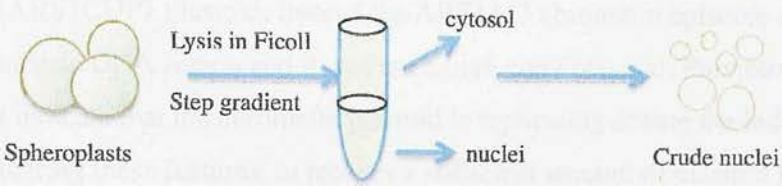
In the previous section a ChIP assay on MVY201 and MVY202 was described. This experiment was performed prior to purification of the chromatin plasmid to confirm that Orc2-3X-FLAGp, an ORC complex subunit that interacts with replication origins, was still bound to ARS1413 after it had been excised. Considering that the aim of this project was to isolate the ARS1413 episome to study the associated proteins it was essential to demonstrate that excision did not interfere with the protein-origin interaction. Given the positive results obtained with the Orc2-3X-FLAGp ChIP analysis (Fig.56 panels A and B) I could proceed with the following step of this unbiased assay, which consisted of the isolation of the chromatin plasmid from the rest of the genome.

For this aim I took advantage of the technique published by Kim and colleagues (Kim, *et al*, 2004). The protocol outlined in this report described the purification of a chromatin plasmid from *S. cerevisiae*; in particular they isolated a 2462 bp plasmid from yeast cells to then analyze its nucleosomal organization. A schematic outline of the procedure is presented in Fig.58. The technique can be divided into three main steps: formation of spheroplasts (Fig.58.1), preparation of nuclei (Fig.58.2) and final extraction and purification of plasmid chromatin (Fig.58.3). Briefly, digestion of the cell wall to obtain spheroplasts was achieved by treatment with a lytic enzyme (step1); crude nuclei preparation was recovered from spheroplasts by step gradient (step2). Nuclei treatment with a low ionic detergent (NLB buffer; a detailed protocol with buffer composition can be found in the Material and Methods chapter, paragraph 2.33) promoted their lysis and release of the chromatin plasmid that was recovered through a sucrose cushion (step 3). The sample is then filtered and purified to eliminate the sucrose.

Before proceeding with the extraction of the excised ARS1413 episomal plasmid, different requirements had to be taken into consideration in order to set up the right conditions for the isolation.

1. Spheroplasting

2. Preparation of nuclei



3. Extraction and purification of plasmid chromatin

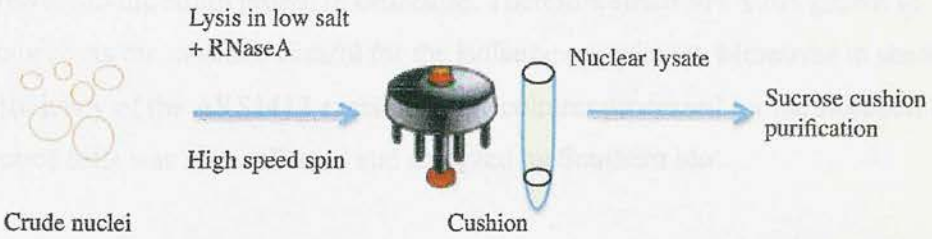
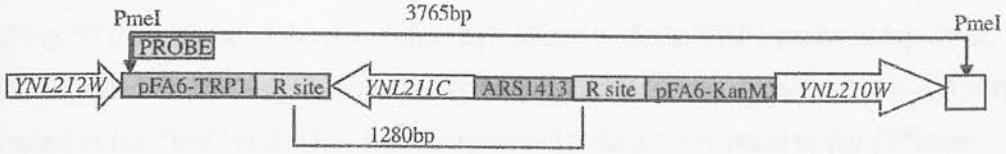


Fig.58 Schematic representation of the technique adopted for ARS1413 plasmid purification (adapted from Kim, *et al*, 2004). Cells are treated with a lytic enzyme to obtain spheroplasts (1); a crude nuclei preparation is then recovered by density gradient (2). Nuclei are lysed and the plasmid is recovered through sucrose cushion ultracentrifugation; plasmid is purified with centricon columns (3).

One of the most critical steps in the isolation was represented by the sucrose cushion centrifugation where the length of the spin is inversely proportional to the size of the plasmid. The small size of the ARS1413 plasmid, which is 1280 bp, increased the difficulty of the experiment especially regarding the length of the sucrose cushion spin. However using the information provided in the Kim, *et al* paper as a guide to the spin time of the ultracentrifugation (2.5 hours for a 2.5 kbp plasmid at 367000 g), the sample containing the ARS1413 chromatin plasmid was spun for 5 hours. After purification the plasmid has to be analyzed by mass spectrometry; therefore enough material had to be recovered from the isolation for the proteomic study. The protocol described in the paper was based on one liter of yeast culture grown to log phase to recover the high copy centromeric TRP1ARS1CUP1 plasmid. Instead the ARS1413 chromatin episome did not contain centromeric DNA region and it was not a high copy plasmid. Previous data (Fig.53) might indicate that the chromatin plasmid is replicating during the induction. Considering these features, to recover a sufficient amount of plasmid for the analysis of the associated proteins three liters of culture were prepared.
The isolation was performed on MVY201 that was separately grown in Raffinose and Galactose. In the previous chapters it has been demonstrated that the excision only occurred in the induced samples; indeed the ARS1413 plasmid was detected exclusively in the strain grown in Galactose. Therefore strain MVY201 grown in Raffinose was the negative control for the isolation experiment. Moreover to show the efficiency of the ARS1413 excision in the cultures processed for the isolation an aliquot of cells was also collected and analyzed by Southern blot.

4.6 CONFIRMATION OF ARS1413 EXCISION BEFORE PLASMID PURIFICATION

For the analysis of ARS1413 excision, DNA was extracted from the collected samples and digested with PmeI that, as described before, produces fragments of 3.7 kbp and 2.5 kbp depending on the presence or not of the replication origin on Chr. XIV rspectively (Fig.59 A).



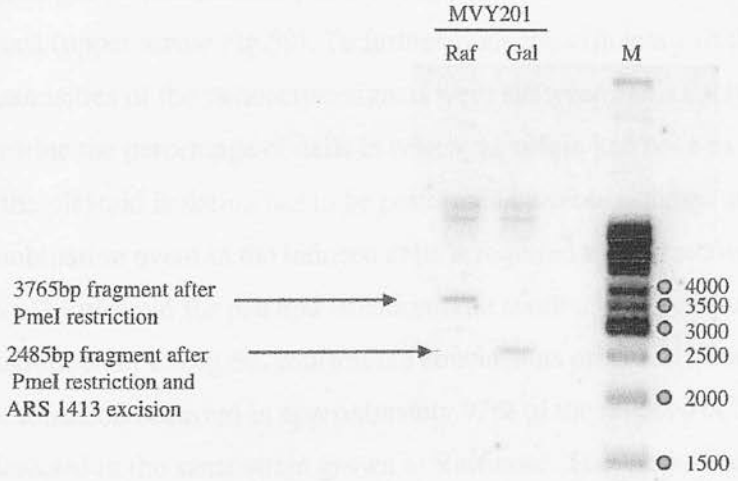


Fig.59 ARS1413 excision in samples treated for plasmid isolation. A) Schematic diagram of ARS1413 region; positions of TRP1 probe and PmeI sites are indicated. B) Southern blot on MVY201 grown in Raffinose (Raf) and Galactose (Gal). M indicates the 1 kbp plus DNA MW ladder; the arrows point to the 3765 bp and 2485 bp fragments detected in the Raf and Gal samples respectively. ARS1413 excision occurred exclusively in cells induced in Gal for five hours.

Southern blot analysis was performed on the digested samples and the membrane was hybridized with the TRP1 probe. In Fig.59 A a schematic representation of the Chr. XIV region containing ARS1413, is reported; the TRP1 probe (orange box) anneals within the tryptophan gene adjacent to the RsiteB. The position of the probe allowed the detection of both the PmeI chromosome digestion fragments. In panel B of Fig.59 the picture of the membrane hybridized with the TRP1 probe is reported. The digested DNA samples collected from the Raffinose and Galactose cultures were loaded in the "Raf" and "Gal" lanes respectively; the arrows point to the different bands detected in the two samples.

In the induced sample (Gal) the only visible band, except for the crosshybridization of the probe with the endogenous *TRP1* gene, is the 2.5 kbp fragment that demonstrates the excision of ARS1413. Instead the fragment produced if the late replication origin was not excised from its chromosomal context is detected only in the cells grown in Raffinose, confirming that in this condition recombination was not induced (upper arrow Fig.59). To further study the efficiency of the recombination the intensities of the radioactive signals were analyzed. This analysis is required to determine the percentage of cells in which the origin had been excised. Considering that the plasmid isolation has to be performed on these samples an efficient recombination event in the induced cells is required to then recover enough ARS1413 plasmid for plasmid isolation. The results, which are reported in the histogram chart in Fig.60, confirm the conclusions previously drawn. Whilst recombination occurred in approximately 97% of the induced cells no excision could be detected in the same strain grown in Raffinose. These results confirm the hypothesis that the episome plasmid should be isolated exclusively from the induced cells; the 97% of excision in MVY201 grown in Galactose should be indicative of a high probability of recovering the chromatin plasmid. The absence of Recombination in the uninduced samples supports the choice of MVY201 grown in Raffinose as the negative control for the experiment.

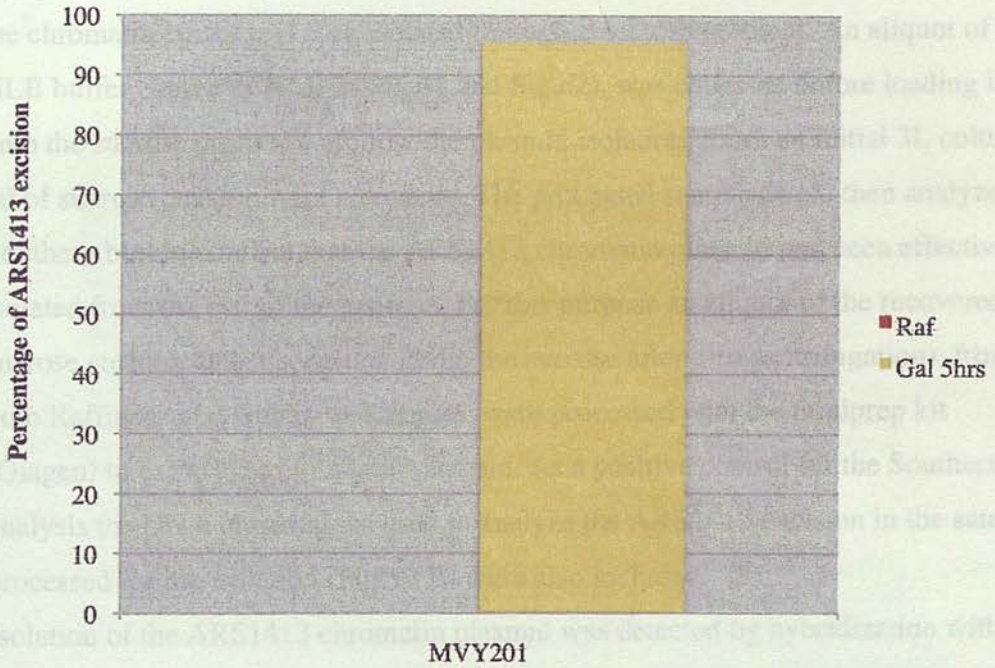


Fig.60 Quantification of ARS1413 excision in induced (Gal 5 hrs) and uninduced (Raf) MVY201. The late replication origin is excised in approximately 95% of the cells induced in 2% Galactose for five hours.

4.7 DEMONSTRATION OF ARS1413 PLASMID ISOLATION FROM THE REST OF THE GENOME

After demonstrating ARS1413 excision in MVY201 (Fig.59) I proceeded with isolation of the chromatin plasmid from the yeast cells. As mentioned in the previous paragraph (4.6) the cells were treated with a lytic enzyme to promote spherolplast formation, which was confirmed by microscopy analysis (data not shown). The crude nuclei preparation, obtained by step gradient analysis, is treated with a low ionic detergent buffer (NLB buffer, Fig.58 step 3; see paragraph 2.33 for NLB buffer composition) that lyses the nucei and release the chromatin plasmid in the supernatant (Kim, *et al*, 2004). The nuclear debris is removed by centrifugation and the chromatin episome is then isolated through a sucrose gradient. An aliquot of the NLB buffer (stated as NLB in Fig.61 and Fig.62), was collected before loading it onto the sucrose cushion to follow the plasmid isolation. From an initial 3L colture 4 ml of sucrose cushion were recovered. The processed samples were then analyzed by Southern blot to confirm that the ARS1413 chromatin plasmid had been effectively isolated from the rest of the genome. For this purpose an aliquot of the recovered sucrose cushion and supernatant (from the sucrose after ultracentrifugation), from both Raffinose and Galactose samples, were processed with the miniprep kit (Qiagen) to extract the chromatin plasmid. As a positive control for the Southern analysis the DNA preparations used to analyze the ARS1413 excision in the samples processed for the isolation (Fig.59 B) were also included.

Isolation of the ARS1413 chromatin plasmid was detected by hybridization with the ARS1413 probe. The Chr. XIV region containing the replication origin and the excised ARS1413 chromatin plasmid are presented in panel A of Fig.61. The green box indicates the position of the ARS1413 probe that anneals within the replication origin allowing detection of both the chromatin plasmid and the 3.7 kbp genomic fragment obtained after PmeI digestion. The plasmid isolation was performed twice and the picture reported in Fig.61 B is representative of one of the two experiments.

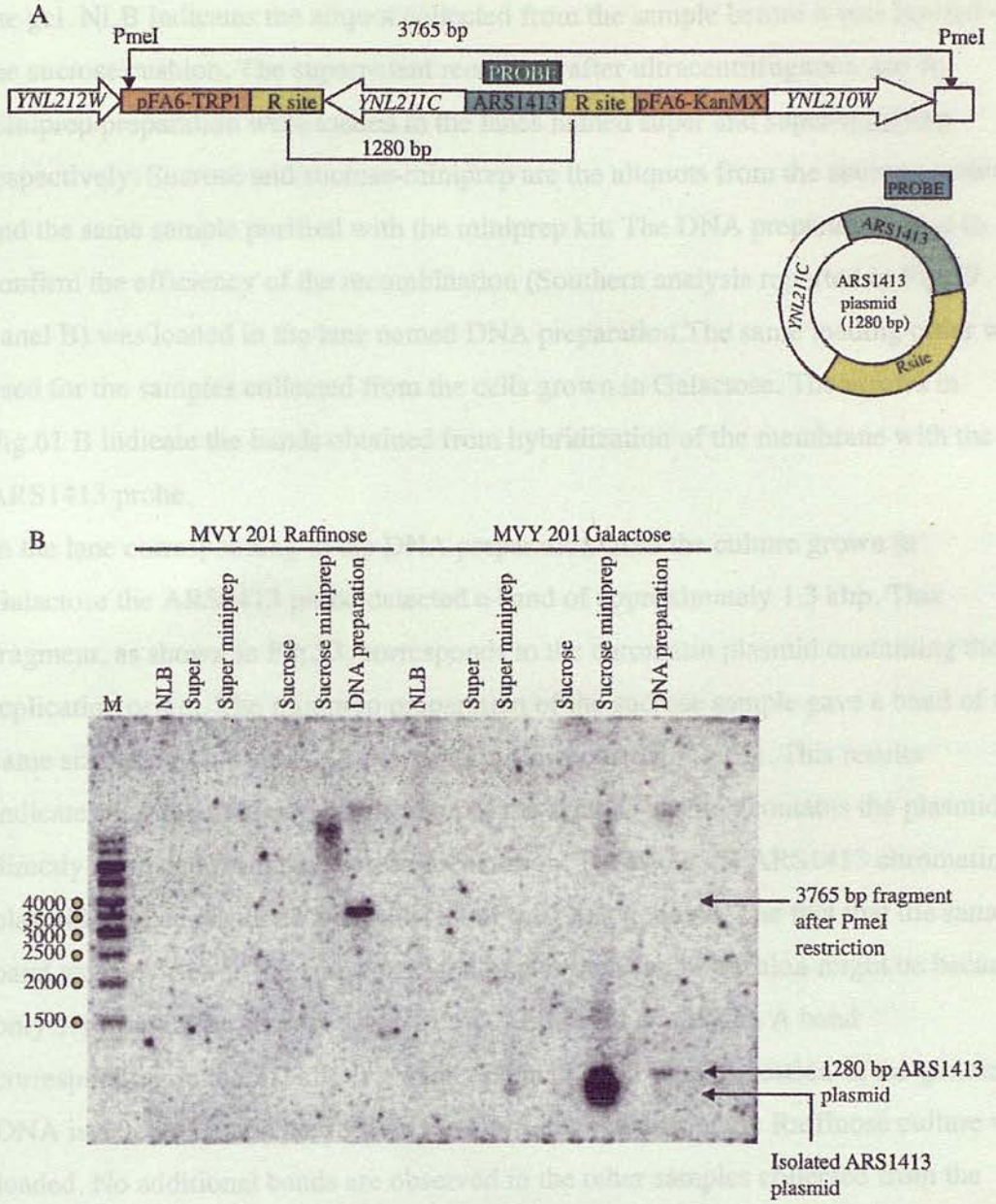


Fig.61 Isolation of the ARS1413 chromatin plasmid. A) Diagram of ARS1413 genomic region; the position of the ARS1413 probe and PmeI sites are indicated. The ARS1413 chromatin plasmid is also shown. B) Southern analysis to confirm isolation of the chromatin plasmid from the rest of the genome. Aliquots from the NLB buffer (20 µl), the sucrose cushion (40 µl) and the supernatant after ultracentrifugation (40 µl) were loaded. The super miniprep and sucrose miniprep lanes indicate the sucrose and superntant samples treated with the Qiagen kit for plasmid isoaltion. The arrows indicate the different bands detected by the ARS1413 probe; the two fragments in the sucrose miniprep lane, from the induced sample, are the two structural forms of the ARS1413 plasmid (1.3 kbp).

The samples derived from the Raffinose culture were loaded in the first part of the gel. NLB indicates the aliquot collected from the sample before it was layered on the sucrose cushion. The supernatant recovered after ultracentrifugation and its miniprep preparation were loaded in the lanes named super and super-miniprep respectively. Sucrose and sucrose-miniprep are the aliquots from the sucrose cushion and the same sample purified with the miniprep kit. The DNA preparation used to confirm the efficiency of the recombination (Southern analysis reported in Fig.59 panel B) was loaded in the lane named DNA preparation.The same loading order was used for the samples collected from the cells grown in Galactose. The arrows in Fig.61 B indicate the bands obtained from hybridization of the membrane with the ARS1413 probe.

In the lane corresponding to the DNA preparation from the culture grown in Galactose the ARS1413 probe detected a band of approximately 1.3 kbp. This fragment, as shown in Fig.53, corresponds to the chromatin plasmid containing the replication origin. The miniprep preparation of the sucrose sample gave a band of the same size as that for the DNA preparation (lower arrow Fig.61). This results indicates that the miniprep preparation of the sucrose cushion contains the plasmid directly extracted from the sucrose preparation. Therefore the ARS1413 chromatin plasmid has been isolated from the rest of the yeast genome. The fact that the same band was not seen in the lane corresponding to the sucrose cushion might be because only 1/100 th of the sucrose sample could be loaded on the gel. A band corresponding to the 3.7 kbp fragment obtained upon PmeI digestion of the genomic DNA is detected in the lane where the DNA extracted from the Raffinose culture was loaded. No additional bands are observed in the other samples collected from the same culture. All together these data demonstrate that the modifications applied to the original protocol have led to the isolation of the 1.3 kbp ARS1413 chromatin plasmid. Moreover the results reported in Fig.61 B indicate that the isolation occurred exclusively in the MVY201 strain induced in Galactose for five hours. The ARS1413 probe was previously used in the Southern performed to demonstrate that the chromatin plasmid was maintained inside the cells after induction (Fig.53). The data obtained also indicate that the probe did not recognize any other genomic

regions. From the results reported in Fig.61 panel B it could not be excluded that other genomic regions were isolated with the plasmid. To exclude this possibility the same membrane was hybridized with a different probe.

To eliminate the radioactive signals the membrane was stripped and only when no radioactive signals were present it was hybridized with the labeled TRP1 probe.

Panel A of Fig.62 shows the genomic region where the TRP1 probe anneals.This probe allows the detection of the 3.7 kbp and 2.5 kbp fragments, obtained after PmeI digestion of Chr. XIV containing or not the ARS1413 origin, respectively.

The picture of the membrane after the hybridization is shown in panel B (Fig.62 B); the arrows on the right edge of the picture, indicate the bands visualized on the membrane.

As expected the TRP1 probe detected the 3.7 kbp and 2.5 kbp fragments in the lanes where the DNA from the Raffinose and Galactose cultures were loaded, respectively. This is in agreement with the results reported in Fig.59 panel B. However it must be noticed that a signal is present in the sucrose miniprep preparations from both the Raffinose and Galactose samples. These data indicate that the isolated sample does not contain only the ARS1413 plasmid but also other genomic regions, such the *TRP1* gene. All together these data demonstrate that the adjustments made to the published protocol (Kim, *et al*, 2004) have allowed the isolation of the 1.3 kbp chromatin plasmid containing the late replication origin ARS1413. However this technique did not allow to exclusively purify the episome because residual chromosomal chromatin was still present in the sample after isolation.

