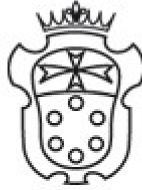
SCUOLA NORMALE SUPERIORE DI PISA



PH. D. THESIS

Molecular Biology

DNA-protein interaction dynamics at the Lamin B2 replication origin

Luca Puzzi

ADVISOR

Prof. Mauro Giacca

2014

This work is dedicated to the memory

of Professor Arturo Falaschi

and to all the people who always trusted in me

TABLE OF CONTENTS

List of Figures and Tables..........vi

List of Abbreviations..........viii

List of Publications..........x

ABSTRACT..........xii

1. INTRODUCTION

1.1 Regulatory mechanisms of eukaryotic DNA replication..........2

1.1.1 Eukaryotic DNA replication: an overview..........2

1.1.2 Multiple levels of DNA replication regulation within cell-cycle progression..........5

1.1.3 From initiation to elongation in eukaryotic DNA replication..........8

1.1.4 Additional requirement for DNA replication..........19

1.1.5 The still obscure determinants of origin specification..........23

1.2 The AP-1 protein family..........26

1.2.1 AP-1 complex protein composition and general features..........26

1.2.2 Structure of AP-1 proteins..........28

1.2.3 AP-1 expression, regulation and activity..........32

2. AIM OF THE WORK

2.1 AP-1 and DNA replication..........39

3. RESULTS

3.1 Characterization of T98G-HF stable cell clone expressing HA/Flag HOXC13..........41

3.2 HOXC13-HF Cp-immunoprecipitation analysis..........46

3.3 Analysis of AP-1 binding onto the LaminB2 origin..........50

3.4 High resolution analysis of the LaminB2 origin of DNA replication..........51

3.4.1 Validation of the high resolution ChIP technology..........52

3.4.2 Analysis of the LaminB2 origin of DNA replication by high resolution ChIP..........53

3.5 Spatial and temporal analysis of AP-1 proteins binding onto the LaminB2 origin..........61

3.6 Effects of the disruption of topological structure on protein binding across the LaminB2 origin region..........69

4. DISCUSSION

4.1 c-Fos and c-Jun are involved in the process of DNA replication..........74

4.2 c-Fos and c-Jun interact with the replicative machinery during late G1............................75

4.3 c-Fos and c-Jun interaction with LaminB2 origin relies on chromatin topology...............75

4.4 A speculative model for origin activation...77

5. MATERIAL AND METHODS

5.1 Cell culture, synchronization and Merbarone treatment...79

5.1.1 Cell synchronization and FACS analysis...80

5.1.2 Merbarone treatment..80

5.2 Antibodies..80

5.3 *in vivo* topoisomerase I and II mapping..81

5.4 Chromatin immuno precipitation analyses (ChIP)..82

5.4.1 X-ChIP..82

5.4.2 Native chromatin immuno precipitation analyses (NChIP)..83

5.4.3 Sequential chromatin immunoprecipitation analyses (Re-ChIP)....................................83

5.5 PCR analyses..84

5.5.1 Competitive PCR..84

5.5.2 High resolution LaminB2 PCR..85

REFERENCES

AKNOWLEDGMENTS

List of Figures and Tables

Figure 1.1: Regulation of DNA replication by origin usage..3

Figure 1.2: Different types of DNA replication origins …..7

Figure 1.3: Model of the regulation of DNA replication..10

Figure 1.4: Schematic pictures of ORC1-5 and Cdc6 proteins from *S. cerevisiae*.................12

Figure 1.5 The many functions of the origin recognition complex..14

Figure 1.6: Proposed features to determine the selection and activation of replication origin..19

Figure 1.7: Model for the progressive restriction of initiation potential during G1 phase...20

Figure 1.8: A/T-rich and C/G-rich islands at DNA replication origins..................................24

Figure 1.9: Coiled coil protein structure..29

Figure 1.10: The Jun-Fos heterodimer...30

Figure 1.11: Transcriptional and post-translational activation of AP-1.................................33

Table 1.1: Different AP-1 genes regulation in tumor development and suppression..............36

Figure 3.1: Flow cytometry and nascent DNA analyses on the T98G-HA/Flag stable clone...41

Figure 3.2: X-linked HA ChIP analysis on T98G HA/Flag stable clone...............................43

Figure 3.3: *in vivo* cleavage analyses of Topoisomerase I and II...45

Figure 3.4: Co-IP partners of HOXC13-HA after X-linked ChIP and Native ChIP...............47

Figure 3.5: ChIP analysis for AP-1 proteins on the LaminB2 origin.....................................48

Figure 3.6: Schematic representation of the position of the 12 overlapping PCR segments...50

Figure 3.7: Primer analysis after amplification of a DNA amplicon obtained using the 1L-12U primers or total genomic DNA from starved (G0) T98G cells...52

Figure 3.8: Primer efficiency analyses performed using G0 DNA..53

Figure 3.9: Input chromatin profile in the presence and absence of formaldehyde crosslink...54

Figure 3.10: H2B and acetylated H3(K14) NChIP analysis....55

Figure 3.11: Pre-RC proteins binding profiles across LaminB2 origin....56

Figure 3.12: c-Fos and c-Jun ChIP analysis....57

Figure 3.13: ChIP analysis to detect binding, across the LaminB2 origin, of other potential protein candidates....59

Figure 3.14: Flow cytometry profile and expression levels of AP-1 proteins during the cell cycle....60