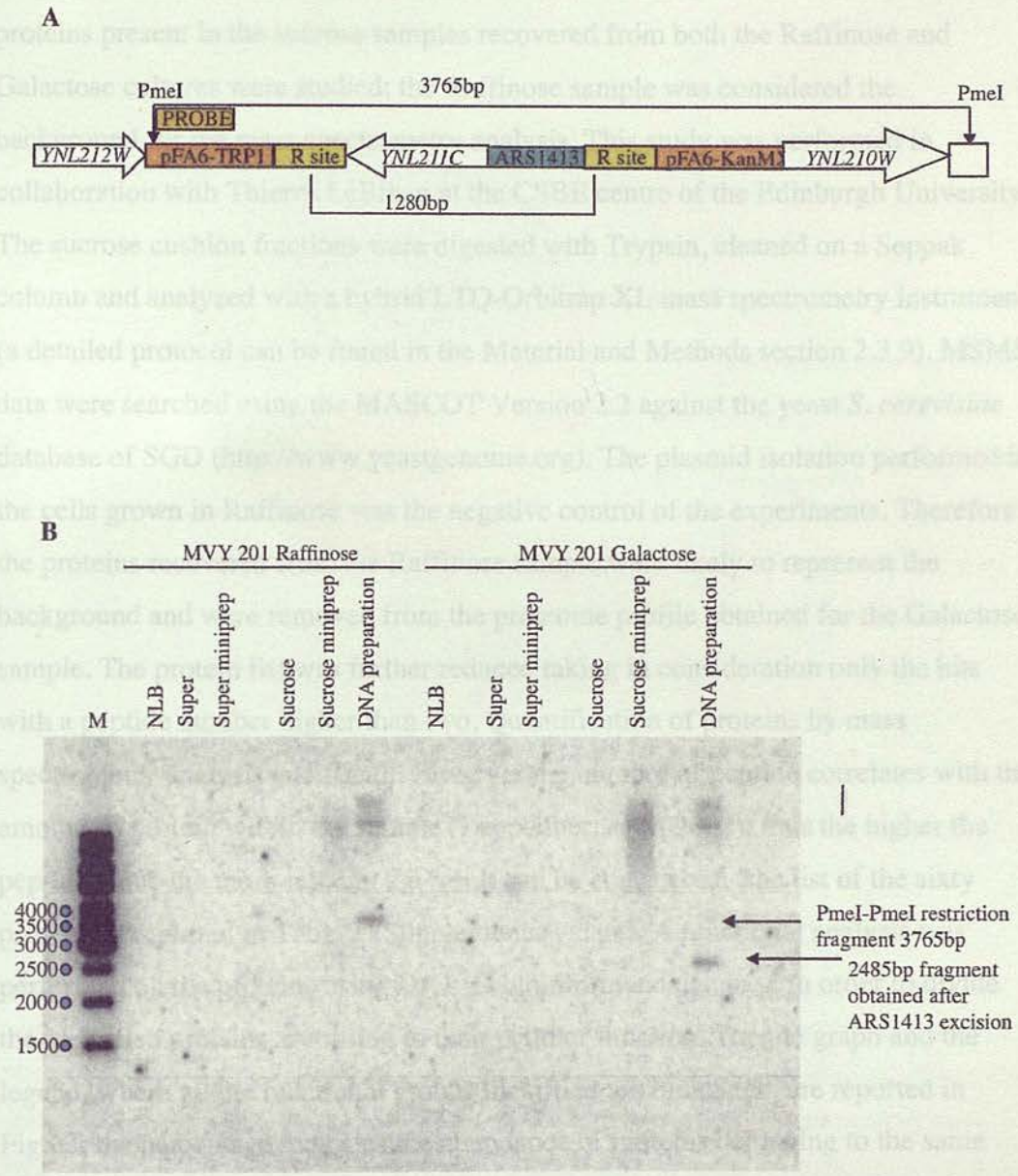


Fig.62 TRP1 hybridization of the samples recovered from the plasmid isolation. A) Schematic representation of the ARS1413 region; position of the TRP1 probe is indicated. B) The same membrane (Fig.61) dehybridized and hybridized with TRP1 probe. The presence of the 3.7 kbp and 2.5 kbp fragments only in the DNA preparations samples demonstrate that no other genomic regions homologous to *TRP1* are isolated with the plasmid.

4.8 MASS SPECTROMETRY ANALYSIS AND FUTURE PERSPECTIVES

After purification of the ARS1413 plasmid, the final step of this project consisted of mass spectrometry analysis of the isolated replication origin. For this purpose the proteins present in the sucrose samples recovered from both the Raffinose and Galactose cultures were studied; the Raffinose sample was considered the background for the mass spectrometry analysis. This study was performed in collaboration with Thierry LeBihan at the CSBE centre of the Edinburgh University. The sucrose cushion fractions were digested with Trypsin, cleaned on a Seppak column and analyzed with a hybrid LTQ-Orbitrap XL mass spectrometry instrument (a detailed protocol can be found in the Material and Methods section 2.3.9). MSMS data were searched using the MASCOT Version 2.2 against the yeast *S. cerevisiae* database of SGD (http://www.yeastgenome.org). The plasmid isolation performed in the cells grown in Raffinose was the negative control of the experiments. Therefore, the proteins recovered from the Raffinose sample were likely to represent the background and were removed from the proteome profile obtained for the Galactose sample. The protein list was further reduced taking in consideration only the hits with a peptide number higher than two. Quantification of proteins by mass spectrometry analysis is difficult. However the number of peptide correlates with the amount of protein within the sample (Rappsilber, *et al*, 2002); thus the higher the peptide value the more reliable the result can be considered. The list of the sixty proteins is reported in Table 1 (Supplementary data). A functional analysis was performed on the proteins using DAVID bioinformatic database in order to divide the identified proteins according to their cellular function. The pie graph and the legend, where all the functional groups identified are indiciated, are reported in Fig.63; the percentage indicates the abundance of proteins belonging to the same group.

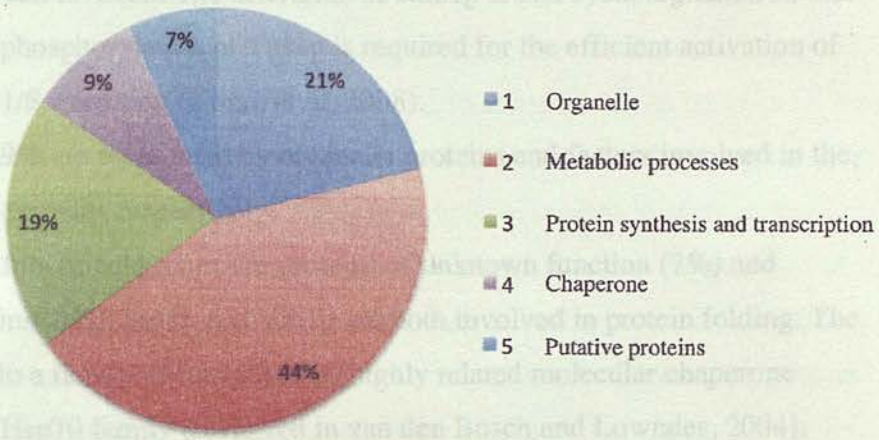


Fig.63 Mass spectrometry of the isolated ARS1413 chromatin plasmid. Functional analysis was performed on the proteins recovered in the induced sample using the DAVID bioinformatic program. The different functional groups are reported in the pie graph according to the percentage of proteins in each group; the cellular pathways in which the recovered proteins are involved are reported in the legend. No replication proteins were identified with the mass spectrometry analysis.

As can be seen most of the recovered factors are involved in metabolic pathways (44%). Tsa1p and Rho1p, which belong to the metabolic functional group, are among the proteins with the higher peptide (Table 1, supplementary data). Tsa1p, an ubiquitous peroxidase, is involved in the antioxidative stress response protecting genome stability and preventing large chromosomal rearrangements (Huang and Kolodner, 2005). Rho1p belongs to the GTP binding proteins and plays a pivotal role on asymmetric cell division. The activation of Rho1p is cell cycle regulated so that Cln2p/Cdc28p phosphorylation of Tus1p is required for the efficient activation of Rho1p at the G1/S transition (Kono, *et al*, 2008).

The 21% and 19% are repesented by organelle proteins and factors involved in the biosynthesis of proteins respectively.

The remaining functional groups are proteins of unknown function (7%) and chaperon proteins (9%). Ssa1p and Ydj1p are both involved in protein folding. The former belong to a family of four (Sss1-4) highly related molecular chaperone proteins of the Hsp70 family (reviewed in van den Bosch and Lowndes, 2004). Ssa1p and Ssa2p are subunits of the large Rad9 complex; it has been speculated that Ssa1p and Ssa2p may facilitate the Mec1-dependent remodeling of the large complex (reviewed in van den Bosch and Lowndes, 2004). Ydj1p was identified in a genetic screen for synthetic lethal mutants with Polδ (Chanet and Heude, 2003). Therefore, it could be that DNA polymerase δ, as well as other replication proteins may require chaperone to reach their native conformation (Chanet and Heude, 2003).

Unfortunately none of the proteins identified by mass spectrometry analysis are involved in DNA replication, such as the ORC and MCM complexes, or a components of chromatin, such as histone proteins.

4.9 CONCLUSION

The aim of the project was the proteomic analysis of a late replication origin. The strategy designed to engineer a suitable yeast strain has allowed the construction of a yeast strain containing two pairs of recombination sites flanking ARS1413. Difficulties were encountered during the induction of recombination leading to the

inability to control the expression of the R gene. These problems were bypassed by integrating the R gene under the Gal promoter within the yeast genome. Indeed in the constructed strains MVY171 and MVY172 the efficency of ARS1413 excision increased reaching almost the 100 % in the induced cells (Fig. 51 and 52).

The chromatin episome was isolated from the rest of the genome taking advantage of the technique described by Kim and colleagues (Kim, *et al*, 2004). However several techincal issues relating to the purification, such as volume and consistency of the sucrose samples, have to be taken in consideration. The recovered sucrose cushion could not be filtered or cleaned up with the centricon columns due to consequent loss of material. This could have been due, as also specified by Kim and colleagues (Kim, *et al*, 2004) to the presence of contaminant chromosomal regions apart from the plasmid DNA. Moreover the three liter samples of culture used for the experiment described in this chapter were split into one liter sample before proceeding with the plasmid purification; however after the ultracentrifugation step the sucrose samples were combined. In this way I could have increased the amount of chromosomal contaminants in the sucrose cushion impairing the filtering process and the sample purification. As a consequence the mass spectrometry analysis was performed directly on the sucrose cushion. However both the presence of the sucrose and the high volume of the sample (approximately 4 ml) impaired efficient mass spectrometry analysis. It is essential to overcome these technical problems in order to reduce the volume of the sample and to purify the plasmid from the sucrose, which interferes with the subsequent proteome analysis. Therefore it can be concluded that the adopted protocol requires further modifications to improve the plasmid preparation and the subsequent mass spectrometry analysis.

5 PROTEOMICS OF STALLED REPLICATION FORKS

5.1 AIM OF THE PROJECT

It is essential during cell division that daughter cells receive a complete copy of the genome from the mother cell. This is achieved by the coordinate duplication of the entire genetic material during S-phase and segregation of the produced sister chromatids. DNA replication initiates from specific sites, called replication origins, which are scattered throughout the chromosomes. Assembly of pre-Replication Complexes (pre-RCs) licenses origins for replication while initiation of replication is triggered by CDK and DDK, which promotes transition from the pre-RC to the RC. Especially in budding yeast, replication origins are activated according to a temporal order (Raghuraman, *et al*, 2001) and thus they can be distinguished as early or late consistent with their time of activation (Ferguson, *et al*, 1991; Yabuki, *et al*, 1992). Replication timing is not an intrinsic property of the origins (Ferguson, *et al*, 1992) but rather depends on trans-acting factors and on chromatin structure. Indeed histone modifications have been shown to influence origin activation. Depletion of Rpd3p, a histone deacetylase, results in earlier firing of late origins due to premature association of Cdc45p (Vogelauer, *et al*, 2002) and DNA polymerase ε (Aparicio, *et al*, 2004). Control of replication origin activation is essential to ensure the genome is completely replicated before completion of the S-phase. Moreover the existence of a temporal program is particularly important in the presence of DNA damage. Indeed activation of the S-phase checkpoint inhibits late origin firing (Tercero, *et al*, 2003). However it has been shown that in *Δrpd3* mutants treated with HU, late origins can escape checkpoint control (Aparicio, *et al*, 2004) and this is not a consequence of failure to activate the checkpoint response. The molecular mechanisms that allow late origins to escape the S-phase checkpoint control in *Δrpd3* cells have still to be understood. It might be that components of the replication complex that have escaped identification are involved in this process.

Therefore the aim of the project was to develop an unbiased assay to study, *in vivo*, the proteomics of stalled replication forks. Through immunoprecipitation of Cdc45p and mass spectrometry analysis of the associated proteins I would be able to identify the proteins associated with the replisome complex. The analysis would be

performed in both WT and *Δrpd3* strains treated with HU, to compare fork proteomes in the absence and in the presence of late origin firing respectively. Moreover comparing the protein profiles obtained from the mass spectrometry analysis on WT strains treated or not with HU would lead to the identification of the proteins required for stabilization and for normal progression of replication forks. Discovery of different protein patterns and of new components of the replisome machinery could lead to the unveiling of mechanisms that regulate replication fork stabilization in response to replication stress. The results would then have to be validated to demonstrate the role of the proteins indentified in DNA replication.

5.2 EXPERIMENTAL STRATEGY

The aim of this project was the identification of proteins associated with replication forks in the presence and in the absence of replication stress. The S-phase checkpoint is a surveillance mechanism that cells have evolved to deal with the presence of abnormal structures or DNA damage during DNA replication. In response to S-phase checkpoint activation, the replication rate is decreased and forks are stalled. This can also be achieved experimentally with HU, which depletes the pool of dNTPs, slows down S-phase and induces stalling of replication forks. In these conditions a stable association between the replisome components is essential to avoid fork collapse (Lopes, *et al*, 2001). Another hallmark of S-phase checkpoint activation is the inhibition of late origin firing (Fig.64 A) (Santocanale and Diffley, 1998). However in cells mutant for the *RPD3* gene and incubated with HU, late origins can escape the S-phase checkpoint control (Fig.64 B) (Aparicio, *et al*, 2004).

It has been shown that in *Δrpd3* mutants the activation of late replication origins in presence of the checkpoint is correlated with an earlier binding of Cdc45p (Aparicio, *et al*, 2004). However it could also be that other proteins are involved. Thus it would be interesting to study if different factors are associated with stalled replication forks in WT and *Δrpd3* strains. Indeed most of the proteins that take part in replisome assembly have been discovered but it is likely that others have still to be identified. To elucidate the mechanisms that regulate origin firing in the presence of replication stress I developed an unbiased assay to study, *in vivo*, the proteome of stalled replication forks. Cdc45p is required for initiation (Zou, *et al*, 1997) and elongation (Tercero, *et al*, 2003) of DNA replication. It moves with the replication fork and remains associated with the replisome even when forks are stalled (Fig.64 A and B). The presented project consisted of the *in vivo* immunoprecipitation of Cdc45p from cells arrested in S-phase with HU (Fig.64 C). The technique adopted allowed separation of chromatin bound and unbound proteins in order to specifically pull down the Cdc45p fraction associated with replication forks. The recovered material was then analyzed by mass spectrometry (Fig.64 C) to identify the proteins that are associated with Cdc45p.

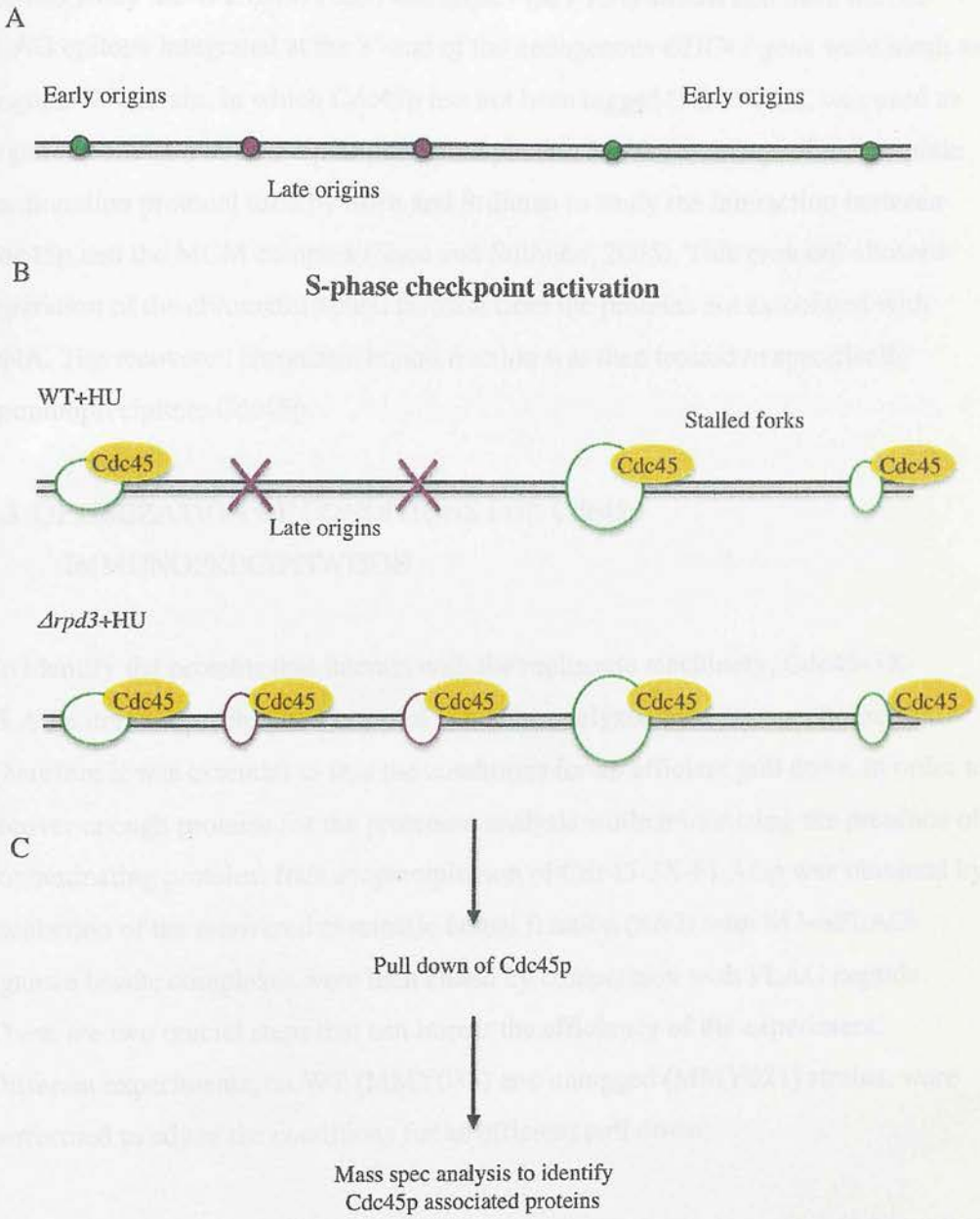


Fig.64 Schematic overview of the project. A-B) Representation of replication origin firing in the presence of HU in WT and Δrpd3 strains. Activation of the S-phase checkpoint inhibits late origin firing (A). However in *Δrpd3* mutants treated with HU late origins can escape checkpoint inhibition (B). To study the proteomes of stalled forks Cdc45p bound to chromatin was immunoprecipitated and the recovered proteins were analyzed by mass spectrometry.

This analysis was performed on WT and *Δrpd3* strains treated with HU and on WT strain in the presence or in the absence of the S-phase checkpoint.
For this study the WT (MMY033) and *Δrpd3* (MVY51) strains that have the 3X-FLAG epitope integrated at the 3'-end of the endogenous *CDC45* gene were used; an isogenic WT strain, in which Cdc45p has not been tagged (MMY021), was used as negative control. For Cdc45p immunoprecipitation I took advantage of a chromatin fractionation protocol used by Sheu and Stillman to study the interaction between Cdc45p and the MCM complex (Sheu and Stillman, 2006). This protocol allowed separation of the chromatin bound fraction from the proteins not associated with DNA. The recovered chromatin bound fraction was then treated to specifically immunoprecipitate Cdc45p.

5.3 OPTIMIZATION OF CONDITIONS FOR Cdc45p IMMUNOPRECIPITATION

To identify the proteins that interact with the replisome machinery, Cdc45-3X-FLAGp immunoprecipitated material would be analyzed by mass spectrometry. Therefore it was essential to find the conditions for an efficient pull down in order to recover enough proteins for the proteome analysis while minimizing the presence of contaminating proteins. Immunoprecipitation of Cdc45-3X-FLAGp was obtained by incubation of the recovered chromatin bound fraction (SN2) with M2-αFLAG-agarose beads; complexes were then eluted by competition with FLAG peptide. These are two crucial steps that can impair the efficiency of the experiment. Different experiments, on WT (MMY033) and untagged (MMY021) strains, were performed to adjust the conditions for an efficient pull down.

5.4 Cdc45-3X-FLAGp IMMUNOPRECIPITATION IN WT AND *Δrpd3* STRAINS

The Cdc45-3X-FLAGp immunoprecipitation was performed on WT and *Δrpd3* strains to compare the proteomes of stalled replication forks in the two strains. As a negative control the experiment was carried out, in parallel, in the untagged strain.

Immunoprecipitation was performed in strains arrested in S-phase by HU treatment (description of HU treatment is reported in the Material and Method section, paragraph 2.36). The samples from the three strains were then analyzed by mass spectrometry to identify the proteins associated with Cdc45-3X-FLAGp.

5.5 DETECTION OF Cdc45 IN IMMUNOPRECIPITATED SAMPLES

Before proceeding with the mass spectrometry analysis a western blot analysis was performed on the immunoprecipitated samples to confirm the presence of Cdc45-3X-FLAGp by hybridization with αFLAG antibody. The results of the experiment are reported in Fig. 65. The Flow Through (FT), which represents the proteins not bound to the beads, the IP (eluted proteins) and the B (proteins that remained bound to the beads) samples from WT and *Δrpd3* strains were loaded on the same gel (Fig.65 A). The samples derived from the analysis of the untagged strain (MMY021) were loaded on a separate gel (panel B, Fig.65).

Cdc45-3XFLAGp was present in the IP and B fractions in both WT and *Δrpd3* strains, as indicated by the arrow. As expected pull down of Cdc45-3XFLAGp could not be observed in any samples from the untagged strain.

These data confirmed that I was able to pull down Cdc45-3XFLAGp exclusively in WT and *Δrpd3* strains. However from this analysis I could not detect if there were differences between the WT and *Δrpd3* strains concerning the proteins associated with Cdc45-3XFLAGp at stalled replication forks.