Figure 3.15: Western Blot analyses of AP-1 proteins in different phases of the cell cycle....62

Figure 3.16: ChIP analysis for AP-1 and ORC4 proteins through the cell cycle....63

Figure 3.17: Co-IP of ORC4 in AP-1-ChIP samples across the cell cycle....64

Figure 3.18: Development of the Re-ChIP procedure....66

Figure 3.19: AP-1/ORC4 chromatin binding profiles by re-ChIP analyses....67

Figure 3.20: Effect of the topoisomerase II inhibitor merbarone on DNA synthesis....68

Figure 3.21: Effects of merbarone on protein:DNA interactions across the LaminB2 origin region....70

Figure 4.1: Alignment of the “non-canonical sequences” to the 1.1 Kbps region analyzed....76

Figure 4.2: Proposed model....77

Table 5.1: Primers used for high resolution protein binding analysis at the LaminB2 replication origin....86

List of Abbreviations

AP-1 = Activator protein 1

ACS = ARS consensus sequence

APC = Anaphase promoting complex

ARS = Autonomously replicating sequence

ATP = Adenosine triphosphate

BAH = Bromo adjacent homology domain

BrdU = Bromo deoxy-Uridine

bps = Base pairs

CDK =Cyclin-dependent-kinase

ChIP = Chromatin immunoprecipitation

Co-IP = Co-immunoprecipitation

CRE = Cyclic AMP responsive element

Ct = Cycle threshold

DDK = Dbf4-dependent-kinase

DTT = Dithiothreitol

EDTA = Ethylenediamine tetra-acetic acid

HA = Hemaagglutinin protein

MCM = Mini-Chromosome-Maintenance proteins

ORC = Origin Recognition Complex

PBS = Phosphate buffered saline

PCR = Polymerase chain reaction

PI = Propidium iodide

pre-IC = pre-initiative complex

pre-RC = pre-replicative complex

RC = replicative complex

SSB = single-stranded binding protein

ssDNA = single-stranded DNA

TAD = transactivation domain

TPA = 12-O-tetradecanoylphorbol 13-acetate

TRE = TPA responsive element

WB = Western Blot

WCL = whole cell lysate

WCE = whole cell extract

List of Publications

Luca Puzzi, Laura Marchetti, Fiorenzo A. Peverali, Giuseppe Biamonti, Mauro Giacca, "DNA-protein interaction dynamics at the Lamin B2 replication origin", Cell Cycle, 2015, 14 (1): 64-73

Publications not included in this thesis

L. Marchetti, L. Comelli, B. D'Innocenzo, **L. Puzzi**, S. Luin, D. Arosio, M. Calvello, R. Mendoza-Maldonado, F. Peverali, F. Trovato, S. Riva, G. Biamonti, G. Abdurashidova, F. Beltram, A. Falaschi, "Homeotic proteins participate in the function of human DNA replication origins", Nucleic Acids Research, 2010 (22): 8105-19

ABSTARCT

The regulation of human DNA replication operates via a time-defined program of activation and deactivation of approximately 30,000 replication origins distributed along the genome. Due to the complexity of this process, each step requires a sequence of cascade checkpoints and licensing events, most of which are well conserved from yeasts to humans. A multi-protein complex assembles onto each origin causing the local unwinding of the DNA double helix and the start of two oppositely moving replicative forks. Despite the *cis*-acting elements necessary for origin firing are almost elucidated, the mechanism that governs the selection of a specific DNA sequence as human (and, more generally, metazoan) origin, in the course of G1 phase of the cell cycle, is still poorly understood. The lack of DNA-sequence consensus between replication origins characterized so far, together with the poor binding-specificity displayed by the Origin Recognition Complex, suggest that origin selection might rather be determined by local chromatin structures and/or *trans*-acting factors. With regard to the latter possibility, it was interesting to find out that a DNA region specifically bound by the AP-1 proteins, is located close to the start site of the human Lamin B2 replication origin.

In the study conducted during this Ph.D. program, the possible role of AP-1 transcription factors in origin specification was explored by investigating the involvement the principal moieties of this protein family, c-Fos and c-Jun, within the replicative complexes in living human cells. The data reported in this thesis provides evidence that both c-Fos and c-Jun interact with the LaminB2 origin of DNA replication and indeed participates in origin function. Participation of these proteins to origin binding is consistent with their interaction with both ORC4 and HOXC13, two members of the replicative complex, and is cell cycle defined, occurring before origin firing. Furthermore the observations point to the existence of specific and

dynamic structural reorganizations of the complexes assembled at the origin region along with origin activation. In this view, AP-1 proteins could contribute to recruit and stabilize the replicative complexes onto the LaminB2 origin, in presence of specific chromatin and topological configurations.

Chapter 1

Introduction

The work reported in this dissertation explores the possible connection between two traditionally separated fields of biology, the regulation of DNA replication and the function of activator protein 1 (AP-1) transcription factors. To provide the conceptual frame of my experimental work, the two following paragraphs will focus on a description of the mechanisms of DNA replication (paragraph 1.1) and of the structure and function of AP-1 proteins (paragraph 1.2). I will try to summarize what appears to be missing for a satisfactory understanding of DNA replication regulation in metazoan organisms, and to what extent the AP-1 proteins could be involved in this process.

1.1 Regulatory mechanisms of eukaryotic DNA replication

1.1.1 Eukaryotic DNA replication: an overview

DNA is the most important molecule for all living organisms, which has to be maintained intact to allow the survival of the organisms themselves. The importance of this process is demonstrated by the presence of many steps that are finely regulated, involving different proteins both for the process itself, and as controllers. One of the most valuable contributions to our understanding of DNA replication is the replicon model proposed in 1963 by Jacob and Brenner [1], who postulated the existence of two fundamental elements that regulate DNA replication: a cis-acting sequence within the genome, called the "replicator" from which replication starts and a positive trans-acting factor called the ''initiator'' able to recognize specifically the sequence of the replicator within the genome. In response to appropriate cellular signals, the initiator directs the local unwinding of the replicator sequence and recruits additional factors to initiate the process of DNA replication. Once DNA replication starts, the replication fork proceeds until genome duplication is completed.

The replicon theory was initially verified by using a bacterial chromosome (Figure 1a). In *Escherichia coli*, the initiator protein DnaA binds with high affinity to the replication origin oriC, which contains multiple DnaA-binding sites [2]. Eukaryotic genomes are very large and their replication rate is slow, when compared to the prokaryotic replicon model. Nevertheless, the process of DNA replication is made possible by the start of DNA replication at 30.000–50.000 different chromosomal locations, known as origins of DNA replication, that are specifically selected and activated in each cell cycle [3] (Figure 1b). The process of DNA synthesis relies on a spatio-temporal coordinated cycle of activation and deactivation of the origins,

restricted to a relatively narrow window of the eukaryotic cell cycle, namely the S phase. The two main advantages of this mechanism are that the overall time required to duplicate the entire genome is reduced and that the generation of single-stranded DNA (ssDNA) is much more localized and transient, helping preserving genome integrity [4,5]. Actually, the activation of all origins localized in eukaryotic genomes leads to the formation of tandemly arranged replication units, each of which can conceptually be considered as an analog of the bacterial replicon [6]. Studies performed in different organisms have clearly demonstrated that more origins are prepared for replication in G1 than those that are actually used during the S phase. This phenomenon, which is known as origin redundancy, is likely to represent a foolproof mechanism, ensuring that replication restarts through the activation of "dormant origins" when replication forks are arrested [7] (Figure 1.2).

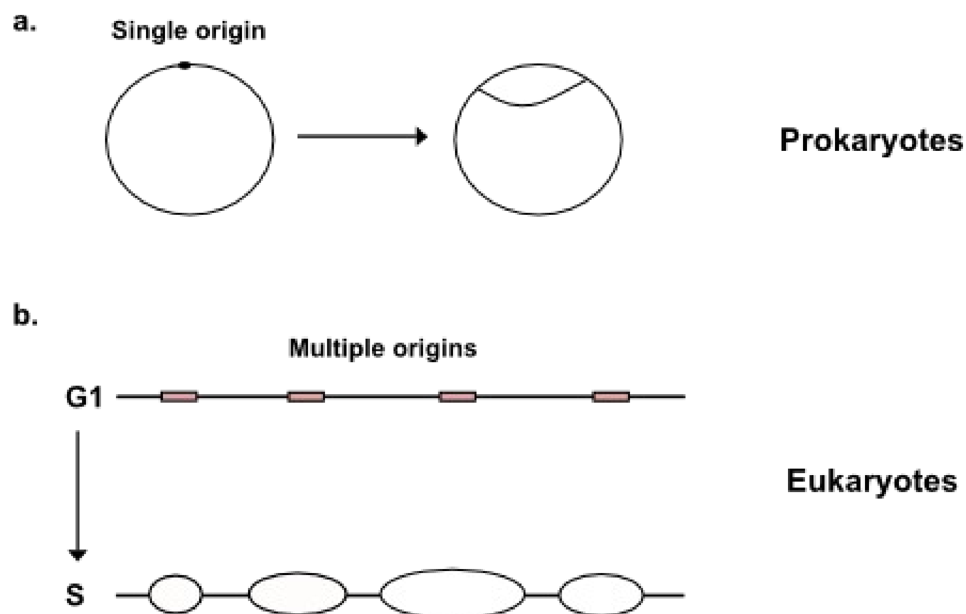


Figure 1.1: Regulation of DNA replication by origin usage. While prokaryotes have a single origin on a circular chromosome (a) in eukaryotes instead, multiple origins are found on a single, linear chromosome (b). This is useful to achieve a faster replication.

Considering DNA replication as a process, it can be divided into three steps: initiation, elongation and termination. During initiation [8], a specific DNA sequence is selected to be an origin (i.e. the start site) of DNA replication, usually in correspondence to loci of actively transcribed genes and AT-rich sequences, and initiator proteins assemble thereon. This results in the formation of a multi-protein complex which is responsible for the local melting of the DNA duplex, which is necessary for proteins to have access to the template strands. Subsequently, the complex stabilizes the ssDNA that is formed, and two replication forks, thanks to DNA helicases and polymerases, start to replicate the two parental DNA strands in opposite directions. The elongation step [9] is actually the continuation of the unwinding activity by the two fork complexes. It ensures simultaneous replication of both parental DNA strands also outside of the origin sequence. When two replication forks converge, they merge and termination of replicon duplication occurs [10]. Significant differences exist between DNA replication mechanisms in lower and higher eukaryotes. In the former organisms, such as the budding yeast *Saccharomices cerevisiae*, replication origins are well defined genetic elements containing the conserved and essential autonomously replicating sequence (ARS) consensus sequence (ACS) directly bound by the origin recognition complex (ORC) proteins in an ATP- dependent manner [8] which serves as a platform for the assembly of the pre-replicative complex (Pre-RC). On the contrary, in the fission yeast *Schizosaccharomyces pombe,* origins are much larger and the ORC complex does not interact with any conserved consensus sequence but binds to AT-rich origin sequences thanks to the AT-hook DNA binding domain of the ORC4 subunit [11], showing an evolution in the origin recognition process in eukaryotes. Conversely, higher eukaryotic organisms display a number of origins. These are at least 100-fold more abundant and, at present, no sequence specific replicators have been found [12]. In spite of these disparities between lower and higher eukaryotes, the proteins that regulate replication are highly conserved in function from yeast to Drosophila, from Xenopus to man, suggesting a common mechanism in the replication function that does not depend on the origin sequence itself [13].