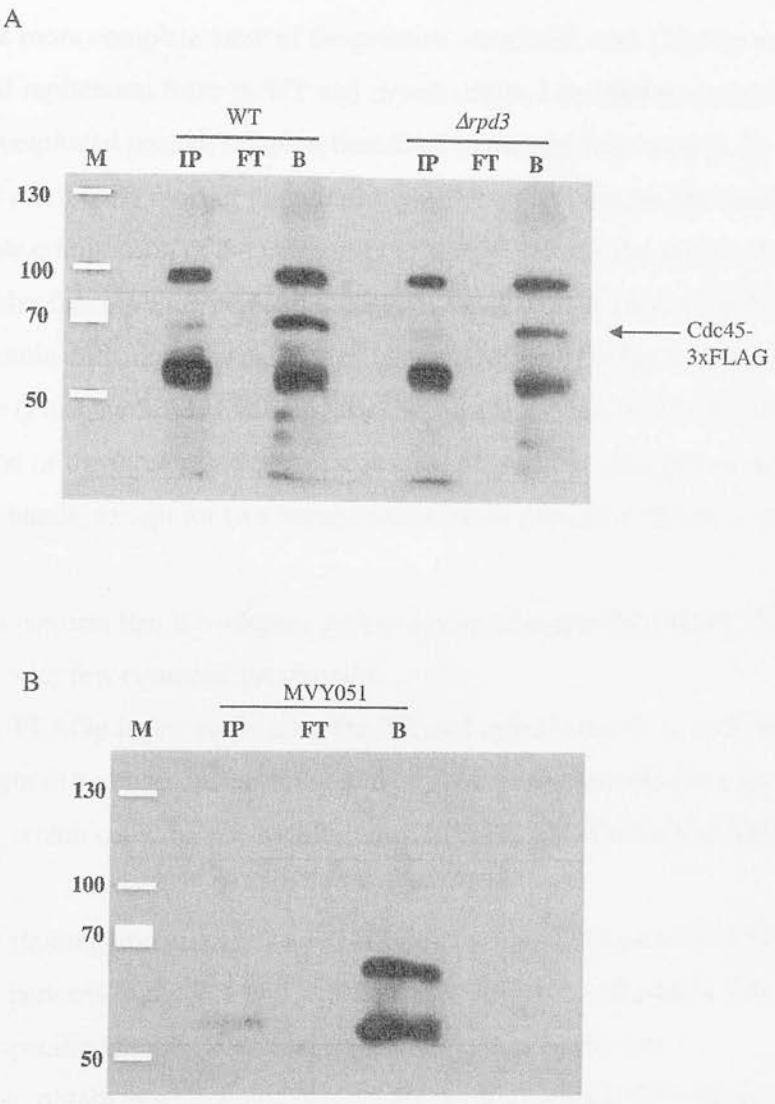


Fig.65 Cdc45-3X-FLAGp immunoprecipitation in WT, *Δrpd3* and untagged (MVY51) strains arrested by HU treatment. Membranes were hybridized with αFLAG antibody to detect Cdc45-3X-FLAGp (80 kDa). IP (immunoprecipitated proteins), Flow Through (FT; proteins not bound to the beads) and B (proteins that remained bound to the beads) samples were loaded for each strain. Sizes of protein ladder (M) bands are indicated on the side of each gel (kDa). A) Western Blot of WT and *Δrpd3* samples; the arrow points to the Cdc45-3X-FLAGp band that is present in both samples. B) Samples from the untagged strain; no Cdc453X-FLAGp pull down is observed in this strain.

5.6 SILVER STAINING ON IMMUNOPRECIPITATED SAMPLES

To obtain a more complete view of the proteins associated with Cdc45p and thus with stalled replication forks in WT and *Δrpd3* strains I decided to analyze the immunoprecipitated protein samples, described in the previous section, by silver staining. By detecting most of the proteins present in the samples, this analysis would allow comparison of the proteomes of stalled forks in the strains of interest. Thus samples from WT, *Δrpd3* and untagged strains were loaded on an SDS-page gel and protein detection was performed using the silver staining kit from GE Healthcare (PlusOne Silver Staining kit). The results are reported in Fig.66. Comparison of the three profiles shows that staining of the untagged samples did not detect any bands, except for two background signals present in all the samples (Fig.66).

These data confirm that the adopted protocol ensured a specific Cdc45-3X-FLAGp pull down with few contaminant proteins.

Cdc45-3X-FLAGp is present in both the WT and *Δrpd3* strains, as indicated by the arrow. Slight differences between the two profiles were observed (orange lane band in Fig.66), which could be due to different Cdc45-3X-FLAGp interactions in the two strains.

The silver staining analysis allowed detection of a few differences between the proteomic patterns of the WT and *Δrpd3* strains however at the same time it did not allow the specific identification of Cdc45-3X-FLAGp interactors.

Therefore to obtain more detailed data on the proteins associated with stalled forks the protein samples immunoprecipitated from the three strains were studied by mass spectrometry.

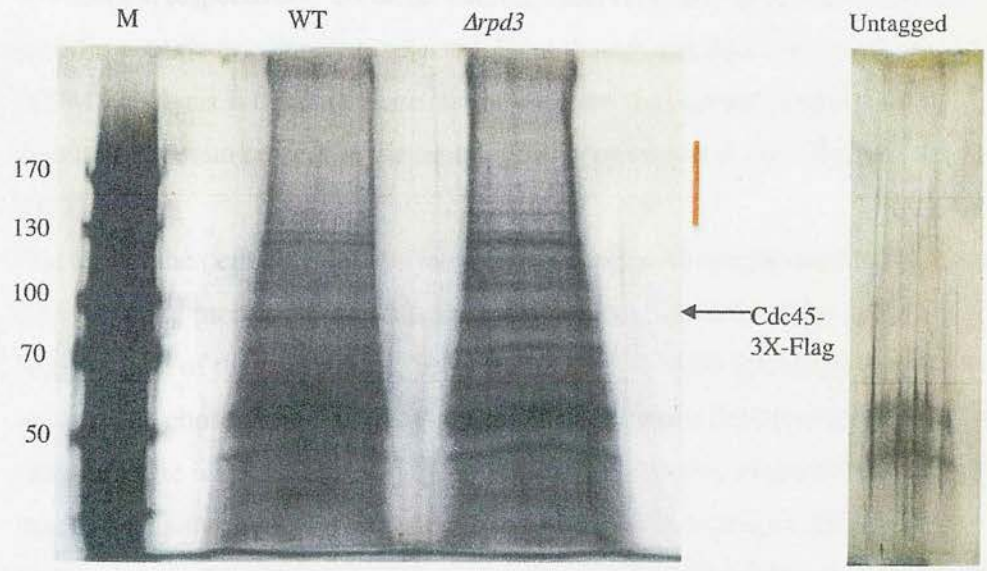


Fig.66 Silver Staining of WT, *Δrpd3* and untagged strains. The arrow indicates the band corresponding to Cdc45-3X-FLAGp, which is only present in the protein samples from the WT and *Δrpd3* strains. The orange bar indicates a zone in the gel where the patterns from the two strains slightly differ.

5.7 MASS SPECTROMETRY ANALYSIS ON IMMUNOPRECIPITATED SAMPLES

The IP fractions from the WT, *Δrpd3* and untagged strains were loaded on an SDS-page gel and the electrophoresis was stopped after the samples were inside the separating gel. After staining with Coomassie Blue solution, bands from the three strains were cut from the gel and analyzed by Mass Spectrometry (in collaboration with Juri Rappsilber, ICB, Edinburgh).

The proteins pulled down in the untagged strain represent the background control and thus they were removed from the proteomic profiles of the WT and *Δrpd3* strains. The protein lists of both WT and *Δrpd3* contained approximately 200 and 300 entries, respectively. To futher shorten these lists only the proteins with a specific peptide number were selected. Even though quantification of protein by MS/MS analysis is difficult a correlation between the peptide number and the amount of protein present in the samples has been demonstrated (Rappsilber, *et al*, 2002).

The higher the peptide value the more reliable the result can be considered. Thus of the remaining proteins those with a peptide number above two were selected. A complete list of the proteins can be found in Table 2 in the Supplementary data section.The chosen proteins were then analyzed for their functions using the DAVID bioinformatic website. The pie graphs in Fig.67 report the results obtained from the functional analyses on WT and *Δrpd3* cells arrested in S-phase with HU.

Each colour in the pie graphs represents a specific functional cluster, as indicated in the figure legend. The same colors are used for both pie graphs. The percentage value is indicative of the number of proteins belonging to each specific functional group.

As it can be seen in Fig.67, the most abundant groups in both samples are the proteins involved in metabolic and catabolic processes. Protein with unknown funtion represent a high percentage of the recovered protein, specifically 17.3% and 12.3% in WT and *Δrpd3* respectively. DNA replication factors have been identified in both samples with similar abundance: 13.3 % in WT and 13.3% in *Δrpd3* (Fig.67).

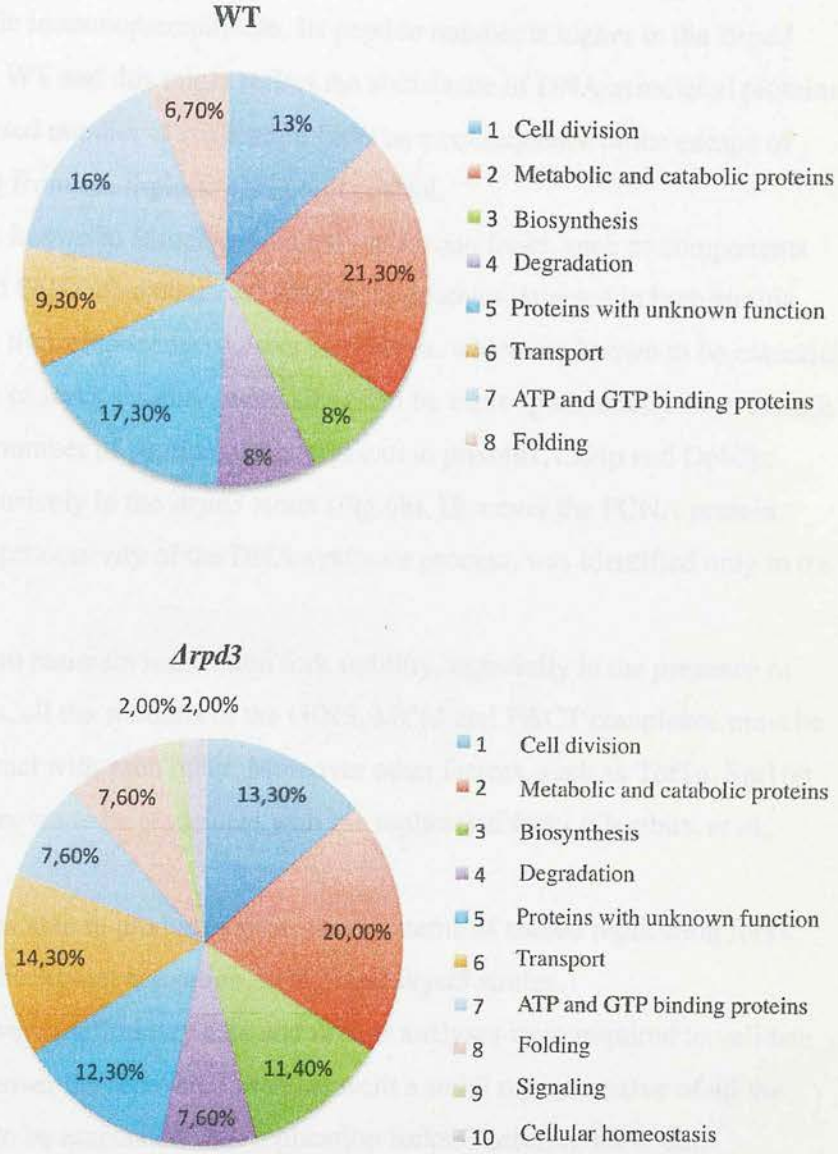


Fig.67 Pie graphs reporting the functional analysis on the proteins identified by mass spectrometry analysis of WT and *Δrpd3* cells treated with HU. The colours correspond to different functional groups as indicated in the legend; the asterisk in the legend indicates the cluster present only in *Δrpd3* cells. The protein abundance in each group is indicated as a percentage. Of the proteins identified in WT and *Δrpd3* cells 13% and 13.3% are involved in cell cycle regulation respectively.

To study the proteomics of replication forks the proteins strictly involved in DNA replication identified in each strain were compared.

The graph presented in Fig.68 indicates the number of peptides hits for each of these proteins. Cdc45p is the most abundant protein in both samples. This might also be due to its specific immunoprecipitation. Its peptide number is higher in the *Δrpd3* sample than the WT and this might reflect the abundance of DNA associated proteins due to an increased number of replication forks as a consequence of the escape of late origin firing from the S-phase chckpoint control.

Several proteins known to associate with the replication forks, such as components of the MCM and GINS complexes are among the proteins detected in both strains. However not all the components of these complexes, which are known to be essential for maintenance of forks stability, were identified by mass spectrometry even though with a not high number of peptide. Other replication proteins, Ctf4p and Dpb3p, were found exclusively in the *Δrpd3* strain (Fig.68). However the PCNA protein, required for the processivity of the DNA synthesis process, was identified only in the WT strain.

It is known that to maintain replication fork stability, especially in the presence of replication stress, all the subunits of the GINS, MCM and FACT complexes must be present and interact with each other. Moreover other factors, such as Tof1p, Spt16p and Pob3p are known to be associated with the replication forks (Gambus, *et al*, 2006),

In summary I was able to produce a proteomic patterns of stalled replication forks due to S-phase checkpoint activation, in WT and *Δrpd3* strains.

However these were preliminary data and further analyses were required to validate the results. Moreover the recovered proteins were a small representative of all the proteins known to be associated with replication forks. Therefore these data demonstrated the designed technique allowed the pulldown of Cdc45-3X-FLAG associated proteins. At the same time the low abundance of replication factors indicate that some steps of the technique require an improvement in order to detect more proteins and in particular the one known to travel with the replication fork (for example subunits of the GINS and FACT complexes).

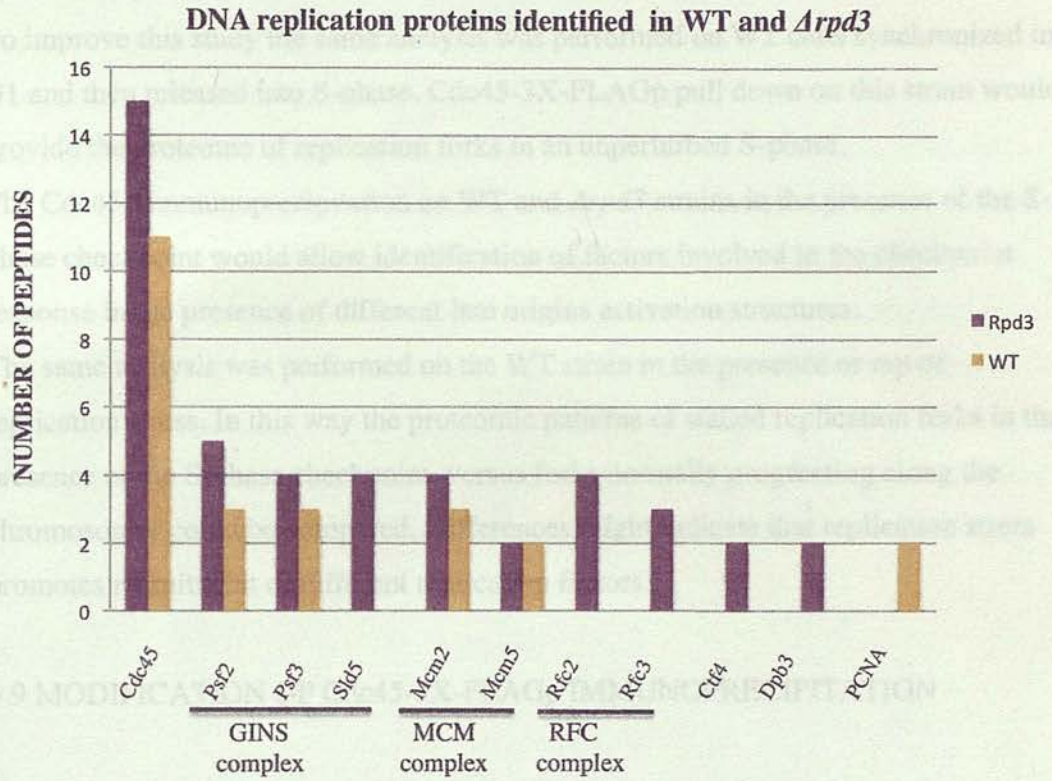


Fig.68 Replication proteins immunoprecipitated with Cdc45p to determine the proteomes of stalled forks in WT (purple) and *Δrpd3* (orange) strains treated with HU. The number of peptides (Y axis) obtained for each protein, by mass spectrometry analysis, is reported; the names of the proteins are shown on the X axis. Although many proteins known to be associated with replication forks were detected in both strains others escaped identification or were only detected in one of the strains.

5.8 PROTEOMICS OF REPLICATION FORKS IN WT AND *Δrpd3* IN THE PRESENCE OR IN THE ABSENCE OF THE S-PHASE CHECKPOINT

The aim of the project was to study the proteomics of replication forks in order to improve our knowledge of the proteins that regulate the DNA replication process. The analysis previously described was performed on WT and *Δrpd3* strains, treated with HU to promote S-phase checkpoint activation.

To improve this study the same analysis was performed on WT cells synchronized in G1 and then released into S-phase. Cdc45-3X-FLAGp pull down on this strain would provide the proteome of replication forks in an unperturbed S-phase.

The Cdc45p immunoprecipitation on WT and *Δrpd3* strains in the presence of the S-phase checkpoint would allow identification of factors involved in the checkpoint response in the presence of different late origins activation structures.

The same analysis was performed on the WT strain in the presence or not of replication stress. In this way the proteomic patterns of stalled replication forks in the presence of the S-phase checkpoint, versus forks normally progressing along the chromosomes could be compared. Differences might indicate that replication stress promotes recruitment of different replication factors.

5.9 MODIFICATION OF Cdc45-3X-FLAGp IMMUNOPRECIPITATION

The data presented in the previous section demonstrated that chromatin fractionation combined with Cdc45p immunoprecipitation cells allowed the study of the proteomes of stalled replication forks in HU arrested cells. Most of the proteins known to be associated with fork paused by the S-phase checkpoint activation could be detected. However some proteins were present with a low peptide number while other replication factors escaped pull down. Therefore to obtain a wider view of the proteomics of replication forks some modifications, such as the use of Protein G Dynabeads and the trypsin digestion of the immunoprecipitated samples, were applied to the protocol.

5.10 ARREST OF CELLS IN S-PHASE

WT, *Δrpd3* and untagged strains were arrested in S-phase by treatment with HU that induces activation of the S-phase checkpoint and stalling of replication forks. The same analysis was carried out on the WT strain arrested in G1 with α-factor and released into S-phase for 30 minutes. An aliquot was collected from each culture to confirm by FACS analysis that the cells were arrested in S-phase.

The FACS profiles are reported in Fig.69 (panels A and B); 1C and 2C indicate unreplicated and replicated (G2 phase) DNA content respectively.

The profile of the WT strain arrested in G1 and then released into S-phase for 30 minutes is presented in panel A. The sample has a DNA content close to 1C but the width of the peak indicates that cells were starting to progress into S-phase.

Profiles of the WT, Untagged and *Δrpd3* strains, treated with HU, are reported in panel B (Fig.69). The DNA content is mostly 1C, as expected for cells arrested in S-phase by HU treatment.

All together these data demonstrated that the proteomic data will be representative of replication forks progressing through S-phase (WT, Fig.69 A) and stalled by HU treatment (WT, Untagged and *Δrpd3* Fig.69 B).

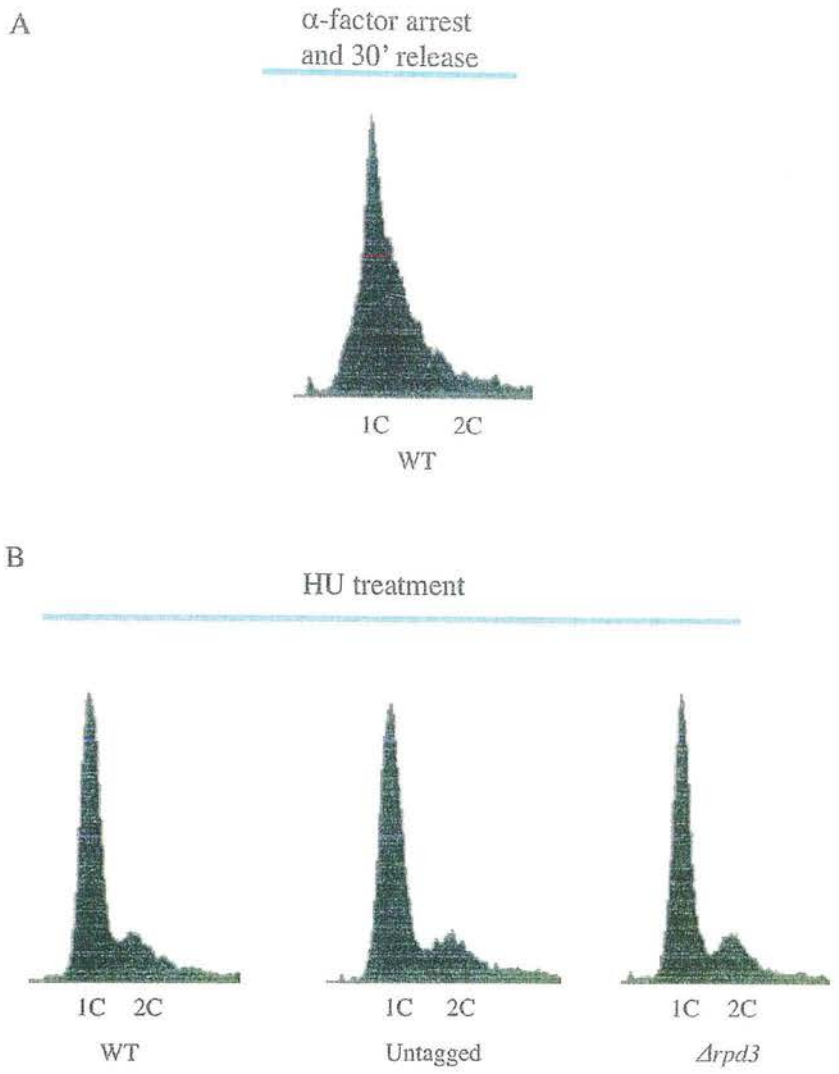


Fig.69 FACS profiles of the samples processed for Ccd45p immunoprecipitation. 1C and 2C indicates DNA content. A) Profile of WT strain (MMY033) arrested in G1 with α-factor and then released into S-phase for 30 minutes. B) Profiles of WT (MMY033), Untagged and (MMY021) *Δrpd3* (MVY51) strains arrested in S-phase by two hours incubation with HU. As expected all four profiles indicate that most of the cells have a DNA content close to 1C.