1.1.2 Multiple levels of DNA replication regulation during cell cycle progression

The complexity of DNA replication in higher eukaryotes implies that any deregulating factor could lead cells to enter apoptosis or progress to tumorigenesis. Because such a strict relationship between replication and tumor proliferation exists, this process has to be strictly controlled by many levels of regulation. The first regulatory step of DNA replication concerns the activation, or initiation of the replication origins. This starts from the end of M phase and the onset of G1 phase [14], when several proteins take part in the pre-RC complex assembly by selecting the DNA sequences that are going to become replication origins and, by binding to these regions, promote the recruitment of other proteins involved in origin activation. During the G1 phase, for all the sequences selected as origins, a timing time of replication initiation is assigned and only a subset of the origins will fires immediately after entry into the S phase (early origins). The remaining ones (middle and late origins) are programmed to fire in an ordered manner after early origins. This results in an organized spatio-temporal activation of replication clusters of different subchromosomal domains at different times during the S phase [15]. In response to genotoxic damage, the DNA damage response pathway prevents entry into S phase by the activation of the G1/S border checkpoint, the components of which are highly conserved in eukaryotes. DNA damage is detected by the ataxia telangiectasia mutated related (ATR) protein which acts as a sensor and leads to the activation of CHK1 protein kinase, which in turn activates effectors (p53 and Cdc25A) that interact with the cell cycle machinery to inhibit cell cycle progression by controlling the association of Cdc45 with chromatin, preventing the transition between G1 to S phase [16, 17].

After origin firing, at the beginning of S phase, the pre-RC is re-organized due to the degradation or modification of several of its members, as a regulatory mechanism to avoid re-replication [18, 19]. Moreover, the temporal separation of pre-RC assembly

from origin activation is actually another key event ensuring that new pre-RC cannot assemble on origins that have already fired [18]. These mechanisms rely on the activity of CDKs (cyclin-dependent kinases), cell-cycle regulated kinases which act on several target proteins [8] controlling the time of replication initiation at specific origins [15, 20]. Because the activity of these kinases remains high from the S phase onset to the end of the following mitosis, re- licensing cannot occur until the beginning of the next subsequent cell cycle [21].

The mechanism that regulates the timing of replication is not completely understood. Originally, it was thought that early replication is a prelude to transcription because transcriptionally active euchromatic regions replicate early and inactive heterochromatic regions late. The molecular relationship between transcription and replication in regulating these temporal programs is unclear, and certainly goes beyond the actual need of DNA-binding proteins to access regions in which chromatin is unfolded [22, 23]. Studies in metazoa have indeed confirmed the recurring correspondence between initiation of DNA replication and transcriptionally active regions [24, 25]. Nevertheless, the timing of origin activation has been reported to correlate with a developmental program rather than with transcription *per se* [26, 27].

Given the complexity and importance of DNA replication elongation for the maintenance of genome integrity, many different checkpoint pathways are active within the S phase, as demonstrated by studies in the yeast model [28]. These checkpoints encompass the whole phase of DNA synthesis, as well as the switch to G2 phase, and comprise a variety of mechanisms to prevent replication defects, repair damaged replication forks and enable fork reactivation. For their role in the overall control of the cell cycle progression, as well as the control of genomic stability, they are often referred to as cell cycle checkpoints [29]. Very interestingly, the induction of a cell cycle checkpoint often results in the retroactive regulation of the recruitment of key members of the pre-RC to the origin site. For example, in the budding yeast, hydroxyurea treatment not only blocks fork progression from early origins but also prevents the firing of late origins, and this mechanism was shown to depend on

Rad53 and Mec1 [30]. The same conclusion was also obtained following induction of double strand breaks, and the protein involved in this regulation was shown to be ORC2 [31].