5.11 PROTEOMICS OF REPLICATION FORKS IN DIFFERENT CONDITIONS

Cells from the four cultures described above were collected and processed at the same time for chromatin fractionation. The recovered chromatin bound fractions were digested with DNaseI and the released proteins incubated with αFLAG antibody and ProteinG Dynabeads. The eluted samples were then digested with Trypsin and the recovered peptides were analyzed by mass spectrometry (in collaboration with Juri Rappsilber, ICB, University of Edinburgh).

The proteins present in the untagged strain, which likely represent contaminant factors, were removed from the results for the WT strain arrested in α-factor, and the WT and *Δrpd3* strains treated with HU. As for the previous analysis a cut off was decided in order to analyze the proteins that were likely to be more abundant in the prepared samples. Thus the proteins with a peptide number higher than two were selected for the following analysis. A complete list of the proteins can be found in the Supplementary data section (Table 3-6).

A further analysis was performed in order to compare the proteins involved in DNA replication identified in the different growth conditions. In the graph in Fig.70 replication factors recovered in the three strains are reported according to their peptide number. In general most of the proteins known to be involved in DNA replication elongation, such as the MCM and GINS subunits were detected in the three strains. Indeed all the helicase proteins were immunoprecipitated with Cdc45p and their peptide number is higher than in the previous experiment (Fig.68) and comparable to Cdc45p. Also Ctf4p and Tof1p and Csm3p were present in all the samples. In addition all the FACT complex subunits, Spt16p and Pob3p, which were not recovered in the previous analysis (Fig.68) were identified in the three strains analyzed. These results are more consistent with data in the literature that have demonstrated the association of these proteins with replication forks in an unperturbed S-phase and in the presence of the S-phase checkpoint (Calzada, et al, 2005; Katou, *et al*, 2003).

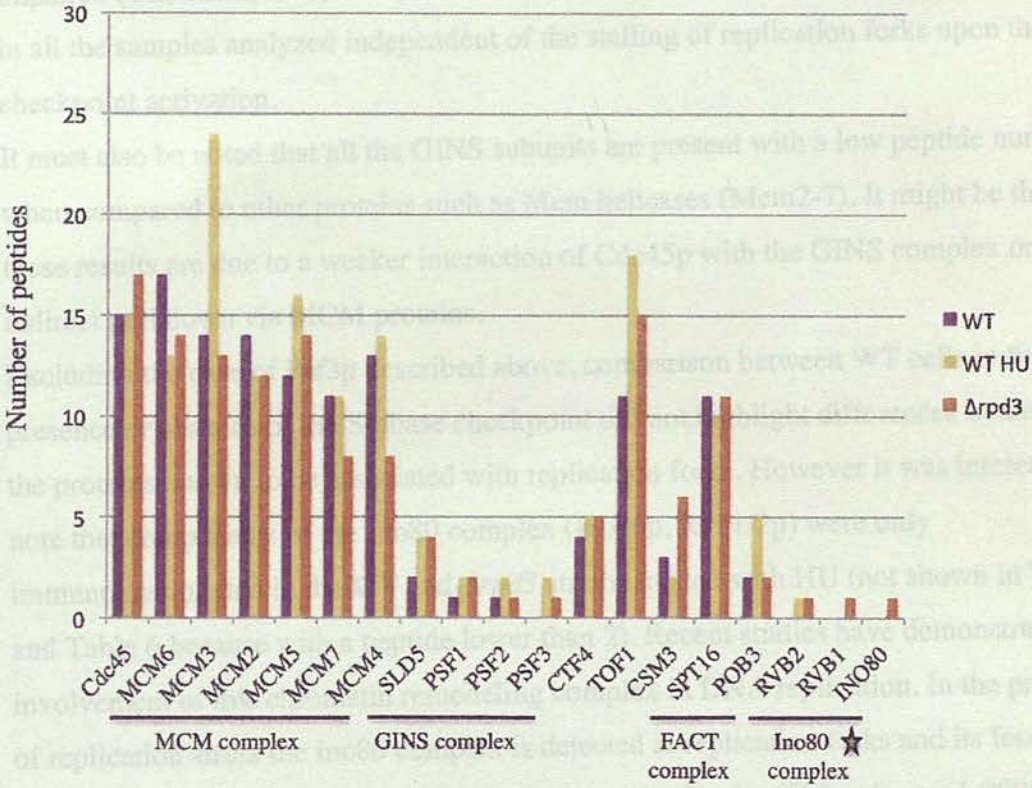


Fig.70 Proteins recovered from the Cdc45p immunoprecipitation on WT and *Δrpd3* strains treated with HU and on a WT strain in an unperturbed S-phase. The number of peptides obtained for each protein by mass spectrometry analysis, is reported. This new experiment gave a more complete view of replication forks. Comparison between the three strains does not show particular differences. However it is interesting to note that components of the Ino80 complex, recently demonstrated to be involved in the checkpoint response, are identified in the protein sample from the *Δrpd3* strain (asterisk).

Thus after comparison with the proteomic patterns presented in Fig.68, these results demonstrate that the changes applied to the protocol have allowed a more efficient pull down of Cdc45p and its associated proteins.

However the technique requires further improvements. Indeed the Psf3p subunit of the GINS complex was only detected in the samples treated with HU and with a low peptide number. It is known that a functional GINS complex is an essential replisome component required to ensure replication fork stability. It has been demonstrated in fission yeast that DNA replication is arrested in *psf3-1* mutants and GINS assembly is impaired (Yabuuchi, *et al*, 2006). Therefore it is likely that this protein should be present in all the samples analyzed independent of the stalling of replication forks upon the checkpoint activation.

It must also be noted that all the GINS subunits are present with a low peptide number when compared to other proteins such as Mcm helicases (Mcm2-7). It might be that these results are due to a weaker interaction of Cdc45p with the GINS complex or to the indirect pull down via MCM proteins.

Excluding the case of Psf3p described above, comparison between WT cells in the presence or absence of the S-phase checkpoint did not highlight differences in terms of the proteins known to be associated with replication forks. However it was interesting to note that components of the Ino80 complex (Ino80p, Rvb1/2p) were only immunoprecipitated in the WT and *Δrpd3* strains treated with HU (not shown in Table 5 and Table 6 because with a peptide lower than 2). Recent studies have demonstrated an involvement of this chromatin remodeling complex in DNA replication. In the presence of replication stress the Ino80 complex is detected at replication forks and its Ies4 subunit becomes phosphorylated by the S-phase checkpoint (Shimada, *et al*, 2008; Morrison, *et al*, 2007). These results validate the use of this technique for the analysis of the proteins associated with replication forks.

Thus, it can be concluded that I have developed an unbiased assay for the *in vivo* study of the proteomics of replication fork. The technique allowed pull down of most of the replication proteins know to be part of the replisome complex. Moreover the results obtained are also in agreement with emerging data regarding the proteins associated with stalled replication forks. Comparison of the results between the three

strains did not reveal many differences. However a more in depth study must be performed in order to validate the data and improve our knowledge of the proteins associated with replication forks. Indeed this analysis could be performed on the same strains but immunoprecipitating other proteins, such as MCM or GINS subunits, which in association with Cdc45p form the core of the replisome progression complex (reviewed in Aparicio, *et al*, 2006). These new experiments could confirm the data presented here and allow identification of different proteins that might have escaped detection because of a low association with Cdc45p.

6 DISCUSSION

Since the discovery of regions that could act as autonomously replicating sequences our knowledge of the mechanisms that regulate genome duplication has been significantly improved. The proteins involved in DNA replication have been identified and extensively studied. Indeed, many of the factors that have a role in the formation of pre-RC and RC and that regulate origin firing have been unveiled. Moreover the involvement of novel proteins, such as Mcm8p and Mcm9p, and chromatin remodeling complexes, such as Ino80, in the regulation of DNA replication has been recently observed (Morrison, *et al*, 2008). Therefore, it is likely that other factors are still to be identified.

The projects described in the thesis were focused on the identification of new proteins involved in DNA replication in *S. cerevisiae*. Identification of conserved sequences has led to the determination of replication origin locations and time of origin firing. Therefore *S. cerevisiae* is a good model organism for studies based on the identification of factors that are involved in DNA replication. Considering the presence of redundant mechanisms that regulate DNA replication and the fact that most proteins are essential for cell survival, genetic means cannot be applied for this analysis. Two different strategies were designed in order to study, by mass spectrometry, the proteins associated with replication origins and replication forks. These analyses could lead to the identification of unknown proteins.

To pursue this aim I developed a system for the *in vivo* study of the proteomes of eukaryotic replication origins, taking into consideration the influence of the chromosomal context in determining origin firing. The assay consisted of three main steps: excision of a replication origin from the chromosome, isolation of the excised chromatin plasmid and mass spectrometry analysis of the proteins associated with the replication origin contained in the purified episome. Excision of the late origin ARS1413 was performed using the pSR1 site-specific recombination system. The constructed pFA6a-TRP1/RsiteB and pFA6a-KanMX/RsiteA plasmids were used to integrate two pairs of recombination sites flanking ARS1413. However it must be noted that these constructs represent an important genetic tool for the analysis of different genomic regions (Mariotti, *et al*, manuscript in preparation). Indeed, by designing the appropriate primers, the R sites can be integrated at any location in the

S. cerevisiae genome. Therefore plasmids pFA6a-TRP1/RsiteB and pFA6a-KanMX/RsiteA can be used at a genome-wide level to study the involvement of any region of interest in various cellular processes.

Integration of the gene encoding for the Recombinase enzyme under the Gal promoter into the yeast genome allowed the control of the recombination event leading to ARS1413 excision only in induced cells (Fig.51 and 52). Preliminary experiments performed to adjust the conditions for more efficient origin excision demonstrated that the pSR1 recombination system is completely independent from the one present in *S. cerevisiae* (Fig.49) and that the copy number of the R gene positively influence the efficiency of recombination (Fig. 51 and 52). For efficient use of plasmid DNA it is essential that the vector is stably maintained into the cells over generations; this can be achieved by introducing sequences, such as centromeric sequences, that ensure proper segregationof the plasmid to the daughter cells. To maintain intact chromosomal chromatin no centromeric sequence were introduced surrounding the chosen late replication origin. Southern blot analysis demonstrated that the chromatin episome was maintained within the cells during five hours of induction (Fig.53). Therefore these data indicate that pSR1 represented a good choice for the *in vivo* excision of the late replication origin.

However further experiments, such as bi-dimensional gel analysis, should be performed to study if the ARS1413 episome replicates during treatment with Galactose. This will be essential for several reasons. Knowing the replication state of the plasmid will be important to investigate the relationship between chromatin structure and DNA replication. Moreover it must be taken in consideration that the step following excision in the described strategy consists of the isolation of the episome from the rest of the genome and its analysis by mass spectrometry. Therefore replication of the plasmid is an important factor that has to be considered when setting up the condition in order to recover enough material for the proteomic analysis.

The isolation of the ARS1413 plasmid was performed taking advantage of the protocol developed by Kim and colleagues (Kim, *et al*, 2004) to isolate a high copy plasmid from the rest of the genome, in *S. cerevisiae*. Considering both that the

plasmid might be present in one or two copies per cell and the small size of the chromatin episome (1.3 kbp) adjustments, such as increasing the volume of the starting material and the length of the sucrose cushion centrifugation, were applied to the protocol in order to improve the efficiency of the isolation. The purification of the ARS1413 episome was confirmed by Southern blot analysis using probes annealing within the late replication origin (Fig.61) and the *TRP1* gene (Fig.62). Only the ARS1413 probe detected the band corresponding to the excised chromatin plasmid. However both primers gave bands of high molecular size (over 10000 kbp) that do not correspond to the episome; these signal were present in both the induced and uninduced cells. These data indicate that other chromatin regions, apart from the ARS1413 episome, could be present in the samples implying that the adopted technique did not result in specific isolation of the excised origin. Indeed in the detailed explanation of the isolation by Kim and colleagues it was specified that the sucrose cushion recovered after ultracentrifugation is usually dirty probably due to the presence of contaminants such as chromosomal chromatin (Kim, *et al*, 2004). For the aim of the project it is essential that chromatin plasmid is exclusively isolated from the rest of the genome.

Technical issues were also encountered during the purification of the sucrose cushion sample containing the chromatin plasmid impairing the mass spectrometry analysis. The results reported in Table 1 indicate that the proteins ientified by mass spectrometry analysis are not involved in DNA replication. However it should be taken in consideration that the plasmid was isolated from asynchronous cells and thus the concentration of replication proteins could have been lower than other factors. Nevertheless the Orc2 protein, which was previously shown by ChIP analysis to be bound to the excised ARS1413 (Fig.56 panels A and B) was not identified by mass spectrometry analysis. This discrepancy might be due to the difficulties, described above, encountered during the isolation step.

It is essential to overcome these technical problems in order to reduce the volume of the sample and to purify the plasmid from the sucrose, which interferes with the subsequent proteome analysis. Changes could be applied to specific steps of the protocol such as to the nuclei preparation in order to increase the quality of the

sample. Different lengths of centrifugation of the step gradient could be tested to study how the time of centrifugation can influence the nuclei preparation.

An alternative could be to utilise the Orc2-3X-FLAG protein. Thus, an immuno-affinity purification step could be performed on the isolated ARS1413 episome to specifically enrich the samples for Orc2-3X-FLAGp and its associated proteins. However it must be considered that the adopted isolation technique was originally designed to specifically study the nucleosome position of the recovered *CUP1* gene and not for a proteomic analysis. Histone proteins are more tightly bound to DNA than replication proteins and therefore the difficulties encountered in the mass spectrometry analysis could also be due to the protocol per se.

Other plasmid purification methods were developed that are based on the Minichromosome Affinity Purification (MAP) technique. Plasmids containing a single Lac operator (LacO) or 21 Tet operator (Tet O) sequences were purified using immobilized Lac repressor or Tet-R proteins respectively (Ducker and Simpson, 2000; Ivanov and Nasmyth, 2005). However, these systems did not allow the recovery of enough or clean material for proteomic analysis. Further improvement to the MAP technique came from a recent system proposed by Unnikrishnan and colleagues (Unnikrishnan, *et al*, 2010). They constructed a 1.8 kbp plasmid containing ARS1, the TRP1 marker gene, seven well-mapped nucleosomes and eight Lac operator sequences. The plasmid was extracted from whole cell extracts using Protein G Dynabeads and 3X-FLAG LacI. Adjustments applied to the existing methods resulted in recovery of sufficient material (1-2 μg/10 L of early-log phase culture) and improved the purity of samples. Indeed mass spectrometry analysis on the isolated plasmid allowed the relationship between histone modifications and DNA replication to be studied. This system could be adopted for the purification of the ARS1413 chromatin plasmid. Insertion of the eight LacO sites into the *S. cerevisiae* genome between the integrated inverted repeats could improve plasmid purification and the following proteomic analysis. Even though in this way the chromatin context might be altered, the insertion of DNA binding sites represent at the moment the best solution to overcome the described technical issues.

Therefore it can be concluded that although this approach could potentially provide a new strategy for the *in vivo* analysis of proteins required for DNA replication it stil requires further adjustments. Indeed until the chromatin plasmid can be exclusively purified from the rest of the genome this technique cannot be applied to study the proteomic of replication origins.

Plasmid DNA has been widely used to study different biological processes, such as chromatin organization (Walter and Newport, 2000) and DNA replication (Takeda, *et al*, 2005). Indeed Takeda and colleagues took advantage of chromatin vector to investigate the recruitment of ORC and Cdc6 to replication origins (Takeda, *et al*, 2005). They created an artificial origin *in vivo* using pFR_Luc, which contains five GAL4-binding sites. Fusing pFR-Luc DNA-binding domains to human initiation factors they demonstrated the requirement of ORC and Cdc6 proteins for pre-RC assembly. Moreover these data indicated that the artificial recruitment of any eukaryotic replication initiation proteins to a DNA sequence can create a functional origin of replication (Takeda, *et al*, 2005). It has been demonstrated that "nucleus-free" extracts of Xenopus eggs represent a model system for initiation of DNA replication because they support efficient plasmid replication (Walter, *et al*, 1998). Walter and Newport took advantage of this system to study the link between DNA replication and changes in the DNA structure (Walter and Newport, 2000). Analysing changes in the topology of the plasmid DNA they demonstrated that with the proceeding of initiation of DNA replication an increase in negative supercoiling could be detected. Changes in the topology of plasmid DNA required the replicative helicase, MCM, as well as Cdc45 and RPA (Walter and Newport, 2000). All these data indicate that the use of vectors represent an important system for studies focused on the mechanisms that regulate DNA replication.

Moreover as mentioned before Raghuraman and colleagues already used the same pSR1 site-recombination system to excise the late replication origin ARS305 from its chromosomal context (Raghuraman, *et al*, 1997). However the technique was applied to study telomere position effect on origin firing. Except for the work published by Unnikrishnan and colleagues (Unnikrishnan, *et al*, 2010) none of the other studies were focused on the proteomic analysis of the isolated plasmids. Therefore the

unbiased assay described in this thesis is the first to combine the *in vivo* isolation of a eukaryotic replication origin from its chromosomal context and a proteomic analysis.

The peculiarity of this project compared to the one proposed by Unnikrisnan and colleagues (Unnikrishnan, *et al*, 2010) is in the maintenance of the origin of replication in its chromosomal context. Indeed all the MAP techniques were applied to constructed plasmids containing the sequence of a replication origin. Instead I developed an assay with the aim of studying the proteins associated with origins of replication that are kept in their chromosomal context, which is known to influence replication origin activity. Therefore, it is essential to carefully evaluate and further optimise the conditions for cleaner purification of the excised 1.3 kbp ARS1413 chromatin plasmid. Indeed once this is achieved the applications of this type of strategy are vast. The ability to tightly control R gene expression allows the induction of recombination, and thus of ARS1413 excision, in different phases of the cell cycle in order to study various proteomic complexes. Therefore this approach could be useful to identify the proteins that are involved in the formation of the pre-RC when applied to cells arrested in G1 with α-factor. On the other hand treatment of the cells with HU, which depletes the pool of dNTPs, will induce the activation of the S-phase checkpoint. Excision of ARS1413 in cells arrested in S-phase with HU will lead to the identification of the factors that interact with late replicating origins and inhibit their firing.

Moreover, as mentioned above, plasmids pFA6a-KanMX6/RsiteA and pFA6a-TRP1/RsiteB are useful tools for the insertion of the Recombination sites into any region of the *S. cerevisiae* genome (Mariotti, *et al*, manuscript in preparation). Their integration into the regions flanking an early replication origin will enable the comparison between the proteomics of early and late replication origins. Thus proteins that specifically interact with the replication origins according to their time of firing could be identified. The data obtained also by Cdc45-3X-FLAG immunoprecipitation led to the identification of the proteins that are involved in DNA replication both during DNA synthesis and checkpoint activation. Extending

the analysis to the MCM and GINS complexes could improve our knowledge of the mechanisms that regulate replication fork progression and stalling.

The possibility of applying the proposed strategies to cells grown in various conditions will allow the study of proteins required in the different phases of DNA replication and in response to checkpoint activation. However these analyses might also unveil novel functions for factors known to be involved in other processes. At the same time it must also be considered that replication protein might be required for other cellular processes. The ORC complex is a significant example of this correlation. Indeed it localizes, in budding yeast, at the mating type loci where it promotes heterochromatin formation. In DNA replication ORC is an essential component of the pre-RC and it is required for origin licensing. Moreover, it has been demonstrated that Orc1p controls centrosome and centriole copy number in human cells (Hemerly, *et al*, 2009). Other regulators of DNA replication, such as Mcm5p and Cyclin E, have been reported to also be involved in centrosome duplication (Ferguson and Maller, 2008). Recently a novel role for JMJD2A, the human H3K9/K36 demethylase, has been discovered (Black, *et al*, 2010). Indeed JMJD2A regulates DNA replication by controlling chromatin accessibility. Therefore, the wide range of applications of these techniques could improve our knowledge of the mechanisms not only that regulate DNA replication but also that link together the different cellular processes, such as DNA replication and transcription.

7 SUPPLEMENTARY DATA

TABLE 1

Table1. Mass spectrometry analysis of the proteins identified after ARS1413 chromatin plasmid isolation. For each protein the score, the peptide number and the function are reported.

Protein name	score	Peptide n	
YMR051C	79	14	Retrotransposon TYA Gag gene
ALD5	125	12	Mitochondrial aldehyde dehydrogenase
HXT6	92	12	High-affinity glucose transporter
RHO1	95	11	GTP-binding protein of the rho subfamily
TSA1	81	8	Ubiquitous housekeeping thioredoxin peroxidase
SSA1	55	8	ATPase involved in protein folding and nuclear localization
FUM1	37	8	Fumarase
ACS2	148	7	Acetyl-coA synthetase isoform
TCB1	84	7	Lipid-binding protein
PUT2	55	7	Delta-1-pyrroline-5-carboxylate dehydrogenase
WTM1	71	6	Transcriptional repressor; regulation of meiosis and silencing
SHM1	62	6	Mitochondrial serine hydroxymethyltransferase
PEP4	61	6	Vacuolar aspartyl protease (proteinase A)
CYB2	57	6	Cytochrome b2 (L-lactate cytochrome-c oxidoreductase)
ERG24	48	6	C-14 sterol reductase, acts in ergosterol biosynthesis
RNR4	39	6	Ribonucleotide-diphosphate reductase (RNR), small subunit
HSC82	128	5	Cytoplasmic chaperone of the Hsp90 family
PDX1]	99	5	Dihydrolipoamide dehydrogenase (E3)-binding protein
SSH1	98	5	Subunit of the Ssh1 translocon complex
HSP78	62	5	Oligomeric mitochondrial matrix chaperone
POR1	53	5	Mitochondrial porin (voltage-dependent anion channel)
PRB1	49	5	Vacuolar proteinase B
ILV2	45	5	Acetolactate synthase
TAO3	40	5	Protein involved in cell morphogenesis and proliferation
VAS1	32	5	Mitochondrial and cytoplasmic valyl-tRNA synthetase
RPO26	135	4	RNA polymerase subunit ABC23
CYS3	106	4	Cystathionine gamma-lyase
SSS1	104	4	Subunit of the Sec61p translocation complex
KGD2	89	4	Dihydrolipoyl transsuccinylase, complex
ERG10	80	4	Acetyl-CoA C-acetyltransferase (acetoacetyl-CoA thiolase)
ALT1	68	4	Putative alanine transaminase
RPS4B	61	4	Protein component of the small (40S) ribosomal subunit
HSP104	60	4	Heat shock protein; cooperates with Ydj1p and Ssa1p
MAM3	60	4	Protein required for normal mitochondrial morphology
SIS1	59	4	Type II HSP40 co-chaperone
FMP12	55	4	Protein of unknown function
PCS60 n	53	4	Peroxisomal AMP-binding protei

LAP4	48	4	Vacuolar aminopeptidase
VMA4	37	4	Subunit E of the V1 peripheral membrane domain
TIF2	36	4	Translation initiation factor eIF4A, identical to Tif1p
GLK1	33	4	Glucokinase
RPL28	30	4	Ribosomal protein L29 of the large (60S) ribosomal subunit
KES1	53	3	Member of the oxysterol binding protein family
TPD3	50	3	Regulatory subunit A of the protein phosphatase 2A
HFD1	49	3	Hypothetical protein
LHS1	44	3	Molecular chaperone of the endoplasmic reticulum lumen
YLL029W	43	3	Putative protein of unknown function
PHB2	42	3	Subunit of the prohibitin complex (Phb1p-Phb2p)
LEU4	41	3	Alpha-isopropylmalate synthase
RPP2A	40	3	Ribosomal protein P2 alpha
BRR2	39	3	RNA-dependent ATPase RNA helicase
ALR1	39	3	Plasma membrane Mg(2+) transporter
DLD2	36	3	D-lactate dehydrogenase, located in the mitochondrial matrix
TCB3	34	3	Lipid-binding protein
VPS34	33	3	Phosphatidylinositol 3-kinase
EXG2	29	3	Exo-1,3-beta-glucanase
ILV3	27	3	Dihydroxyacid dehydratase
VID28	26	3	Protein involved in catabolite degradation of fructose-
BAT1	25	3	Mitochondrial branched-chain amino acid aminotransferase
HOM2	24	3	Aspartic beta semi-aldehyde dehydrogenase
PDH1	24	3	Mitochondrial protein that participates in respiration

Table1. Proteins recovered from the Raffinose samples were considered as a background and therefore they were removed from the profile of the Galctose samples. Of the remaining hits, the proteins with a peptide number lower than 2 were not considered.

TABLE 2

Table 2. Mass spectrometry analysis of the proteins identified after Cdc45-3X-FLAGp immunopreicpitation in WT cells treated with HU. For each protein the score, the peptide number and the function are reported.

Protein name	Peptides	Score	Protein Description
CDC45	11	571	Cell division control protein 45
ADH1	9	443	Alcohol dehydrogenase 1 (YADH-1)
TDH3	9	516	Glyceraldehyde-3-phosphate dehydrogenase 3
GRS1	7	311	Glycyl-tRNA synthetase 1
PTH2	6	216	Peptidyl-tRNA hydrolase 2
YMM5	5	208	Hypothetical 35.3 kDa protein in HMGS-TUB3
YHB1	5	234	Flavohemoprotein (Hemoglobin-like protein)
GRP78	5	246	78 kDa glucose-regulated protein homolog precursor
ERG27	5	249	3-keto-steroid reductase
ARF1	5	289	ADP-ribosylation factor 1
YM32	4	141	Hypothetical 29.1 kDa protein in IMP1-HLJ1 intergenic region
CTP1	4	148	Tricarboxylate transport protein (Citrate transport protein) (CTP)
ARO10	4	167	Transaminated amino acid decarboxylase
ENO2	4	170	Enolase 2 (2-phosphoglycerate dehydratase 2)
YE12	4	180	Hypothetical 27.7 kDa protein in ISC10 3'region
TRS31	4	190	Transport protein particle 31 kDa subunit
YPT1	4	200	GTP-binding protein YPT1 (Rab GTPase YPT1)
YBQ2	4	215	Hypothetical 22.9 kDa protein in REG2-YRO2 intergenic region
QCR2	4	226	Ubiquinol-cytochrome-c reductase complex core protein 2
PLSC	3	96	Probable 1-acyl-sn-glycerol-3-phosphate acyltransferase
ERG9	3	98	Squalene synthetase
GET3	3	103	ATPase GET3 (Arsenical pump-driving ATPase)
PSF2	3	105	DNA replication complex GINS protein PSF2
RSR1	3	113	Ras-related protein RSR1
MCM2	3	116	DNA replication licensing factor MCM2
ERG13	3	119	Hydroxymethylglutaryl-CoA synthase (HMG-CoA synthase)
CCT4	3	123	T-complex protein 1 subunit gamma (TCP-1-gamma)
GFA1	3	131	Glucosamine--fructose-6-phosphate aminotransferase
TBA1	3	136	Tubulin alpha-1 chain
YJG6	3	137	Hypothetical 28.5 kDa protein in SMC3-MRPL8 intergenic region
ASC1	3	139	Guanine nucleotide-binding protein beta subunit-like protein
CCC1	3	159	Protein CCC1
MRT4	3	161	mRNA turnover protein 4
HHF2	3	164	Histone H4
SPC3	3	164	Microsomal signal peptidase subunit 3
YG4Y	3	165	Hypothetical 26.9 kDa protein in YHB1-PFK1 intergenic region
HTB1	3	182	Histone H2B.1
ARB1	3	190	Probable ATP-dependent transporter YER036C
PSF3	3	200	DNA replication complex GINS protein PSF1
SLC1	3	96	Probable 1-acyl-sn-glycerol-3-phosphate acyltransferase
YJ58	2	59	Hypothetical 33.9 kDa protein in STE18-GRR1 intergenic region
MBF1	2	59	Multiprotein-bridging factor 1
SEC61A	2	59	Protein transport protein SEC61 alpha subunit
RNA1	2	59	Ran GTPase-activating protein 1
YBL059W	2	61	Protein YBL095W
GPA1	2	66	T-complex protein 1 subunit delta (TCP-1-delta) (CCT-delta)
SEC66	2	66	Translocation protein SEC66 (Protein HSS1)
ADE5,7	2	67	Bifunctional purine biosynthetic protein ADE5,7

YKU6	2	67	Hypothetical 30.7 kDa protein
PCNA	2	67	Proliferating cell nuclear antigen (PCNA)
YRB4	2	70	Importin beta-4 subunit (Karyopherin beta-4 subunit)
ALG5	2	71	Dolichyl-phosphate beta-glucosyltransferase
ARL3	2	72	ADP-ribosylation factor-like protein 3 (Arf-like GTPase 3)
YPT10	2	72	GTP-binding protein YPT10
YNP6	2	73	Hypothetical 33.7 kDa protein in YGP1-YCK2 intergenic region
ORM1	2	74	Protein ORM1
AAC3	2	75	ADP, ATP carrier protein 3 (ADP/ATP translocase 3)
MCM5	2	78	Minichromosome maintenance protein 5
RPN2	2	84	26S proteasome regulatory subunit RPN2
CDC16	2	84	Anaphase-promoting complex subunit CDC16
MST27	2	84	Multicopy suppressor of SEC21 protein 27
YHI0	2	84	Putative prolyl-tRNA synthetase
YM23	2	85	Hypothetical 62.8 kDa protein
GPA1	2	91	Guanine nucleotide-binding protein alpha-1 subunit (GP1-alpha)
SEC66	2	91	Hypothetical 28.6 kDa protein in DBP6-COQ2 intergenic region
PRE9	2	95	Proteasome component Y13
YPT6	2	97	GTP-binding protein YPT6
YBJ5	2	103	Hypothetical 31.1 kDa protein in SIP18-SPT21 intergenic region
RNA1	2	110	Tubulin beta chain (Beta tubulin)
TCP1	2	113	T-complex protein 1 subunit alpha (TCP-1-alpha) (CCT-alpha)
SAM1	2	114	S-adenosylmethionine synthetase 2
PSA1	2	116	Proteasome component PRE5
SEC2	2	118	Coatomer subunit epsilon (Epsilon-coat protein) (Epsilon-COP)
PNC1	2	126	Nicotinamidase (Nicotine deamidase) (NAMase)
AYR1	2	143	NADPH-dependent 1-acyldihydroxyacetone phosphate reductase
YEM6	2	162	Mannose-1-phosphate guanyltransferase

Table 2. Cdc45-3X-FLAGp immunoprecipitation was performed with the FLAG peptide. Protein with a petide number lower than 2 were not considered.

TABLE 3

Table 3. Mass spectrometry analysis of the proteins identified after Cdc45-3X-FLAGp immunopreicpitation in *rpd3* mutant cells treated with HU. For each protein the score, the peptide number and the function are reported.

Acc Num	Peptides	Score	Protein Description
CDC45	15	803	Cell division control protein 45
ADH1	12	633	Alcohol dehydrogenase 1
TDH3	11	660	Glyceraldehyde-3-phosphate dehydrogenase 3
YHB1	9	471	Flavohemoprotein
YMM5	8	489	Hypothetical 35.3 kDa protein
GRS1	8	449	Glycyl-tRNA synthetase 1
ERG27	6	285	3-keto-steroid reductase
MPG1	5	335	Mannose-1-phosphate guanyltransferase
YM32	5	249	Hypothetical 29.1 kDa protein
PLSC	5	224	Probable 1-acyl-sn-glycerol-3-phosphate acyltransferase
YRB4	5	213	Importin beta-4 subunit (Ran-binding protein YRB4)
TBA1	5	197	Tubulin alpha-1
GFA1	5	193	Glucosamine--fructose-6-phosphate aminotransferase
PSF2	5	179	DNA replication complex GINS protein PSF2
GRP78	4	223	78 kDa glucose-regulated protein homolog precursor
PSF3	4	223	DNA replication complex GINS protein PSF1
RAS2	4	220	Ras-like protein 2 precursor
YBQ2	4	219	Hypothetical 22.9 kDa protein
SLD5	4	213	DNA replication complex GINS protein SLD5
HHF2	4	205	Histone H4
AYR1	4	204	NADPH-dependent 1-acyldihydroxyacetone phosphate reductase
TRS31	4	201	Transport protein particle 31 kDa subunit
SEC72	4	200	Translocation protein SEC72 (p23)
ODC1	4	192	Mitochondrial 2-oxodicarboxylate carrier 1
MCM2	4	181	DNA replication licensing factor
CCC1	4	176	Protein CCC1
ERG25	4	167	C-4 methylsterol oxidase (Methylsterol monooxygenase)
RFC2	4	152	Activator 1 41 kDa subunit (Replication factor C subunit 2)
ARO10	4	146	Transaminated amino acid decarboxylase
SEC28	3	249	Coatomer subunit epsilon
ASC1	3	195	Guanine nucleotide-binding protein beta subunit-like protein
PTH2	3	192	Peptidyl-tRNA hydrolase 2 (PTH 2)
YJG6	3	190	Hypothetical 28.5 kDa protein
SPC3	3	182	Microsomal signal peptidase subunit
CYS4	3	174	Cystathionine beta-synthase
HTB1	3	170	Histone H2B.1
SAM1	3	161	S-adenosylmethionine synthetase 1
GET3	3	156	ATPase GET3 (Arsenical pump-driving ATPase)
RPN11	3	151	26S proteasome regulatory subunit RPN11 (MPR1 protein)
PRS8	3	145	26S protease regulatory subunit 8 homolog (SUG1 protein)
YHR020W	3	140	Putative prolyl-tRNA synthetase YHR020W
TUB2	3	140	Tubulin beta chain (Beta tubulin)
RSR1	3	133	Ras-related protein RSR1
YN8O	3	132	Hypothetical 28.6 kDa protein in DBP6-COQ2 intergenic region

MRT4	3	132	mRNA turnover protein 4
RFC3	3	131	Activator 1 40 kDa subunit (Replication factor C subunit 3)
OLA1	3	129	Putative GTP-binding protein YBR025C
YCP4	3	126	Hypothetical 26.4 kDa protein in CDC10-CIT2 intergenic region
STM1	3	121	Suppressor protein STM1 (GU4 nucleic-binding protein 2)
HOSC	3	119	Homocitrate synthase, cytosolic isozyme
ORT1	3	116	Mitochondrial ornithine carrier protein
RAS1	3	116	Ras-like protein 1 precursor
YPT1	3	115	GTP-binding protein YPT1 (Rab GTPase YPT1)
ADE5,7	3	108	Bifunctional purine biosynthetic protein ADE5,7
GRX7	3	96	Putative glutaredoxin-like protein YBR014C precursor
RPN9	3	94	26S proteasome regulatory subunit RPN9
YMC2	3	91	Carrier protein YMC2, mitochondrial precursor
SEC66	3	91	Translocation protein SEC66 (Protein HSS1)
CTF4	2	134	DNA polymerase alpha-binding protein (CTF4 protein)
TCP1	2	133	T-complex protein 1 subunit alpha (TCP-1-alpha) (CCT-alpha)
SDH1	2	126	Succinate dehydrogenase iron-sulfur protein
ARB1	2	124	Probable ATP-dependent transporter YER036C
YG4Y	2	121	Hypothetical 26.9 kDa protein in YHB1-PFK1 intergenic region
PDB1	2	118	Pyruvate dehydrogenase E1 component subunit beta
PRS4	2	116	26S protease regulatory subunit 4 homolog
TF2B	2	114	Transcription initiation factor IIB
QCR2	2	109	Ubiquinol-cytochrome-c reductase complex core protein 2
CTP1	2	103	Tricarboxylate transport protein (Citrate transport protein)
ODPA	2	99	Pyruvate dehydrogenase E1 component alpha subunit
CCT4	2	99	T-complex protein 1 subunit gamma (TCP-1-gamma)
ACO1	2	96	Acyl-CoA desaturase 1 (Stearoyl-CoA desaturase 1)
SEC11	2	96	Signal sequence-processing protein SEC11
SEC17	2	95	Alpha-soluble NSF attachment protein (SNAP-alpha
APA1	2	95	Protein APA1
GRX2	2	94	Glutaredoxin-2, mitochondrial precursor
DPB3	2	94	Protein YBR287W
GCH1	2	93	GTP cyclohydrolase I (GTP-CH-I)
YNP6	2	92	Hypothetical 33.7 kDa protein in YGP1-YCK2 intergenic region
ERG9	2	92	Squalene synthetase
SC61	2	87	Protein transport protein SEC61 alpha subunit
COP1	2	86	Coatomer subunit delta (Delta-coat protein) (Delta-COP)
YGL039W	2	86	Hypothetical protein YGL039W
HTZ1	2	83	Histone H2A variant (Histone H2A.Z)
YN28	2	81	Hypothetical 41.2 kDa protein in PLC1-SEC21 intergenic region
RV161	2	81	Reduced viability upon starvation protein 161
CKA1	2	77	Casein kinase II subunit alpha' (CK II)
ERG13	2	77	Hydroxymethylglutaryl-CoA synthase (HMG-CoA synthase)
NAT2	2	77	N-terminal acetyltransferase 2
ORM1	2	77	Protein ORM1
GOS1	2	75	Protein transport protein GOS1 (Golgi SNARE protein 1)
ALG5	2	73	Dolichyl-phosphate beta-glucosyltransferase
IRC4	2	73	Putative oxidoreductase YIR036C
ERG1	2	72	Squalene monooxygenase (Squalene epoxidase)
CDS1	2	69	Phosphatidate cytidylyltransferase
YPT10	2	67	GTP-binding protein YPT10
RPN10	2	64	26S proteasome regulatory subunit RPN10
PPT1	2	63	26S protease regulatory subunit 7 homolog (CIM5 protein)
MCK1	2	63	Protein kinase MCK1

YN86	2	61	Hypothetical 99.5 kDa protein in URK1-SMM1 intergenic region
MCM5	2	61	Minichromosome maintenance protein 5
CCT6	2	61	T-complex protein 1 subunit theta (TCP-1-theta) (CCT-theta)
ABF2	2	60	ARS-binding factor 2, mitochondrial precursor
FAR3	2	60	Factor arrest protein 3
YK68	2	60	Hypothetical 38.3 kDa protein
TYB	2	55	Transposon Ty1 protein B

Table 3. Cdc45-3X-FLAGp immunoprecipitation was performed with the FLAG peptide. Protein with a petide number lower than 2 were not considered.

TABLE 4

Table 4. Mass spectrometry analysis of the proteins identified after Cdc45-3X-FLAGp immunopreicpitation in WT cells treated with HU. For each protein the score, the peptide number and the function are reported.

Description	Score	Peptides	Description
MCM3	1372	24	Component of the Mcm2-7 hexameric complex
TOF1	1240	18	Subunit of a replication-pausing checkpoint complex
MCM7	927	16	Component of the Mcm2-7 hexameric complex
CDC45	897	15	DNA replication initiation factor
CDC54	928	14	Component of the Mcm2-7 hexameric complex
MCM6	690	13	Component of the Mcm2-7 hexameric complex
MCM2	795	12	Component of the Mcm2-7 hexameric complex
MCM5	643	11	Component of the Mcm2-7 hexameric complex
SPT16	472	10	Subunit of the heterodimeric FACT complex
FET4	515	9	Low-affinity Fe(II) transporter of the plasma membrane
PET9	455	7	Major ADP/ATP carrier of the mitochondrial membrane
CDC19	381	7	Pyruvate kinase
MIR1	335	6	Mitochondrial phosphate carrier
VMA2	269	6	Subunit B of the peripheral membrane domain
ACT1	286	5	Actin, structural protein
POB3	246	5	Subunit of the heterodimeric FACT complex
CTF4	296	5	Chromatin-associated protein
ILV2	270	5	Acetolactate synthase
YDJ1	270	5	Protein chaperone regulating of the HSP90 and HSP70
YHM2	288	4	Carrier protein
SLD5	231	4	Subunit of the GINS complex
TEF1	191	4	Translational elongation factor EF-1 alpha
PGA3	210	4	Putative cytochrome b5 reductase
SEC4	179	4	Rab family GTPase
GGC1	149	4	Mitochondrial GTP/GDP transporter
TCB1	220	3	Lipid-binding protein
SSA2	146	3	ATP binding protein
HHF1	128	3	Histone H4
PSF1	118	2	Subunit of the GINS complex
SSA1	73	2	ATPase
PSF3	122	2	Subunit of the GINS complex
HSP26	132	2	Small heat shock protein
PHO86	122	2	Endoplasmic reticulum (ER) resident protein
PSA1	111	2	GDP-mannose pyrophosphorylase
HTA2	104	2	Histone H2A
KAR2	102	2	ATPase involved in protein import into the ER
PHO84	98	2	High-affinity inorganic phosphate (Pi) transporter
BNA4	88	2	Kynurenine 3-mono oxygenase,
YHR048W	60	2	Presumed antiporter of the DHA1 family
TUB1	78	2	Alpha-tubulin
CSM3	73	2	Replication fork associated factor
RPP0	81	2	Conserved ribosomal protein P0 of the ribosomal stalk
POR1	76	2	Mitochondrial porin

RHO3	68	2	Non-essential small GTPase of the Rho/Rac subfamily
OAC1	67	2	Mitochondrial inner membrane transporter
ORT1	54	2	Ornithine transporter of the mitochondrial membrane

Table 4. Cdc45-3X-FLAGp immunoprecipitation was performed with PorteinG-Dynabeads. Protein with a petide number lower than 2 were not considered.

TABLE 5

Table 5. Mass spectrometry analysis of Cdc45-3X-FLAgp immunoprecipitated samples in *rdp3* cells after HU treatment. For each protein the score, the peptide number and the function are reported.