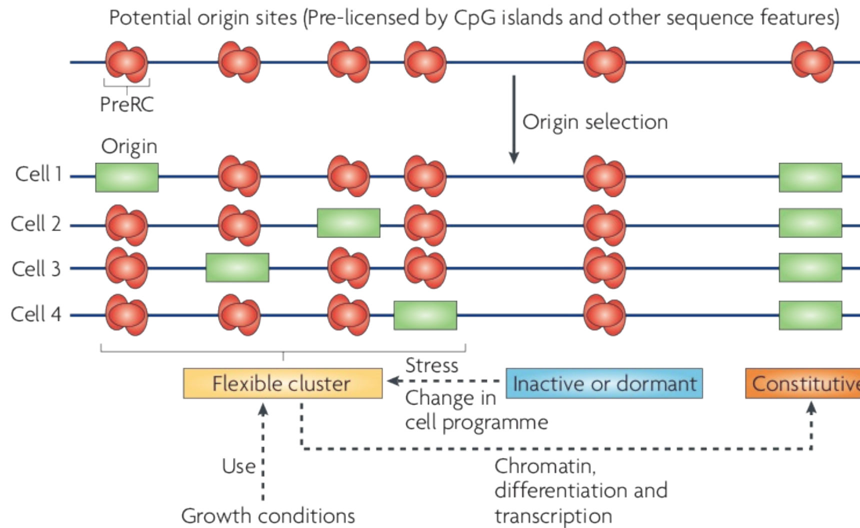


Figure 1.2: Different types of DNA replication origins. The origins that will be activated at the following S phase are selected during G1 phase and may vary according to several parameters, such as the environmental conditions or cell fate. Four examples of DNA replication origin position in a growing cell population are shown. A cluster of flexible origins contains origins that can be activated differently in different cells; according to physiological or abnormal growth conditions. Inactive (or dormant) origins are not frequently used or not used at all; whereas constitutive origins are fixed and always set at the same position by chromatin or transcriptional constraints, inactive origins can be activated increasing the number of origins per replication cluster. Adapted from [13].

Moreover, a post initiation role was recently proposed for the protein Cdc6 which does not affect replication elongation, nor checkpoint activation *in vivo* due to the absence of CHK1 activation [32, 33]. In contrast, Cdc6 seems to be crucial for activation of S phase checkpoints in a Xenopus cell-free DNA replication system [34]. Altogether, this information indicate that the same factors involved in DNA replication initiation are also important actors in the regulation of the replication process at different stages during the cell cycle, being the targets of many checkpoint

controls.

1.1.3 From initiation to elongation in eukaryotic DNA replication

DNA replication starts from the stepwise recruitment of the replication machinery to the various origins on the chromosome. The recruitment process is an essential part of the initiation process, to be distinguished from the subsequent replication of the DNA by the replisome (named elongation). As reported in the previous paragraph, initiation is a major step at which DNA replication is regulated: the ordered recruitment of the pre-replication proteins onto the origin is indeed responsible for controlling the process of initiation of DNA replication in terms of both space and time. Furthermore, the subsequent inactivation or removal of some of the protein prevents re-replication during S phase. For these reasons, initiator proteins are crucial in regulating origin activity. The basic mechanism of initiation occurs in several steps that finally lead to bidirectional replication from the origin. These steps can be summarized as follows.

1. Recognition: labeling of the origin by ORC, Cdc6 and Cdt1;
2. Licensing or initiative assembly: loading of the DNA-helicase (MCM complex or minichromosome maintenance complex), to form the pre-RC;
3. Unwinding: activation of the DNA helicase or by protein kinase activity;
4. Elongative assembly: loading of the complete replisome, including DNA polymerase enzymes and SSB (single-stranded DNA binding protein).

The ordered sequence of these four steps allows the switch from initiation to elongation; each of these steps is briefly summarized in figure 1.3

1. Recognition: in this step, the ORC complex recognizes and marks the origins, which is proposed to occur between the late M and G1 phase [14], and provides an anchorage point for two other proteins entering the complex during the course of G1

phase, Cdc6 and Cdt1. ORC is a six-protein heterocomplex containing ORC1–6 proteins (Figure 1.3) in an equal stoichiometric ratio. It was first isolated in yeast cells due to its specific binding to origin sequences [35]. Although the ORC1–6 proteins are evolutionary conserved in all eukaryotes, the recognition of specific sequences is a property lost in ORC except for the fission yeast *Schizosaccharomices pombe,* in which a preference for AT-rich sequences exists. This is direct evidence for the absence of a consensus DNA sequence in metazoa and, at the same time, leads to the conclusion that ORC cannot be considered as a true "initiator" protein by itself. The most impressive proof of this concept, and also of the preservation of ORC among eukaryotes, is when recombinant ORC1–6 proteins from human were found to replace the frog ORC1–6 proteins *in vitro* to initiate DNA replication in a sequence-independent manner [36]. To date, it is not clear which DNA or chromatin structure ORC recognizes. Most likely this is a particular chromatin structure governed by epigenetic determinants and not primary DNA sequence. This possibility is supported by several, recent observations and will further be discussed later.

Most ORC subunits belong to the superfamily of AAA+ ATPases (ATPases Associated with various cellular Activities) and share conserved motifs [37] (Figure 1.4). The ATP-binding activity is required in the process of origin DNA recognition. Indeed, in *Saccharomyces cerevisiae,* the ORC1 ATPase activity is inhibited until the Cdc6 protein, which is also an AAA+ ATPase, is recruited and activates ORC1 ATPase, thus resulting in the specific recognition of the origin [38]. The role of the ORC6 protein in DNA binding and pre-RC assembly is controversial and represents a sort of enigma. ORC6 is an essential protein for viability in yeast but is not required for DNA binding *in vitro*. In metazoan cells, complexes with lower amounts of ORC6 than the other ORC1–5 proteins are still active, whilst in Drosophila ORC6 was shown to have intrinsic DNA binding activity and any point mutation in its DNA binding domain negatively affects DNA synthesis [39, 40]. In yeast, ORC is bound to origins throughout the cell cycle and re-replication is avoided by phosphorylation of ORC2 and ORC6 by CDK1.

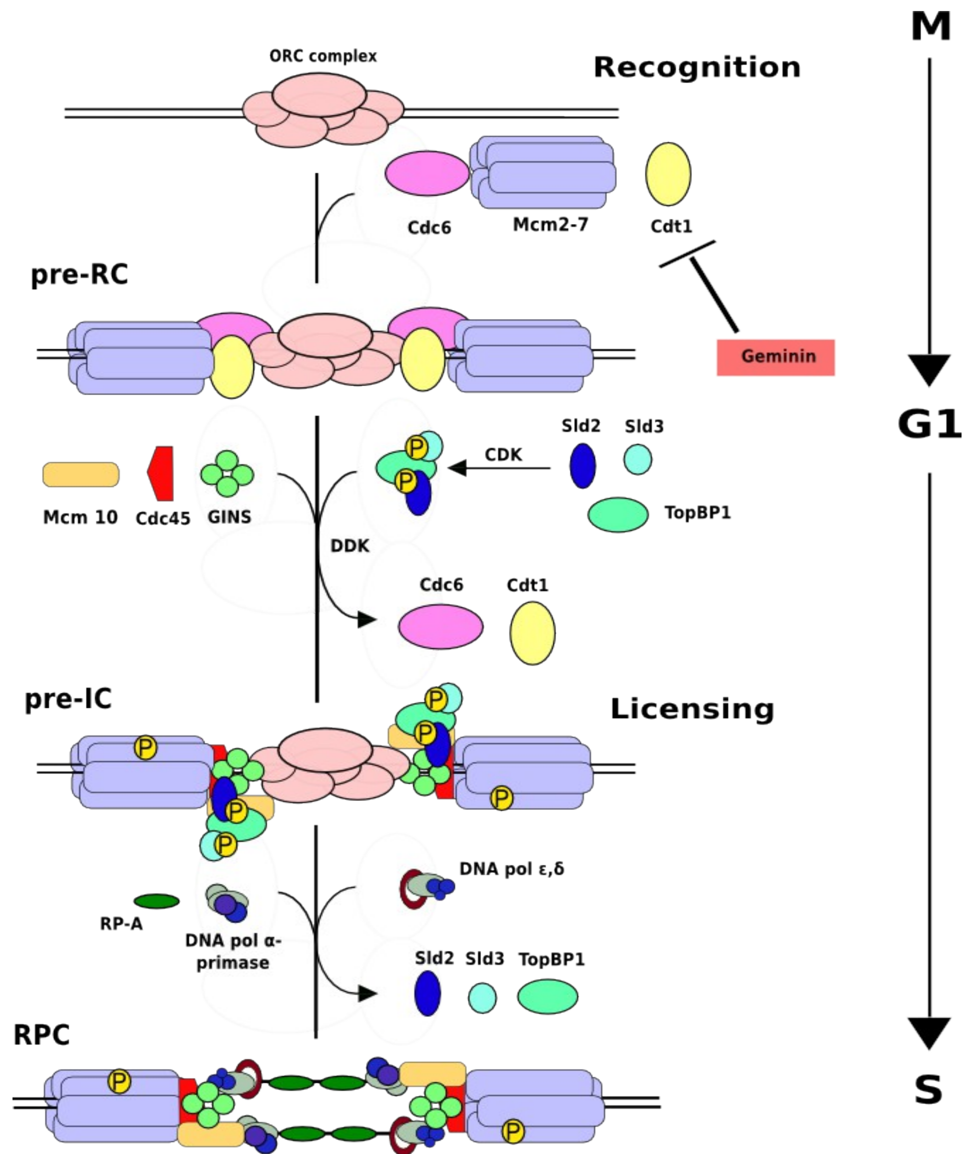


Figure 1.3: Model of the regulation of DNA replication. In eukaryotes, a replication origin is recognized by ORC at the M/G1 transition. Then, Cdc6 and Cdt1 proteins load the MCM helicase to form the pre-RC complex in G1 phase. Geminin inhibits Cdt1 and consequently pre-RC re-formation. CDK and DDK become active in late G1 and activate the MCM helicase; in addition, CDK inhibits any further licensing. To this end, CDK phosphorylates Sld2 and Sld3 proteins and DDK phosphorylates MCM proteins giving rise to the pre-initiation complex (Pre-IC), Finally loading of primase, polymerase and RPA allows DNA replication to start forming its fundamental unit, the replisome. Modified from S. J. Aves, DNA replication initiation-Methods in Molecular Biology (521) 2009.

However, in other eukaryotes, ORC binding is regulated based on a mechanism known as the “ORC cycle” [41]. This is used to avoid re-replication and consists in the dissociation of ORC1 from the chromatin-bound ORC2–5 complex and its subsequent degradation in cells at the end of G1 phase and beginning of S phase [42]. This process is regulated by CDK1-cyclin A phosphorylation [43]. Of note, recent studies have identified roles for ORC proteins other than the direct control of DNA replication initiation [44] (Figure 1.5). ORC1 has been reported to participate in gene silencing via its BAH domain (Figure 1.4), providing a direct interaction with the silent chromatin protein Sir1 in *S. cerevisiae* [45], as well as with heterochromatin protein 1 (HP1) in Xenopus, Drosophila [46] and mammals [47]. In both cases, ORC1 helps Sir1 and HP1 to propagate silenced chromatin. Other ORC proteins have been reported to be important for heterochromatin maintenance. ORC2 and ORC3 are associated with constitutive heterochromatin and HP1 in Drosophila; in human cells depletion of these proteins causes HP1 disruption leading to compromised gene silencing, sister chromatid cohesion and centromere function in mitosis [48]. Depletion of ORC1 and ORC5 also results in loss of HP1, but from large heterochromatin foci instead of the centric one where ORC2 and ORC3 are present during mitosis [49]. Studies in both Drosophila and mammalian cells have revealed that ORC6 coordinates cytokinesis with pre-RC formation and chromosome segregation, independent from the rest of the complex [50]. It has been proposed that ORC6 may also participate in positioning of the ORC at the origins of DNA replication, similar to the role of TFIIB in positioning transcription pre-initiation complex at the promoter [51]. Human ORC6 was shown also to localize to kinetochores and reticular-like structures around the cell periphery during mitosis, and to be necessary for the proper progression of this stage of the cell cycle [52]. Human ORC2 also is present at the centrosome during all the cell cycle and, when depleted, mitotic defects and multiple centrosomes arise [48]. Recently, human ORC1 was reported to have a similar role in controlling centrosome copy number [53].