Name	Score	Peptides	Description
CDC45	956	17	DNA replication initiation factor
TOF1	949	15	Subunit of a replication-pausing checkpoint complex
FET4	740	14	Low-affinity Fe(II) transporter of the plasma membrane
MCM7	782	14	Component of the Mcm2-7 hexameric complex
MCM6	671	14	Component of the Mcm2-7 hexameric complex
MCM3	777	13	Component of the Mcm2-7 hexameric complex
MCM2	786	12	Component of the Mcm2-7 hexameric complex
SPT16	612	11	Subunit of the heterodimeric FACT complex
MCM4	598	8	Component of the Mcm2-7 hexameric complex
CDC19	435	8	Pyruvate kinase
MCM5	425	8	Component of the Mcm2-7 hexameric complex
SSA1	362	6	ATPase
MIR1	339	6	Mitochondrial phosphate carrier
CSM3	258	6	Replication fork associated factor
ILV2	261	6	Acetolactate synthase
CTF4	363	5	Chromatin-associated protein
VMA2	297	5	Subunit B of the peripheral membrane domain
PSA1	244	5	GDP-mannose pyrophosphorylase
YDJ1	225	5	Protein chaperone regulating of the HSP90 and HSP70
PGA3	298	4	Putative cytochrome b5 reductase
YHM2	226	4	Carrier protein
PET9	221	4	Major ADP/ATP carrier of the mitochondrial membrane
KAR2	219	4	ATPase involved in protein import into the ER
SEC4	202	4	Rab family GTPase
PMA1	228	4	Plasma membrane H+-ATPase
PHO86	231	4	Endoplasmic reticulum (ER) resident protein
ARO10	219	4	Phenylpyruvate decarboxylase
HHF1	190	4	Histone H4
SLD5	208	4	Subunit of the GINS complex
URA2	194	4	Bifunctional carbamoylphosphate synthetase
POR1	205	4	Mitochondrial porin
GGC1	176	4	Mitochondrial GTP/GDP transporter
GRS1	183	4	Cytoplasmic and mitochondrial glycyl-tRNA synthase
TCB1	202	3	Lipid-binding protein
ACT1	151	3	Actin, structural protein
SSB2	141	3	Cytoplasmic ATPase
TEF1	143	3	Translational elongation factor EF-1 alpha
ALG2	152	3	Mannosyltransferase
OLE1	143	3	Delta(9) fatty acid desaturase
OAC1	130	3	Mitochondrial inner membrane transporter
GFA1	138	3	Glutamine-fructose-6-phosphate amidotransferase
TUB1	124	3	Alpha-tubulin
ATP2	124	3	Beta subunit of the F1 sector
TFP1	99	3	Subunit A of the vacuolar H+-ATPase
AAC3	93	2	Mitochondrial inner membrane ADP/ATP translocator
POB3	117	2	Subunit of the GINS complex
YPT31	125	2	Rab family GTPase
HSP26	150	2	Small heat shock protein
NDE1	103	2	Mitochondrial external NADH dehydrogenase

YHR020W	121	2	Protein of unknown function; may interact with ribosomes
PSF1	108	2	Subunit of the GINS complex
PHO84	115	2	High-affinity inorganic phosphate (Pi) transporter
RSR1	112	2	GTP-binding protein
ADH1	99	2	Alcohol dehydrogenase
RPP0	108	2	Conserved ribosomal protein P0 of the ribosomal stalk
ALG14	107	2	Component of UDP-GlcNAc transferase
HTA2	86	2	Histone H2A
FAA1	96	2	Long chain fatty acyl-CoA synthetase
RPS3	90	2	Protein component of the small (40S) ribosomal subunit
ATP1	88	2	Alpha subunit of the F1 sector
LCB2	65	2	Component of serine palmitoyltransferase
CBR1	55	2	Microsomal cytochrome b reductase

Table 5. Cdc45-3X-FLAGp was immunoprecipitated with ProteinG-Dynalbeads. Proteins with a peptide number lower than 2 were not considered.

TABLE 6

Table 6. Mass spectrometry analysis of Cdc45-3X-FLAgp immunoprecipitated samples in WT cells arrested in G1 and then released into S-phase. For each protein the score, the peptide number and the function are reported.

Name	Score	Peptides	Description
MCM6	1192	17	Component of the Mcm2-7 hexameric complex
CDC45	772	15	DNA replication initiation factor
MCM2	862	14	Component of the Mcm2-7 hexameric complex
MCM3	766	14	Component of the Mcm2-7 hexameric complex
MCM4	750	13	Component of the Mcm2-7 hexameric complex
MCM7	596	12	Component of the Mcm2-7 hexameric complex
SPT16	654	11	Subunit of the heterodimeric FACT complex (Spt16p-Pob3p)
TOF1	586	11	Subunit of a replication-pausing checkpoint complex
MCM5	580	11	Component of the Mcm2-7 hexameric complex
FET4	508	10	Low-affinity Fe(II) transporter of the plasma membrane
TEF1	264	5	Translational elongation factor EF-1 alpha
YHM2	333	4	Carrier mithocondrial protein
ACT1	196	4	Actin, structural protein
CTF4	236	4	Chromatin-associated protein
PET9	224	4	Major ADP/ATP carrier of the mitochondrial inner membrane
PMA1	209	3	Plasma membrane H+-ATPase
CSM3	151	3	Replication fork associated factor
GGC1	118	3	Mitochondrial GTP/GDP transporter
HTB2	92	3	Histone H2B
SLD5	186	2	Subunit of the GINS complex
CDC16	142	2	Subunit of the anaphase-promoting complex/cyclosome
KAR2	102	2	ATPase involved in protein import into the ER,
HHF1	103	2	Histone H4
POB3	90	2	Subunit of the heterodimeric FACT complex
MIR1	98	2	Mitochondrial phosphate carrier
NDE1	97	2	Mitochondrial external NADH dehydrogenase
CBR1	91	2	Microsomal cytochrome b reductase
RPP2B	79	2	Ribosomal protein P2 beta
YBL005W-B	76	2	Transcriptional activator of the drug resistance network
RPL15A	68	2	Protein component of the large (60S) ribosomal subunit

Table 6. Cdc45-3X-FLAGp immunoprecipitation was performed with ProteinA-Dynabeads. Protein with a petide number lower than 2 were not considered.

8 BIBLIOGRAPHY

Alvino, G. M., Collingwood, D., Murphy, J.M., Delrow, J., Brewer, J.B. and Raghuraman, M.K. (2007). "Replication in hydroxyurea: it's a matter of time." *Mol Cell Biol* **27**(18): 6396-406.

Aparicio, J. G., Viggiani, C. J., Gibson, D.G. and Aparicio, O.M. (2004). "The Rpd3-Sin3 histone deacetylase regulates replication timing and enables intra-S origin control in Saccharomyces cerevisiae." *Mol Cell Biol* **24**(11): 4769-80.

Aparicio, O. M., Stout, A. M. and Bell, S.P. (1999). "Differential assembly of Cdc45p and DNA polymerases at early and late origins of DNA replication." *Proc Natl Acad Sci U S A* **96**(16): 9130-5.

Aparicio, O. M., Weinstein, D. M. and Bell, S.P. (1997). "Components and dynamics of DNA replication complexes in S. cerevisiae: redistribution of MCM proteins and Cdc45p during S phase." *Cell* **91**(1): 59-69.

Aparicio, T., Guillou, E., Coloma, J. Montoya, G. and Mendez, J. (2009). "The human GINS complex associates with Cdc45 and MCM and is essential for DNA replication." *Nucleic Acids Res* **37**(7): 2087-95.

Aparicio, T., Ibarra, A. and Mendez, J. (2006). "Cdc45-MCM-GINS, a new power player for DNA replication." *Cell Div* **1**: 18.

Araki, H. (2009). "Regulatory mechanism of the initiation step of DNA replication by CDK in budding yeast." *Biochim Biophys Acta* **1804**(3): 520-3.

Araki, H., Leem, S. H., Phongdara, A. and Sugino, A. (1995). "Dpb11, which interacts with DNA polymerase II(epsilon) in Saccharomyces cerevisiae, has a dual role in S-phase progression and at a cell cycle checkpoint." *Proc Natl Acad Sci U S A* **92**(25): 11791-5.

Araki, H., Nakanishi, N., Evans, B.R., Matsuzaki, H., Jarayam, M. and Oshima, Y. (1992). "Site-specific recombinase, R, encoded by yeast plasmid pSR1." *J Mol Biol* **225**(1): 25-37.

Arias, E. E. and Walter, J. C. (2007). "Strength in numbers: preventing rereplication via multiple mechanisms in eukaryotic cells." *Genes Dev* **21**(5): 497-518.

Baker, T. A. and Bell S. P. (1998). "Polymerases and the replisome: machines within machines." *Cell* **92**(3): 295-305.

Bando, M., Katou, Y., Komata, M., Tanaka, H., Itoh, T., Sutani, T. and Shirahige, K. (2009). "Csm3, Tof1, and Mrc1 form a heterotrimeric complex mediator complex that associates with DNA replication forks." *J Biol Chem* **284**(49): 34355-65.

Barry, E.R. and Bell, S.D. (2006). "DNA replication in the acrhaea." *Microbiol Mol Biol Rev* **70**(4): 876-87.

Bell, S. P. (2002). "The origin recognition complex:from single origins to complex functions." *Genes Dev* **15**(6): 659-72.

Bell, S. P. and Dutta A. (2002). "DNA replication in eukaryotic cells." *Annu Rev Biochem* **71**:333-74.

Bell, S. P. and Stillman, B. (1992). "ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex." *Nature* **357**(6374): 128-34.

Benjamin, J. M., Torke, S. J., Demeler, B. and McGarry, T.J. (2004). "Geminin has dimerization, Cdt1-binding, and destruction domains that are required for biological activity." *J Biol Chem* **279**(44): 45957-68.

Biswas, D., Takahata, S., Xin, H., Dutta-Biswas, R., Yu, Y., FOrmosa, T. and Stillma, D.J. (2008). "A role for Chd1 and Set2 in negatively regulating DNA replication in Saccharomyces cerevisiae." *Genetics* **178**(2): 649-59.

Black, J.C., Allen, A., Van Rechem, C., Forbes, E., Longworth, M., Toschop, K., Rinehart, C., Quiton, J., Walsh, R., Smallwood, A., Dyson, N.J. and Whetsine, J.R. (2010). "Conserved antagonism between JMJD2A/KDM4A and HP1γ during cell cycle progression." *Mol Cell* **40**(5): 736-48.

Bloom, J. and Cross, F. R. (2007). "Multiple levels of cyclin specificity in cell-cycle control." *Nat Rev Mol Cell Biol* **8**(2): 149-60.

Blow, J. and Dutta, A. (2005). "Preventing re-replication of chromosomal DNA." *Nat Rev Mol Cell Biol* **6**(6): 476-86.

Blow, J. and Ge, X. Q. (2008). "Replication forks, chromatin loops and dormant replication origins." *Genome Biol* **9**(12): 244.

Blow, J. and Ge, X. Q. (2009). "A model for DNA replication showing how dormant origins safeguard against replication fork failure." *EMBO Rep* **10**(4): 406-12.

Blow, J., Gillespie, P. J., Francis, D. and Jackson, D.A. (2001). "Replication origins in Xenopus egg extract are 5-15 kilobases apart and are activated in clusters that fire at different times." *J Cell Biol* **152**(1): 15-25.

Bolon, Y.T. and Bielinsky, A.K. (2006). "The spatial arrangement of ORC binding modules determines the functionality of replication origins in budding yeast." *Nucleic Acid Res* **34**(18): 5069-80.

Bowers, J. L., Randell, J. C., Chen, S. and Bell, S.P. (2004). "ATP hydrolysis by ORC catalyzes reiterative Mcm2-7 assembly at a defined origin of replication." *Mol Cell* **16**(6): 967-78.

Branzei, D. and Foiani, M. (2007). "Interplay of replication checkpoint and repair proteins at stalled replication forks." *DNA Repair* **6**(7): 994-1003.

Brewer, B. J. and Fangman, W. L. (1987). "The localization of replication origins on ARS plasmids in S. cerevisiae." *Cell* **51**(3): 463-71.

Brewer, B. J. and Fangman, W. L. (1993). "Initiation at closely spaced replication origins in a yeast chromosome." *Science* **262**(5140): 1728-31.

Brush, G. S., Morrow, D. M., Hieter, P. and Kelly, T.J. (1996). "The ATM homologue MEC1 is required for phosphorylation of replication protein A in yeast." *Proc Natl Acad Sci U S A* **93**(26): 15075-80.

Burgers, P. M. (2009). "Polymerase dynamics at the eukaryotic DNA replication fork." *J Biol Chem* **284**(7): 4041-5.

Calzada, A., Hodgson, B., Kanemaki, M., Bueno, A. and Labib, K. (2005). "Molecular anatomy and regulation of a stable replisome at a paused eukaryotic DNA replication fork." *Genes Dev* **19**(16): 1905-19.

Chanet, R and Heude, M. (2003). "Characterization of mutations that are synthetic lethal with pol3-13, a mutant of DNA polymerase delta in Saccharomyces cerevisiae." *Curr Genet* **43**(5): 337-50.

Chang, Y. P., Wang, G., Bermudez, V., Hurwitz, J. and Chen, X.S. (2007). "Crystal structure of the GINS complex and functional insights into its role in DNA replication." *Proc Natl Acad Sci U S A* **104**(31): 12685-90.

Chen, S., de Vries, M. A. and Bell, S.P. (2007). "Orc6 is required for dynamic recruitment of Cdt1 during repeated Mcm2-7 loading." *Genes Dev* **21**(22): 2897-907.

Chen, Z., Speck, C., Wendel, P., Tang, C., Stillman, B. and Li, H. (2008). "The architecture of the DNA replication origin recognition complex in Saccharomyces cerevisiae." *Proc Natl Acad Sci U S A* **105**(30): 10326-31.

Choudary, C., Kumar, C., Gnad, F., Nielsen, M.L., Rehman, M., Walther, T.C., Olsen, J.V. and Mann, M. (2009). "Lysine acetylation targets protein complexes and co-regulates major cellular functions." *Science* **325**(5942): 834-40.

Cimbora, D. M., Schubeler, D., Reik, A., Hamilton, J., Francastel, C., Epner, E. and Groudine, M. (2000). "Long-distance control of origin choice and replication timing in the human beta-globin locus are independent of the locus control region." *Mol Cell Biol* **20**(15): 5581-91.

Cobb, J. A., Bjergbaek, L., Shimada, K., Frei, C. and Gasser, S.M. (2003). "DNA polymerase stabilization at stalled replication forks requires Mec1 and the RecQ helicase Sgs1." *EMBO J* **22**(16): 4325-36.

Cobb, J. A., Schleker, T., Rojas, V., Bjergbaek, L., Tercero, J.A. and Gasser, S.M. (2005). "Replisome instability, fork collapse, and gross chromosomal rearrangements arise synergistically from Mec1 kinase and RecQ helicase mutations." *Genes Dev* **19**(24): 3055-69.

Cosgrove, A. J., Nieduszynski, C. A. and Donaldson, A.D. (2002). "Ku complex controls the replication time of DNA in telomere regions." *Genes Dev* **16**(19): 2485-90.

Costa, S. and Blow, J.J. (2007). "The elusive determinants of replication origins." *EMBO Rep* **8**(4): 332-4.

Coverley, D., Pelizon, C., Trewick, S. and Laskey, R.A. (2000). "Chromatin-bound Cdc6 persists in S and G2 phases in human cells, while soluble Cdc6 is destroyed in a cyclin A-cdk2 dependent process." *J Cell Sci* **113** (11): 1929-38.

Cvetic, C. A. and Walter, J. C. (2006). "Getting a grip on licensing: mechanism of stable Mcm2-7 loading onto replication origins." *Mol Cell* **21**(2): 143-4.

Czajkowsky, D. M., Liu, J., Hamlin, J.L. and Shao, Z. (2008). "DNA combing reveals intrinsic temporal disorder in the replication of yeast chromosome VI." *J Mol Biol* **375**(1): 12-9.

Da-Silva, L. F. and Duncker, B. P. (2007). "ORC function in late G1: maintaining the license for DNA replication." *Cell Cycle* **6**(2): 128-30.

Davey, M.J., Fang, L., McInerney, P., Gergescu, R.E. and O'Donnell, M. (2002). "The DnaC helicase loader is a dual ATP/ADP switch protein." *EMBO J.* **21**(12): 3148-59.

Davey, M.G., Indiani, C. and O'Donnell, M. (2003). "Reconstitution of the Mcm2-7p heterohexamer, subunit arrangement, and ATP site architecture." *J Biol Chem* **278**(7): 449-9.

DePamphilis, M. L., Blow, J. J., Ghosh, S., Saha, T., Noguchi, K. and Vassilev, A. (2006). "Regulating the licensing of DNA replication origins in metazoa." *Curr Opin Cell Biol* **18**(3): 231-9.

Diffley, J. F., Cocker, J.H., Dowell, S.J., Harwood, J. and Rowley, A. (1995). "Stepwise assembly of initiation compelxes at budding yeast replication origins during cell cycle." *J.Cell.Sci* **19**: 67-72.

Diffley, J. F. (2004). "Regulation of early events in chromosome replication." *Curr Biol* **14**(18): 778-86.

Diffley, J. F. and Labib, K. (2002). "The chromosome replication cycle." *J Cell Sci* **115**(5): 869-72.

Donaldson, A. D. (2005). "Shaping time: chromatin structure and the DNA replication programme." *Trends Genet* **21**(8): 444-9.

Donaldson, A. D., Raghuraman, M. K., Friedman, K., Cross, F.R., Brewer, B.J. and Fangman, W.L. (1998). "CLB5-dependent activation of late replication origins in S. cerevisiae." *Mol Cell* **2**(2): 173-82.

Drury, L. S., Perkins, G. and Diffley, J.F. (2000). "The cyclin-dependent kinase Cdc28p regulates distinct modes of Cdc6p proteolysis during the budding yeast cell cycle." *Curr Biol* **10**(5): 231-40.

Ducker, C.E. and Simpson, R.T. (2000). "The organized chromatin domain of the repressed yeast a cell-specific gene STE6 contains two molecules of the corepressor Tup1p per nucleosome." *EMBO J* **19**(3): 400-9.

Duncker, B. P., Chesnokov, I. N. and McConkey, B.J. (2009). "The origin recognition complex protein family." *Genome Biol* **10**(3): 214.

Duncker, B. P., Shimada, K., Tsai-Pflugfelder, M., Pasero, P. and Gasser, S.M. (2002). "An N-terminal domain of Dbf4p mediates interaction with both origin recognition complex (ORC) and Rad53p and can deregulate late origin firing." *Proc Natl Acad Sci U S A* **99**(25): 16087-92.

Eaton, M. L., Galani, K., Kang, S., Bell, S.P. and MacAlpine, D.M. (2010) "Conserved nucleosome positioning defines replication origins." *Genes Dev* **24**(8): 748-53.

Ebrahimi, H., Robertson, E. D., Taddei, A., Gasser, S.M., Donaldson, A.D. and Hiraga, S. (2010) "Early initiation of a replication origin tethered at the nuclear periphery." *J Cell Sci* **123**(7): 1015-9.

Espinosa, M. C., Rehman, M. A., Chiasamone-Robert, P., Jeffrey, D. and Yankulov, K. (2010) "GCN5 is a positive regulator of origins of DNA replication in Saccharomyces cerevisiae." *PLoS One* **5**(1): e8964.

Evrin, C., Clarke, P., Zech, J. Lurz, R., Sun, J., Uhle, S., Li, H., Stillman, B. and Speck, C. (2009) "A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication." *PNAS* **106**(48):20240-20245.

Fang, L., Davey, M.J. and O'Donnell, M. (1999). "Replisome assembly at oriC, the replication origin of E. coli, reveals an explanation for initiation sites outside an origin." *Mol. Cell* **4**(4): 541-53.

Feng, W., Collingwood, D., Boeck, M.E., Fox, L.A., Alvino, G.M., Fangman, W., Raghuraman, M.K. and Brewer, B.J. (2006). "Genomic mapping of single-stranded DNA in hydroxyurea-challenged yeasts identifies origins of replication." *Nat Cell Biol* **8**(2): 148-55.

Feng, W., Raghuraman, M. K. and Brewer, B.J. (2007). "Mapping yeast origins of replication via single-stranded DNA detection." *Methods* **41**(2): 151-7.

Ferguson, B. M., Brewer, B. J., Reynolds, A.E. and Fangman, W.L. (1991). "A yeast origin of replication is activated late in S phase." *Cell* **65**(3): 507-15.

Ferguson, B. M. and Fangman, W. L. (1992). "A position effect on the time of replication origin activation in yeast." *Cell* **68**(2): 333-9.

Ferguson, R.L. and Maller, J.L. (2008). "Cyclin E-dependent localization of MCM5 regulates centrosome duplication." *J. Cell Sci.* **121**(Pt 19): 3324-32.

Ferreira, M.F., Santocanale, C., Drury, L.S. and Diffley, J.F. (2000). "Dbf4p, an essential S pahse-promoting factor, is targeted for degradation by the anaphase-promoting complx." *Mol Cell Biol* **20**(1): 242-8.

Forsburg, S.L. (2004). "Eukaryotic MCM proteins:beyond replication initiation." *Microbiol Mol Biol Rev* **68**(1): 109-31.

Fox, C.A., Ehrenhofer-Murray, A.E., Loo, S. and Rine, J. (1997). "The origin recognition complx, SIR1, and the S pahse requirement for silnecing." *Science* **276**(5318): 1547-51.

Francis, L. I., Randell, J. C., Takara, T.J. Uchim, L. and Bell, S.P. (2009). "Incorporation into the prereplicative complex activates the Mcm2-7 helicase for Cdc7-Dbf4 phosphorylation." *Genes Dev* **23**(5): 643-54.

Friedman, K. L., Brewer, B. J. and Fangman, W.L. (1997). "Replication profile of *Saccharomyces cerevisiae* chromosome VI." *Genes Cells* **2**(11): 667-78.

Friedman, K. L., Diller, J. D., Ferguson, B. M., Nyland, S.V., Brewer, B.J. and Fangman, W. L. (1996). "Multiple determinants controlling activation of yeast replication origins late in S phase." *Genes Dev* **10**(13): 1595-607.

Friedman, K. L., Raghuraman, M. K., Fangman, W. L. and Brewer, B.J. (1995). "Analysis of the temporal program of replication initiation in yeast chromosomes." *J Cell Sci Suppl* **19**: 51-8.

Fuller, R.S., Funnel, B.E. and Kornberg, A. (1984). "The dnaA protein complex with the E. coli chromosomal replication origin (oriC) and other DNA sites." *Cell* **38**(3): 889-900.

Gambus, A., Jones, R. C., Sanchez-Diaz, A., Kanemaki, M., van Deursen, F., Edmondson, R.D. and Labib, K. (2006). "GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks." *Nat Cell Biol* **8**(4): 358-66.

Ge, X.Q., Jackson, D.A. and Blow, J. (2007). "Dormant origins licensed by excess Mcm2-7 are required for human cells to survive replicative stress." *Genes Dev* **21**(24):3331-41.

Gibson, D. G., Bell, S. P. and Aparicio, O.M. (2006). "Cell cycle execution point analysis of ORC function and characterization of the checkpoint response to ORC inactivation in *Saccharomyces cerevisiae*." *Genes Cells* **11**(6): 557-73.

Gilbert, D. M. (2001). "Nuclear position leaves its mark on replication timing." *J Cell Biol* **152**(2): F11-5.

Gomes, X.V. and Burgers, P.M. (2001). "ATP utilization by yeast replication factor C. I. ATP-mediated interaction with DNA and with prolifrerating cell nuclear antigen." *J Biol Chem* **276**(37): 34768-75.

Goren, A., Tabib, A., Hetch, M. and Cedar, H. (2008). "DNA replication timing of the human beta-globin domain is controlled by histone modification at the origin." *Genes Dev* **22**(10): 1319-24.

Green, B.M., Morreale, R.J., Ozadyn, B., Derisi, J.L. and Li, J.J. (2006). "Genome-wide mapping of DNA synthesis in Saccharomyces cerevisiae reveals that mechanisms preventing reinitiation of DNA replication are not redundant." *Mol Biol Cell* **17**(5): 2401-14.

Gregan, J., Lindner, K., Brimage, L., Franklin, R., Namdar, M., Hart, E.A., Aves, S.J. and Kearsey, E. (2003). "Fission yeast Cdc23/Mcm10 functions after pre-replicative complex fromation to promote Cdc45 chromatin binding." *Mol Biol Cell* **14**: 3876-87.

Guthrie, C. and Frink, G.R. (2002). "Guide to yeast genetics and molecular and cell biology." *Methods of Enzymology* **350** and **351**.

Han, J., Zhou, H., Horazdovsky, B., Zhang, K., Xu, R.M. and Zhang, Z. (2007). "Rtt109 acetylates histone H3 lysine 56 and functions in DNA replication." *Science* **315**(5812): 653-5.

Hardy, C. F. (1997). "Identification of Cdc45p, an essential factor required for DNA replication." *Gene* **187**(2): 239-46.

Hayashi, M., Katou, Y., Itoh, T., Tazumi, A., Yamada, Y., Takahashi, T., Nakagawa, T., Shirahige, K. and Masukata, H. (2007). "Genome-wide localization of pre-RC sites and identification of replication origins in fission yeast." *EMBO J* **26**(5): 1327-39.

Hemerley, A.S., Prasanth, S.G., Siddiqui, K. and Stillman, B. (2009). "Orc1 controls centrole and centrosome copy number in human cells." *Science* **323**(5915): 789-93.

Heun, P., Laroche, T., Raghuraman, M.K. and Gasser, S.M. (2001). "The positioning and dynamics of origins of replication in the budding yeast nucleus." *J Cell Biol* **152**(2): 385-400.

Hiraga, S., Botsios, S. and Donaldson, A.D. (2008). "Histone H3 lysine 56 acetylation by Rtt109 is crucial for chromosome positioning." *J Cell Biol* **183**(4): 641-51.

Hiraga, S., Hagihara-Hayashi, A., Ohya, T. and Sugino, A. (2005). "DNA polymerases alpha, delta, and epsilon localize and function together at replication forks in Saccharomyces cerevisiae." *Genes Cells* **10**(4): 297-309.

Hiratani, I. and D. M. Gilbert (2009). "Replication timing as an epigenetic mark." *Epigenetics* **4**(2): 93-7.

Hiratani, I., Takebayashi, S., Lu, J. and Gilbert, D.M. (2009). "Replication timing and transcriptional control: beyond cause and effect--part II." *Curr Opin Genet Dev* **19**(2): 142-9.

Hodgson, B., Calzada, A. and Labib, K. (2007). "Mrc1 and Tof1 regulate DNA replication forks in different ways during normal S phase." *Mol Biol Cell* **18**(10): 3894-902.

Homesley, L., Lei, M., Kawasaki, Y., Sawyer, S., Christensen, T. and Tye, B.K. (2000). "Mcm10 and the MCM2-7 complex interact to initiate DNA synthesis and to release replication factors from origins." *Genes Dev* **14**(8): 913-26.

Huang, M-E. and Kolodner, R.D. (2005). "A biological network in Saccharomyces cerevisiae prevents deleterious effects of endogenous oxidative DNA damage." *Mol Cell* **17**:709-20.

Ibarra, A., Schwob, E. and Mendez, J. (2008). "Excess MCM proteins protect human cells from replicative stress by licensing backup origins of replication." *Proc Natl Acad Sci U S A* **105**(26): 8956-61.

Ilves, I., Petojevic, T., Pesavento, J.J. and Botchan, M.R. (2010) "Activation of the MCM2-7 helicase by association with Cdc45 and GINS proteins." *Mol Cell* **37**(2): 247-58.

Im, J. S., Ki, S. H., Farina, A., Jung, D.S., Hurwitz, J. and Lee, J.K. (2009). "Assembly of the Cdc45-Mcm2-7-GINS complex in human cells requires the Ctf4/And-1, RecQL4, and Mcm10 proteins." *Proc Natl Acad Sci U S A* **106**(37): 15628-32.

Ivanov, D. and Nasmyth, K. (2005). "A topological interaction between cohesin rings and a circular minichromosome." *Cell* **122**(6): 849-60.

Jacob, F., Brenner, S. and Cuzin, F. (1963). "On the regulation of DNA replication in bacteria." *Cold Spring Harbor Symp Quant Biol* **28**: 329-347

Jackson, A.L., Pahl, P.M., Harrison, K., Rosamond, J. and Sclafani, R.A. (1993). "Cell cycle regualtion of the yeast Cdc7 protein kinase by association with the Dbf4 protein." *Mol Cell Biol* **13**(5): 2899-908.

Jeon, Y., Bekiranov, S., Karnani, N., Kapranov, P., Ghosh, S., MacAlpine, D., Lee, C., Hwang, D.S., Ginferas, T.R. and Dutta, A. (2005). "Temporal profile of replication of human chromosomes." *Proc Natl Acad Sci U S A* **102**(18): 6419-24.

Johansson, E., Garg, P. and Burgers, P.M. (2004). "The Pol32 subunit of DNA polymerase delta contains separable domains for processive replication and proliferating cell nuclear antigen (PCNA) binding." *J Biol Chem* **279**(3): 1907-15.

Kamimura, Y., Masumoto, H., Sugino, A. and Araki, H. (1998). "Sld2, which interacts with Dpb11 in Saccharomyces cerevisiae, is required for chromosomal DNA replication." *Mol Cell Biol* **18**(10): 6102-9.

Kamimura, Y., Tak, Y. S., Sugino, A. and Araki, H. (2001). "Sld3, which interacts with Cdc45 (Sld4), functions for chromosomal DNA replication in Saccharomyces cerevisiae." *EMBO J* **20**(8): 2097-107.

Kanemaki, M. and Labib, K. (2006). "Distinct roles for Sld3 and GINS during establishment and progression of eukaryotic DNA replication forks." *EMBO J* **25**(8): 1753-63.

Karnani, N., Taylor, C. M., Malhotra, A. and Dutta, A. (2010). "Genomic study of replication initiation in human chromosomes reveals the influence of transcription regulation and chromatin structure on origin selection." *Mol Biol Cell* **21**(3): 393-404.

Katou, Y., Kanoh, Y., Bando, M., Noguchi, H., Tanaka, H., Ashikari, T., Sugimoto, K. and Shirahige, K. (2003). "S-phase checkpoint proteins Tof1 and Mrc1 form a stable replication-pausing complex." *Nature* **424**(6952): 1078-83.

Kawasaki, Y., Hiraga, S. and Sugino, A. (2000). "Interactions between Mcm10p and other replication factors are required for proper initiation and elongation of chromosomal DNA replication in Saccharomyces cerevisiae." *Genes Cells* **5**(12): 975-89.

Kelman, L.M. and Kelman, Z. (2003). "Archaea: an archetype for replication initiation studies?" *Mol Microbiol* **48**(3): 605-15.

Kelman, Z., Lee, J.K. and Hurwitz, J. (1999). "The single minichromosome maintenance protein of Methanobacterium thermiautotrophicum DeltaH contains DNA helicase activity." *Proc. Natl. Acad, Sci. USA* **96**(26): 14783-8.

Kim, Y., Shen, C. H. and Clark, D.J. (2004). "Purification and nucleosome mapping analysis of native yeast plasmid chromatin." *Methods* **33**(1): 59-67.

Klemm, R. D., Austin, R. J. and Bell, S.P. (1997). "Coordinate binding of ATP and origin DNA regulates the ATPase activity of the origin recognition complex." *Cell* **88**(4): 493-502.

Klemm, R. D. and Bell, S. P. (2001). "ATP bound to the origin recognition complex is important for preRC formation." *Proc Natl Acad Sci U S A* **98**(15): 8361-7.

Knott, S. R., Viggiani, C. J., Tavare, S. and Aparicio, O.M. (2009). "Genome-wide replication profiles indicate an expansive role for Rpd3L in regulating replication initiation timing or efficiency, and reveal genomic loci of Rpd3 function in Saccharomyces cerevisiae." *Genes Dev* **23**(9): 1077-90.

Kong, D. and DePamphilis, M.L. (2002). "Site-specific ORC binding, pre-replication complex assembly and DNA synthesis at Schizosaccharomyces pombe replication origins." *EMBO J* **21**(20):5567-76.

Kono, K., Nogami, S., Abe, M., Nishizawa, M., Morishita, S., Peelman, D. and Ohya, Y. (2008). "G1/S cyclin-dependent kinase regulates small GTPase Rho1p through phosphoylation of RhoGEF Tus1p in Saccharomyces cerevisiae." *Mol Biol Cell* **19**(4):1763-71.

Koren, A., Soifer, I. and Barakai, N. (2010) "MRC1-dependent scaling of the budding yeast DNA replication timing program." *Genome Res* **20**(6): 781-90.

Kramer, E.R., Scheuringer, N., Podtelejnikov, A.V., Mann, M. and Peters, J.M. (2000). "Mitotic regulation of the PAC activator proteins CDC20 and CDH1." *Mol Biol Cell* **11**(5): 1555-69.

Kumagai, A., Lee, J., Yoo, H.Y. and Dunphy, W.G. (2006). "TopBP1 activates the ATR-ATRIP complex." *Cell* **124**(5): 943-55.

Kunkel, T. A. and Burgers, P. M. (2008). "Dividing the workload at a eukaryotic replication fork." *Trends Cell Biol* **18**(11): 521-7.

Labib, K., Diffley, J.F. and Kearsey, S.E. (1999). "G1-phase and B-type cyclins exclude the DNA-replication factor Mcm4 from the nucleus." *Nature* **1**(7):415-22.

Labib, K. and Gambus, A. (2007). "A key role for the GINS complex at DNA replication forks." *Trends Cell Biol* **17**(6): 271-8.

Labit, H., Perewoska, I., Germe, T., Hyrien, O. and Marheineke, K. (2008). "DNA replication timing is deterministic at the level of chromosomal domains but stochastic at the level of replicons in Xenopus egg extracts." *Nucleic Acids Res* **36**(17): 5623-34.

Lambert, S., Watson, A., Sheedy, D.M., Martin, B. and Carr, A.M. (2005). "Gross chormosomal rearrangements and elevated recombin at an inducible site-specific replication fork barrier." *Cell* **121**(5): 689-702.

Lee, D. G. and Bell, S. P. (1997). "Architecture of the yeast origin recognition complex bound to origins of DNA replication." *Mol Cell Biol* **17**(12): 7159-68.

Lee, D. G., Makhov, A. M., Klemm, R.D., Griffith, J. and Bell, S.P. (2000). "Regulation of origin recognition complex conformation and ATPase activity: differential effects of single-stranded and double-stranded DNA binding." *EMBO J* **19**(17): 4774-82.

Lee, J.K. and Hurwitz, J. (2001). "Processive DNA helicase activity of the minichromosome maintenance proteins 4, 6 and 7 complex requires forked DNA structure." *Proc Natl Acad Sci USA* **98**(1): 54-9.

Lee, J.K., Moon, K.Y., Jiang, Y. and Hurwitz, J. (2001) "The Schizosaccharomyces pombe origin recognition complex interacts with multiple AT-rich regions of the replication origins DNA by means of the AT-hook domain of the spOrc4 protein." *Proc Natl Acad Sci U S A* **98**(24): 13588-94.

Lee, J. K., Seo, Y. S. and Hurwitz, J. (2003). "The Cdc23 (Mcm10) protein is required for the phosphorylation of minichromosome maintenance complex by the Dfp1-Hsk1 kinase." P*roc Natl Acad Sci U S A* **100**(5): 2334-9.

Lee, J. S. and Shilatifard, A. (2007). "A site to remember: H3K36 methylation a mark for histone deacetylation." *Mutat Res* **618**(1-2): 130-4.

Lengronne, A., McIntyre, J., Katou, Y., Kanoh, Y., Hopfned, K.P., Shirahige, K. and Uhlman, F. (2006). "Establishemnt of sister chromatid cohesion at the S. cerevisiae replication fork." *Mol Cell* **23**(6): 787-99.

Li, A. and Blow, J. (2005). "Cdt1 downregulation by proteolysis and geminin inhibition prevents DNA re-replication in Xenopus." *EMBO J* **24**(2): 395-404.

Li, Y.C., Cheng, T.H. and Gartenberg, M.R. (2001). "Establishment of transcriptional silencing in the absence of DNA replication." *Science* **291**(5504): 650-3.

Liang, C., Weinreich, M. and Stillman, B. (1995). "ORC and Cdc6p interact and determine the frequency of initiation of DNA replication in the genome." *Cell* **81**(5): 667-76.

Liku, M. E., Nguyen, V. Q., Rosales, A.W., Irie, K. and Li, J.J. (2005). "CDK phosphorylation of a novel NLS-NES module distributed between two subunits of the Mcm2-7 complex prevents chromosomal rereplication." *Mol Biol Cell* **16**(10): 5026-39.

Lindquist, S. and Kim, G. (1996). "Heat-shock protein 104 expression is sufficient for thermotolerance in yeast." *Proc Natl Acad Sci U S A* **93**(11): 5301-5306.

Lipford, J. R. and Bell, S. P. (2001). "Nucleosomes positioned by ORC facilitate the initiation of DNA replication." *Mol Cell* **7**(1): 21-30.

Longhese, M. P., Clerici, M. and Lucchini, G. (2003). "The S-phase checkpoint and its regulation in Saccharomyces cerevisiae." *Mutat Res* **532**(1-2): 41-58.

Longtine, M. S., McKenzie, A., Demarini, D.J., Shah, N.G., Wach, A., Brachat, A., Philippsen, P. and Pringle, J.R. (1998). "Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae." *Yeast* **14**(10): 953-61.

Lopes, M., Cotta-Ramusino, C., Pellicioli, A., Liberi, G., Plevani, P., Muzi-Faconi, M., Newlon, C.S. and Foiani, M. (2001). "The DNA replication checkpoint response stabilizes stalled replication forks." *Nature* **412**(6846): 557-61.

Lopez-Mosqueda, J., Maas, N.L., Jonsson, Z.O., Defazio-Eli, L.G., Wohlschlegel, J. and Toczyski, D.P. (2010). "Damage induced phosphorylation of Sld3 is important to block late origin firing." *Nature* **467**(7314): 479-83.

Lou, H., Komata, M., Katou, Y., Guan, Z., Reis, C.C., Budd, M., Shirahige, K. and Campbell, J.L. (2008). "Mrc1 and DNA polymerase epsilon function together in linking DNA replication and the S phase checkpoint." *Mol Cell* **32**(1): 106-17.

Louvion, J.F., Havaux-Copf, B. and Picard, D. (1993). "Fusion of GAL4-VP16 to a steroid-binding domain provides a tool for gratuitous induction of galactose-responsive genes in yeast." *Gene* **131**(1): 129-34.

Lucas, I. and Feng, W. (2003). "The essence of replication timing: determinants and significance." *Cell Cycle* **2**(6): 560-3.

Lutzman, M. and Mechali, M. (2008) "MCM9 binds Cdt1 and is required for the assembly of prereplication complexes." *Mol Cell* **31**(2): 190-200.

MacAlpine, D. M., Rodriguez, H. K. and Bell, S.P. (2004). "Coordination of replication and transcription along a Drosophila chromosome." *Genes Dev* **18**(24): 3094-105.

Machida, Y. J., Hamlin, J. L. and Dutta, A. (2005). "Right place, right time, and only once: replication initiation in metazoans." *Cell* **123**(1): 13-24.

Maiorano, D., Cuvier, O., Danis, E. and Mechali, M. (2005a). "MCM8 is and MCM2-7-related protein that functions as a DNA helicase during replication elongation and not initiation." *Cell* **120**:315-28.

Maiorano, D., Krasinska, L., Lutzman, M. and Mechali, M. (2005b). "Recombinant Cdt1 induces rereplication of G2 nuclei in Xenopus egg extracts." *Curr Biol* **15**(2): 146-53.

Maiorano, D., Lutzman, M. and Mechali, M. (2006). "MCM proteins and DNA replication." *Curr. Opin, Cell Biol.* **18**(2): 130-6.

Majka, J. and Burgers, P. M. (2007). "Clamping the Mec1/ATR checkpoint kinase into action." *Cell Cycle* **6**(10): 1157-60.

Mantiero, D., Clerici, M., Lucchini, G. and Longhese, M.P. (2007). "Dual role for Saccharomyces cerevisiae Tel1 in the checkpoint response to double-strand breaks." *EMBO Rep* **8**(4): 380-7.

Marahrens, Y. and Stillman, B. (1992). "A yeast chromosomal origin of DNA replication defined by multiple functional elements." *Science* **255**(5046): 817-23.

Marinsek, N., Barry, E.R., Makarova, K.S., Dionne, I., Koonin, E.V. and Bell, S.D. (2006). "GINS, a cnetral nexus in the archaeal DNA replication fork." *EMBO Rep* **7**(5): 539-45.

Masumoto, H., Sugino, A. and Araki, H. (2000). "Dpb11 controls the association between DNA polymerases alpha and epsilon and the autonomously replicating sequence region of budding yeast." *Mol Cell Biol* **20**(8): 2809-17.

Matsuno, K., Kumano, M., Kubota, Y., Hasimoto, Y. and Takisawa, H. (2006). "The N-terminal noncatalytic region of Xenopus RecQ4 is required for chromatin binding of DNA polymerase alpha in the initiation of DNA replication." *Mol Cell Biol* **26**(13): 4843-52.

Matsuzaki, H., Araki, H. and Oshima, Y. (1988). "Gene conversion associated with site-specific reomcbiantion in yeast plasmid pSR1." *Mol Cell Biol* **8**(2): 955-62.

Matsuzaki, H., Nakajima, R., Nishiyama, J., Araki, H. and Oshima, Y. (1990). "Chromosome engineering in Saccharomyces cerevisiae by using a site-specific recombiantion system of a yeast plasmid." *J Bacteriol* **172**(2): 610-8.

McCune, H. J., Danielson, L. S., Alvino, G.M., Collingwood, D., Delrow, J.J., Fangman, W.L., Brewer, B.J. and Raghuraman, M.K. (2008). "The temporal program of chromosome replication: genomewide replication in Δclb5 Saccharomyces cerevisiae." *Genetics* **180**(4): 1833-47.

McGarry, T.J. and Kirschner, M.W. (1998). "Geminin, an inhibitor of DNA replication, is degraded during mitosis." *Cell* **93**(6): 1043-53.

Mechali, M. (2001). "DNA replication origins: from sequence specificity to epigenetics." *Nat Rev Genet* **2**(8): 640-5.

Mendez, J. and Stillman, B. (2003). "Perpetuating the double helix: molecular machines at eukaryotic DNA replication origins." *Bioessays* **25**(12): 1158-67.

Mendez, J., X. Zou-Yang, H., Kim, S.Y., Hikada, M., Tansey, W.P and Stillman, B. (2002). "Human origin recognition complex large subunit is degraded by ubiquitin-mediated proteolysis after initiation of DNA replication." *Mol Cell* **9**(3): 481-91.

Merchant, A. M., Kawasaki, Y., Chen, Y., Lei, M. and Tye, B.K. (1997). "A lesion in the DNA replication initiation factor Mcm10 induces pausing of elongation forks through chromosomal replication origins in Saccharomyces cerevisiae." *Mol Cell Biol* **17**(6): 3261-71.

Messer, W. (2002). "The bacterial replication initiator DnaA. DnaA and *oriC*, the bacterial mode to initiate DNA replication." *Microbiol Review* **26**: 355-74.

Mimura, S., Seki, T., Tanaka, S. and Diffley, J.F. (2004). "Phosphorylation-dependent binding of mitotic cyclins to Cdc6 contributes to DNA replication control." *Nature* **431**(7012): 1118-23.

Mizushima, T., Takahashi, N. and Stillman, B. (2000). "Cdc6p modulates the structure and DNA binding activity of the origin recognition complex in vitro." *Genes Dev* **14**(13): 1631-41.

Moldovan, G.L., Pfander, B. and Jentsch, S. (2007). "PCNA, the maestro of the replication fork." Cell 129(4): 665-79.

Morohashi, H., Maculins, T. and Labib, K. (2009). "The amino-terminal domain of Dia2 tethers SCF (Dia2) to the replisome progression complex." *Curr Biol* **19**(22): 1943-9.

Morrison, A. J., Kim, J. A., Person, M.D., Highland, J., Xiao, J., Wehr, T.S., Hensley, S., Bao, Y., Shen, J., Collins, S.R., Weissman, J.S., Delrow, J., Krogan, N.J., Haber, J.E. and Shen, X. (2007). "Mec1/Tel1 phosphorylation of the INO80 chromatin remodeling complex influences DNA damage checkpoint responses." *Cell* **130**(3): 499-511.

Moyer, S. E., Lewis, P. W. and Botchan, M.R. (2006). "Isolation of the Cdc45/Mcm2-7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase." *Proc Natl Acad Sci U S A* **103**(27): 10236-41.

Muller, P., Park, S., Shor, E., Huebert, D.J., Warren, C.L., Ansari, A.Z., Weinreich, M., Eaton, M.L., MacAlpine, D.M. and Foc, C.A. (2010). "The conserved bromo-adjacent homology domain of yeast Orc1 functions in the selection of DNA replication origins within chromatin." *Genes Dev* **24**(13): 1418-33.

Muramatsu, S., Hirai, K., Tak, Y.S., Kamimura, Y. and Araki, H. (2010). "CDK-dependent complex formation bteween replication proteins Dpb11, Sld2, Pol (epsilon), and GINS in budding yeast." *Genes Dev* **24**(6): 602-12.

Naylor, M. L., Li, J. M., Osborn, A.J. and Elledge, S.J. (2009). "Mrc1 phosphorylation in response to DNA replication stress is required for Mec1 accumulation at the stalled fork." *Proc Natl Acad Sci U S A* **106**(31): 12765-70.