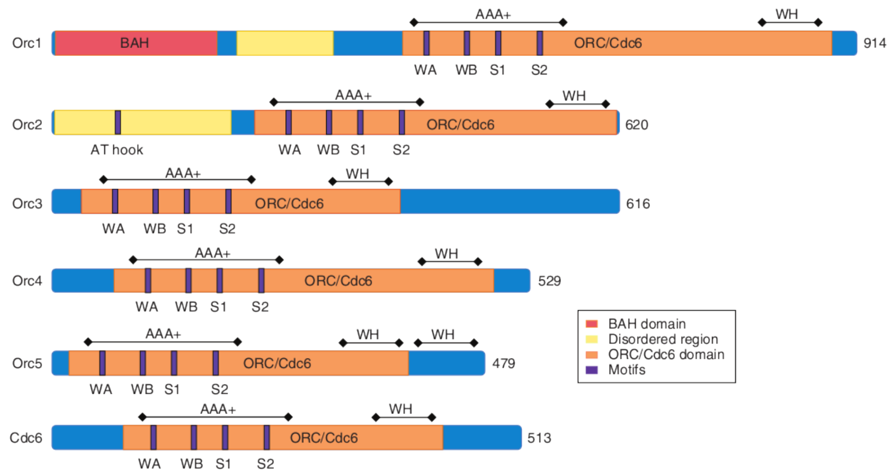


Figure 1.4: Schematic pictures of ORC1-5 and Cdc6 proteins from *S. cerevisiae*. All ORC1-5 and Cdc6 proteins contain an AAA+ domain as part of a larger ORC/Cdc6 homology domain (highlighted in orange). Motifs in the AAA+ domain include Walker A (WA), Walker B (WB), Sensor-1 (S1) and Sensor-2 (S2). The winged-helix domain (WH) is involved in DNA binding. ORC1 contains an additional BAH (bromo-adjacent homology) domain (highlighted in pink). ORC1 and ORC2 have disordered regions (yellow); a DNA-binding, AT- hook motif was identified in *S. cerevisiae* ORC2, and many of these regions have also been found in disordered regions of *S. pombe* ORC4. The total number of amino acids for each protein is indicated at the right side. Adapted from: [37].

2. Initiative assembly. The next step is to load the heterohexameric DNA helicase onto the origin (Figure 1.3). This is accomplished by two proteins, Cdc6 and Cdt1, which recruit the mini chromosome maintenance (MCM) helicase to finally achieve the pre-RC assembly onto the origin. "Replication licensing" is a conventional term that is used to describe the process in which origins are "licensed" when the MCM helicase is loaded onto them in the G1 phase of the cell cycle [54]. Cdc6 is also an AAA+ ATPase (see Figure 1.4), which is required to load the MCM helicase onto the complex in the G1 phase, as shown in experiments performed in budding yeast, which also revealed the importance of its ATPase activity to exert this function [55]. In particular, the Cdc6 and ORC ATPases act sequentially, with Cdc6 required initially. In a recently proposed model, Cdc6 and origin chromatin set off a molecular

switch in ORC for pre-RC assembly [38]. Indeed, in *S. cerevisiae* the ORC1 ATPase activity is inhibited until Cdc6 protein is recruited and activates ORC1 ATPase. This produces a conformational change in the ORC-Cdc6-DNA complex to achieve a ring-like structure with increased specificity for the origin sequence. Origin DNA inhibits ATP hydrolysis by Cdc6 and stabilizes the complex, whereas mutations in the origin sequence can increase Cdc6 ATPase activity, resulting in a less stable Cdc6-DNA complex. This means that ORC binding to the origin is not specific unless Cdc6 is also bound, thus Cdc6 rather than ORC is responsible for origin selection [38]. The structure suggested for Cdc6, which was deduced by comparison with the ORC structure, is similar to the atomic structure of the archaeal homologue, ORC1/Cdc6. ORC1 and Cdc6 proteins are homologues (Figure 1.4), and indeed archaeal species have one protein Orc1/Cdc6 acting both in origin recognition and in MCM helicase loading.

The Cdt1 protein, like Cdc6, is also necessary to load the MCM helicase during G1 phase of the cell cycle of eukaryotes [8]. This protein, which was initially found in fission yeast, is clearly conserved in eukaryotic evolution. As Cdc6 ATPase is required for Cdt1 binding onto the origin in vitro, it has been suggested that a Cdt1-MCM complex is loaded onto the ORC-Cdc6-origin complex during initiation [55] [56]. Cdc6 and Cdt1 then dissociate and, finally, ORC hydrolyzes ATP and this completes the MCM helicase loading reaction [38] [55]. As stated in paragraph 1.1.2, licensing is blocked during S, G2, and M phases of the cell cycle to prevent re-replication. Re-replication is actually avoided by the concurrence of several, redundant mechanisms that block MCM loading during S, G2 and M phases. Pre-RC complexes can be assembled only in the course of G1 phase, but are activated for origin firing only during S phase. A higher level of regulation is catalyzed by CDK, which operates at many redundant levels to avoid licensing in most eukaryotes [57].