Navadgi-Patil, V.M. and Burgers, P.M. (2008). "A tale of two tails: activation of DNA damage checkpoint kinase Mec1/ATR by the 9-1-1 clamp and by Dpb11/TopBP1." *DNA Rep* **8**(9):996-1003.

Nedelcheva-Veleva, M. N., Krastev, D. B. and Stoynov, S.S. (2006). "Coordination of DNA synthesis and replicative unwinding by the S-phase checkpoint pathways." *Nucleic Acids Res* **34**(15): 4138-46.

Nguyen, V. Q., Co, C., Irie, K. and Li, J.J. (2000). "Clb/Cdc28 kinases promote nuclear export of the replication initiator proteins Mcm2-7." *Curr Biol* **10**(4): 195-205.

Nguyen, V. Q., Co, C. and Li, J.J. (2001). "Cyclin-dependent kinases prevent DNA re-replication through multiple mechanisms." *Nature* **411**(6841): 1068-73.

Nick McElhinny, S. A., Gordenin, D. A., Stith, C.M., Burgers, P.M. and Kunkel, T.A. (2008). "Division of labor at the eukaryotic replication fork." *Mol Cell* **30**(2): 137-44.

Nieduszynski, C. A., Hiraga, S., Ak, P., Benham, C.J. and Donaldson, A.D. (2007). "OriDB: a DNA replication origin database." *Nucleic Acids Res* **35**(Database issue): D40-6.

Nieduszynski, C. A., Knox, Y. and DOnaldson, A.D. (2006). "Genome-wide identification of replication origins in yeast by comparative genomics." *Genes Dev* **20**(14): 1874-9.

Nishitani, H., Sugimoto, N., Roukos, V., Nakanishi, Y., Saijo, M., Obuse, C., Tsurimoto, T., Nakayama, K.I., Nakayama, K., Fujita, M., Lygerou, Z. and Nishimoto, T. (2006). "Two E3 ubiquitin ligases, SCF-Skp2 and DDB1-Cul4, target human Cdt1 for proteolysis." *EMBO J* **25**(5): 1126-36.

Nougarede, R., Della Seta, F., Zarvov, P. and Schwob, E. (2000). "Hierarchy of S-phase-promoting factors: yeast Dbf4-Cdc7 kinase requires prior S-phase cyclin-dependent kinase activation." *Mol Cell Biol* **20**(11): 3795-806.

Olsen, G.J. and Woese, C.R. (1997). "Archeal genomics: an overview." *Cell* **89**(7): 991-4.

Osborn, A. J. and Elledge, S. J. (2003). "Mrc1 is a replication fork component whose phosphorylation in response to DNA replication stress activates Rad53." *Genes Dev* **17**(14): 1755-67.

Paciotti, V., Clerici, M., Lucchini, G. and Longhese, M.P. (2000). "The checkpoint protein Ddc2, functionally related to S. pombe Rad26, interacts with Mec1 and is regulated by Mec1-dependent phosphorylation in budding yeast." *Genes Dev* **14**(16): 2046-59.

Pasero, P., Duncker, B. P., Schwob, E. and Gasser, S.M. (1999). "A role for the Cdc7 kinase regulatory subunit Dbf4p in the formation of initiation-competent origins of replication." *Genes Dev* **13**(16): 2159-76.

Patel, P. K., Arcangioli, B., Baker, S.P., Bensimon, A. and Rhind, N. (2006). "DNA replication origins fire stochastically in fission yeast." *Mol Biol Cell* **17**(1): 308-16.

Patel, P. K., Kommajosyula, N., Rosebrock, A., Bensimon, A., Leatherwood, J., Bechloefer, J. and Rhind, N. (2008). "The Hsk1(Cdc7) replication kinase regulates origin efficiency." *Mol Biol Cell* **19**(12): 5550-8.

Perkins, G., Drury, L. S. and Diffley, J.F. (2001). "Separate SCF(CDC4) recognition elements target Cdc6 for proteolysis in S phase and mitosis." *EMBO J* **20**(17): 4836-45.

Porter, A. C. (2008). "Preventing DNA over-replication: a Cdk perspective." *Cell Div* **3**: 3.

Pryde, F., Jain, D., Kerr, A., Curley, R., Mariotti, F.R. and Vogelauer, M. (2009). "H3 k36 methylation helps determine the timing of cdc45 association with replication origins." *PLoS One* **4**(6): e5882.

Pursell, Z. F., Isoz, I., Lundstrom, E.B., Johansson, E. and Kunkel, T.A. (2007). "Regulation of B family DNA polymerase fidelity by a conserved active site residue: characterization of M644W, M644L and M644F mutants of yeast DNA polymerase epsilon." *Nucleic Acids Res* **35**(9): 3076-86.

Raghuraman, M. K. and B. J. Brewer (2010). "Molecular analysis of the replication program in unicellular model organisms." *Chromosome Res* **18**(1): 19-34.

Raghuraman, M. K., Brewer, B. J. and Fangman, W.L. (1997). "Cell cycle-dependent establishment of a late replication program." *Science* **276**(5313): 806-9.

Raghuraman, M. K., Winzeler, E. A., Collingwood, D., Hunt, S., Wodicka, L., Conway, A., Lockhart, D.J., Davis, R.W., Brewer, B.J. and Fangman, W.L. (2001). "Replication dynamics of the yeast genome." *Science* **294**(5540): 115-21.

Randell, J. C., Bowers, J. L., Rodriguez, H.K. and Bell, S.P. (2006). "Sequential ATP hydrolysis by Cdc6 and ORC directs loading of the Mcm2-7 helicase." *Mol Cell* **21**(1): 29-39.

Rao, H. and Stillman, B. (1995). "The origin recognition complex interacts with a bioartite DNA binding site within yeast replicators." *Proc Natl Acad Sci USA* **92**(6): 2224-8.

Rao, P. N. and Johnson, R.T. (1970). "Mammalian cells fusion: Studies on the regulation of DNA synthesis and mitosis." *Nature* **225**: 159-164.

Rappsilber, J., Ishiyama, Y. and Mann, M. (2003). "Stop and go extraction tips for matrix-assosted desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics." *Anal Chem* **75**(3): 663-70.

Rappsilber, J., Ryder, U., Lamond, A.I. and Mann, M. (2002). "Large-scale proteomic analysis of the human spliceosome." *Genome Res* **12**(8): 1231-45.

Raveendranathan, M., Chattopadhyay, S., Bolon, Y.T., Haworth, J., Clarke, D.J. and Bielinsky, A.K. (2006). "Genome-wide replication profiles of S-phase checkpoint mutants reveal fragile sites in yeast." *EMBO J* **25**(15): 3627-39.

Rhind, N. (2006). "DNA replication timing: random thoughts about origin firing." *Nat Cell Biol* **8**(12): 1313-6.

Ricke, R. M. and Bielinsky, A. K. (2004). "Mcm10 regulates the stability and chromatin association of DNA polymerase-alpha." *Mol Cell* **16**(2): 173-85.

Robinson, N. P. and Bell, S. D. (2005). "Origins of DNA replication in the three domains of life." *FEBS J* **272**(15): 3757-66.

Robinson, N.P., Dionne, I., Lundgren, M., Marsh, V.L., Bernarder, R. and Bell, S.D. (2004). "Identification of two oigins of replication in the single chromosome of the archaeon Sulfolobus Sulfactarius." *Cell* **116**(1): 25-38.

Rossi, M.L. and Bambara, R.A. (2006). "Reconstituted Okazaki fragment processing indicates two pathways of primer removal." *J Biol Chem* **281**(36): 26051-61.

Sakwe, A.M., Nguyen, T., Athanasopoulos, V., Shire, K. and Frappier, L. (2007). "Identification and characterization of a novel component of the human minichromosome maintenance complex." *Mol Cell Biol* **27**(8): 3044-55.

Sambrook, J. and Russell, D. (2002). "Molecular cloning: A laboratory manual." *CSH*.

Santocanale, C. and Diffley, J. F. (1996). "ORC- and Cdc6-dependent complexes at active and inactive chromosomal replication origins in Saccharomyces cerevisiae." *EMBO J* **15**(23): 6671-9.

Santocanale, C. and Diffley, J. F. (1998). "A Mec1- and Rad53-dependent checkpoint controls late-firing origins of DNA replication." *Nature* **395**(6702): 615-8.

Santocanale, C., Sharma, K. and Diffley, J.F. (1999). "Activation of dormant origins of DNA replication in budding yeast." Genes Dev **13**(18): 2360-4.

Sawyer, S. L., Cheng, I. H., Chai, W. and Tye, B.K. (2004). "Mcm10 and Cdc45 cooperate in origin activation in Saccharomyces cerevisiae." *J Mol Biol* **340**(2): 195-202.

Schaper, S. and Messer, W. (1995). "Interaction of the initiator protein DnaA of Escherichia coli with its DNA target." *J. Biol. Chem* **270**(29): 17622-6.

Schwacha, A. and Bell, S.P. (2001). "Interactions between catalytically distinct MCM sbgroups are essential for coordinated ATP hydrolysis and DNA replication." *Mol Cell* **8**(5): 1093-104.

Sclafani, R. A. (2000). "Cdc7p-Dbf4p becomes famous in the cell cycle." *J Cell Sci* **113 (Pt 12)**: 2111-7.

Sclafani, R. A. and Holzen, T. M. (2007). "Cell cycle regulation of DNA replication." *Annu Rev Genet* **41**: 237-80.

Sclafani, R. A., Tecklenburg, M. and Pierce, A. (2002). "The mcm5-bob1 bypass of Cdc7p/Dbf4p in DNA replication depends on both Cdk1-independent and Cdk1-dependent steps in Saccharomyces cerevisiae." *Genetics* **161**(1): 47-57.

Segurado, M., de Luis, A. and Antequera, F. (2003). "Genome-wide distribution of DNA replication origins at A+T-rich islands in Schizosaccharomyces pombe." *EMBO Rep* **4**(11): 1048-53.

Segurado, M. and Diffley, J. F. (2008). "Separate roles for the DNA damage checkpoint protein kinases in stabilizing DNA replication forks." *Genes Dev* **22**(13): 1816-27.

Segurado, M. and Tercero, J. A. (2009). "The S-phase checkpoint: targeting the replication fork." *Biol Cell* **101**(11): 617-27.

Seki, T., Akita, M., Kamimura, Y., Muramatsu, S., Araki, H. and Sugino, A. (2006). "GINS is a DNA polymerase epsilon accessory factor during chromosomal DNA replication in budding yeast." *J Biol Chem* **281**(30): 21422-32.

Sheu, Y. J. and Stillman, B. (2010) "The Dbf4-Cdc7 kinase promotes S phase by alleviating an inhibitory activity in Mcm4." *Nature* **463**(7277): 113-7.

Sheu, Y. J. and Stillman, B. (2006). "Cdc7-Dbf4 phosphorylates MCM proteins via a docking site-mediated mechanism to promote S phase progression." *Mol Cell* **24**(1): 101-13.

Shimada, K., Oma, Y., Schleker, T., Kugou, K., Ohta, K., Harata, M. and Gasser, S. (2008). "Ino80 chromatin remodeling complex promotes recovery of stalled replication forks." *Curr Biol* **18**(8): 566-75.

Shirahige, K., Iwasaki, T., Rashid, M.B., Ogasawara, N. and Yoshikawa, H. (1993). "Location and characterization of autonomously replicating sequences from chromosome VI of Saccharomyces cerevisiae." *Mol Cell Biol* **13**(8): 5043-56.

Shirayama, M., Toth, A., Galova', M. and Nasmyth, K. (1999). "APC(Cdc20) promotes exit from mitosis by destroyin the anaphase inhibitor Pds1 and cyclin Clb5." *Nature* **402**(6758): 203-7.

Speck, C., Chen, Z., Li, H. and Stillman, B. (2005). "ATPase-dependent cooperative binding of ORC and Cdc6 to origin DNA." *Nature* **12**(11): 7.

Speck, C. and Stillman, B. (2007). "Cdc6 ATPase regulates ORC x Cdc6 stability and the selection of specific DNA sequences as origins of DNA replication." *J Biol Chem* **282**(16): 11705-14.

Spiesser, T. W., Klipp, E. and Barberis, M. (2009). "A model for the spatiotemporal organization of DNA replication in Saccharomyces cerevisiae." *Mol Genet Genomics* **282**(1): 25-35.

Stillman, B. (2005). "Origin recognition and the chromosome cycle." *FEBS Letters* **579**: 877-884.

Stillman, B. (2008). "DNA polymerases at the rpelication fork in eukaryotes." Mol *Cell* **30**(3): 259-60.

Stillman, B., Bell, S. P., Dutta, A. and Marahrens, Y. (1992). "DNA replication and the cell cycle." *Ciba Found Symp* **170**: 147-56; discussion 156-60.

Stinchcomb, D. T., Thomas, M., Kelly, J., Selker, E. and Davis, R.W. (1980). "Eukaryotic DNA segments capable of autonomous replication in yeast." *Proc Natl Acad Sci U S A* **77**(8): 4559-63.

Tada, S., Kundu, L.R. and Enomoto, T. (2008). "Insight into initiator-DNA interactions: a lesson from the archeal ORC." *BioEssay* **30**: 208-211.

Tak, Y. S., Tanaka, Y., Endo, S., Kamimura, Y. and Araki, H. (2006). "A CDK-catalysed regulatory phosphorylation for formation of the DNA replication complex Sld2-Dpb11." *EMBO J* **25**(9): 1987-96.

Takayama, Y., Kamimura, Y., Okawa, M., Muramatsu, S., Sugino, A. and Araki, H (2003). "GINS, a novel multiprotein complex required for chromosomal DNA replication in budding yeast." *Genes Dev* **17**(9): 1153-65.

Takeda, D.Y. and Dutta, A. (2005). "DNA replication and progression through S-phase." *Oncogene* **24**(17): 2827-43.

Takeda, D.Y., Shibata, Y., Parvin, J.D. and Dutta, A. (2005). "Recruitment of ORC and Cdc6 to DNA is sufficient to create an artificial origin of replication in mammalian cells." *Genes & Dev* **19**: 2827-36.

Tanaka, H., Katou, Y., Yagura, M., Saitoh, K., Itoh, T., Araki, H., Bando, S. and Shirahige, K. (2009). "Ctf4 coordinates the progression of helicase and DNA polymerase α" *Genes to Cells* **14**: 807-20.

Tanaka, S. and Diffley, J. F. (2002). "Interdependent nuclear accumulation of budding yeast Cdt1 and Mcm2-7 during G1 phase." *Nat Cell Biol* **4**(3): 198-207.

Tanaka, S., Umemori, T., Hirai, K., Muramatsu, S., Kamimura, Y. and Araki, H. (2007). "CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast." *Nature* **445**(7125): 328-32.

Tercero, J. A. and Diffley, J. F. (2001). "Regulation of DNA replication fork progression through damaged DNA by the Mec1/Rad53 checkpoint." *Nature* **412**(6846): 553-7.

Tercero, J.A., Labib, K. and Diffley, J.F. (2000). "DNA synthesis at individual forks requires the essential initiation factor Cdc45p." *EMBO J* **19**(9): 2082-93.

Tercero, J. A., Longhese, M. P. and Diffley, J.F. (2003). "A central role for DNA replication forks in checkpoint activation and response." *Mol Cell* **11**(5): 1323-36.

Tourriere, H. and Pasero, P. (2007). "Maintenance of fork integrity at damaged DNA and natural pause sites." *DNA Repair (Amst)* **6**(7): 900-13.

Tourriere, H., Versini, G., Cordon-Preciado, V., Alabert, C. and Pasero, P. (2005). "Mrc1 and Tof1 promote replication fork progression and recovery independently of Rad53." *Mol Cell* **19**(5): 699-706.

Tuduri, S., Tourriere, H. and Pasero, P. (2010). "Defining replication origin efficiency using DNA fiber assays." *Chromosome Res* **18**(1): 91-102.

Ulhmann, F. (2009). "A matter of choice: the establishment of sister chromatid cohesion." *EMBO Rep* **10**(10): 1095-102.

Unnikrishnan, A., Gafken, P. R. and Tsukiyama, T. (2010) "Dynamic changes in histone acetylation regulate origins of DNA replication." *Nat Struct Mol Biol* **17**(4): 430-7.

Usui, T., Ogawa, H. and Petrini, J.H. (2001). "A DNA damage response pathway controlled by Tel1 and Mre11 complex." *Mol Cell* **7**(6): 1255-66.

van den Bosch, M. and Lowndes, N.F. (2004). "Remodeling the Rad9 checkpoint complex." *Cell Cycle* **3**(2): 119-122.

Venditti, P., Costanzo, G., Negri, R. and Camilloni, G. (1994). "ABFI contributes to the chromatin organization of Saccharomyces cerevisiae ARS1 B-domain." *Biochim Biophys Acta* **1219**(3): 677-89.

Vincent, J. A., Kwong, T. J. and Tsukiyama, T. (2008). "ATP-dependent chromatin remodeling shapes the DNA replication landscape." *Nat Struct Mol Biol* **15**(5): 477-84.

Vogelauer, M., Rubbi, L., Lucas, I., Brewer, B.J. and Grunstein, M. (2002). "Histone acetylation regulates the time of replication origin firing." *Mol Cell* **10**(5): 1223-33.

Walter, J. and Newport, J. (2000). "Initiation of eukaryotic DNA replication: origin unwinding and sequential chromatin association of cdc45, RPA and DNA polymerase α." *Mol Cell* **5**: 617-27.

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

Wang, L., Lin, C.M., Brooks, S., Cimbora, D., Groudine, M. and Aladjem, M.I. (2004). "The human beta-globin replication initiation consists of two modular independent replicators." *Mol Cell Biol* **24**(8): 3373-86.

Weinreich, M., Palacios DeBeer, M. A. and Fox, C.A. (2004). "The activities of eukaryotic replication origins in chromatin." *Biochim Biophys Acta* **1677**(1-3): 142-57.

White, E. J., Emanuelsson, O., Scalzo, D., Royce, T., Kosak, S., Oakeley, E.J., Weissman, S., Gerstein, M., Groudine, M., Snyder, M., Schubeler, D. (2004). "DNA replication-timing analysis of human chromosome 22 at high resolution and different developmental states." *Proc Natl Acad Sci U S A* **101**(51): 17771-6.

Willis, N. and Rhind, N. (2009). "Regulation of DNA replication by the S-phase DNA damage checkpoint." *Cell Div* **4**: 13.

Wilmes, G. M. and Bell, S. P. (2002). "The B2 element of the Saccharomyces cerevisiae ARS1 origin of replication requires specific sequences to facilitate pre-RC formation." *Proc Natl Acad Sci U S A* **99**(1): 101-6.

Wohlschlegel, J.a., Dwyer, B.T., Dhar, S.K., Cvetic, C., Walter, J.C. and Dutta, A. (2000). "Inhibition of eukaryotic DNA replication by geminin binding to Cdt1" *Science* **290**(5500): 2309-12.

Woodfine, K., Fiegler, H., Beare, D.M., Collins, J.E., McCann, O.T., Young, B.D., Debernardi, S., Mott, R., Dunham, I. and Carter, N.P. (2004). "Replication timing of the human genome." *Hum Mol Genet* **13**(2): 191-202.

Wu, P. Y. and Nurse, P. (2009). "Establishing the program of origin firing during S phase in fission Yeast." *Cell* **136**(5): 852-64.

Wyrick, J. J., Aparicio, J. G., Chen, T., Barnett, J.D., Jennings, E.G., Young, R.A., Bell, S.P. and Aparicio, O.M. (2001). "Genome-wide distribution of ORC and MCM proteins in S. cerevisiae: high-resolution mapping of replication origins." *Science* **294**(5550): 2357-60.

Xu, W., Aparicio, J.G., Aparicio, O.M. and Tavare', S. (2006). "Genome-wide mapping of ORC and Mcm2p binding sites on tilling arrays and identification of essential ARS consensus sequences in S. cerevisiae." *BMC Genomics* **7**:276.

Yabuki, N., Terashima, H. and Kitada, K. (2002). "Mapping of early firing origins on a replication profile of budding yeast." *Genes Cells* **7**(8): 781-9.

Yabuuchi, H., Yamada, Y., Uchida, T., Sunathavinchkul, T., Nakagawa, T. and Masukata, H. (2006). "Ordered assembly of Sld3, GINS and Cdc45 is distinctly regulated by DDK and CDK for activation of replication origins." *EMBO J* **25**(19): 4663-74.

Yamashita, M., Hori, Y., Shinomiya, T., Obuse, C., Tsurimoto, T., Yoshikawa, H. and Shirahige, K. (1997). "The efficiency and timing of initiation of replication of multiple replicons of Saccharomyces cerevisiae chromosome VI." *Genes Cells* **2**(11): 655-65.

You, Z. and Masai, H. (2008). "Cdt1 forms a complex with the minichromosome maintenance protein (MCM) and activates its helicase activity." *J Biol Chem* **283**(36): 24469-77.

Zegerman, P. and Diffley, J. F. (2007). "Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast." *Nature* **445**(7125): 281-5.

Zegerman, P. and Diffley, J.F. (2010). "Checkpoint-dependent inhibition of DNA replication initiation by SLd3 and Dbf4 phosphorylation." *Nature* **467**(7314): 474-8.

Zhang, Y., yu, Z., Fu, X. and Liang, C. (2002). "Noc3p, a bHLH protein, plays an integral role in the initiation of DNA replication in budding yeast." *Cell* **109**(7): 849-60.

Zou, L. and Elledge, S. J. (2003). "Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes." *Science* **300**(5625): 1542-8.

Zou, L., Mitchell, J. and 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**(2): 553-63.

Zou, L. and Stillman, B. (1998). "Formation of a preinitiation complex by S-phase cyclin CDK-dependent loading of Cdc45p onto chromatin." *Science* **280**(5363): 593-6.

Zou, L. and Stillman, B. (2000). "Assembly of a complex containing Cdc45p, replication protein A, and Mcm2p at replication origins controlled by S-phase cyclin-dependent kinases and Cdc7p-Dbf4p kinase." *Mol Cell Biol* **20**(9): 3086-96.