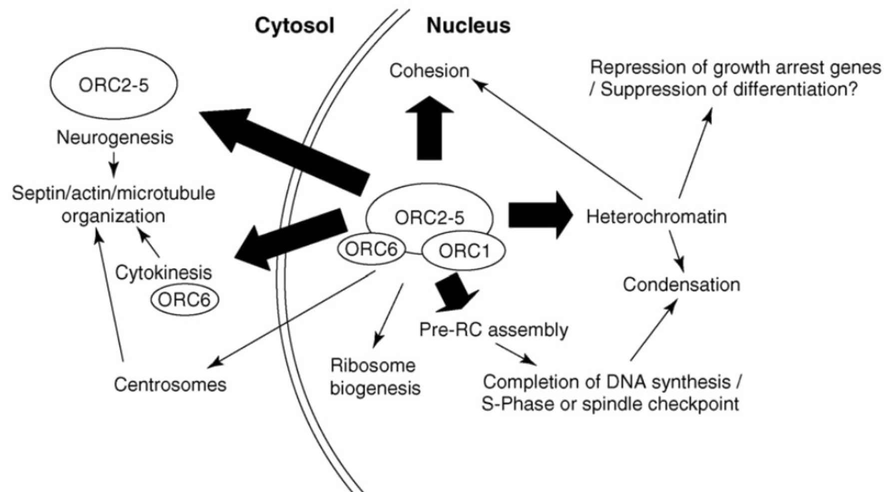


Figure 1.5 The many functions of the origin recognition complex. This diagram describes roles for ORC proteins that are supported by functional evidence (indicated with a thick arrow), while roles that can be explained indirectly or that are supported primarily by localization or physical association of ORC are indicated with a thin arrow. Adapted from [44].

These include the localization and degradation of several pre-RC components. Besides the already mentioned ORC1 protein in higher eukaryotes, another modified protein is Cdc6: in yeast it is degraded after CDK phosphorylation [58]. Another degree of regulation to block re-replication occurs through a protein known as Geminin (Figure 1.3), which was discovered in frog egg extracts [59] and is only found in metazoans. Geminin binds to and inhibits Cdt1 thus preventing replication licensing by blocking the loading of the MCM helicase [59]. The same role for Geminin is reported to also occur in human cells [19]. The redundancy of these mechanisms in avoiding re-replication has been proposed to provide a key driving force in the evolution of licensing control [60]. Because there is no any unique mechanism effective to inhibit pre-RC components, multiple mechanisms are required for an efficient block of the re-replication. However, as the number of inhibitory mechanisms increases, the relative importance of any single mechanism decreases. During evolution, these regulatory mechanisms may be gained or lost. This could explain the different regulatory mechanism of Cdc6 between yeast and human

cells as well as the appearance of Geminin in metazoa to provide an additional mechanism for preventing re-replication, which may have been important in supporting an increase in genome size with respect to lower eukaryotes.

3. Unwinding and 4. Elongative assembly. These two steps refer to the activation of helicase activity at the origin and to replisome assembly respectively; due to their close interconnection they are therefore described together.

The MCM complex is believed to be the engine of the replicative helicase. This complex is a hexamer comprising of six related polypeptides (MCM2–7), all of them with AAA+ ATPase activity. They are coded by a family of six paralogous genes, which are conserved from yeast to human. All six members are essential genes as described in a pioneering work in fission yeast in which MCM was as a complex that contained all six subunits in 1:1:1:1:1:1 stoichiometric ratio having a ring-like structure [61].

In G1 phase, pre-RCs with the MCM2–7 helicase bound are present on almost all origins. Indeed, about 90% of all origins that are bound by ORC also contain the MCM complex bound [62]. Nevertheless, MCM is loaded in an inactive state in the pre-RC, when CDK activity is low. The next step is to activate the MCM helicase. This is achieved by the binding of several other proteins to the origin, up to the loading of the replisome (Figures 1.3). The multi-protein complex assembled on the origin at this stage is referred to as the pre-initiation complex (pre-IC), and is required for the activation of the MCM2–7 helicase. Both helicase activation and replisome loading require phosphorylation by two different kinases which are regulated independently of each other, but by similar mechanisms. The kinase activity is established by a protein heterocomplex, thus both kinases are inactive in monomeric form and are activated by the binding of an activating subunit, Cyclin for CDK and Dbf4/Drf1 proteins for DDK [63], respectively. Thus, CDK is Cyclin-dependent kinase (comprising heterodimer of different Cdk and a Cyclin) and DDK is Dbf4-dependent kinase (comprising of Dbf4/Drf1 and Cdc7). In mammals, while there is only one DDK, there are at least four CDKs (Cdk1–4) and four classes of cyclins (A,

B, D, and E) required for cell cycle progression [64]. Thus, the substrate specificity by different Cdk-cyclin complexes drives the cell cycle. In budding yeast, there is only Cdk1 or Cdc28 enzyme, but there are six B-type cyclins (Clb1–6) needed for S and M phases [64]. Cdk1-Clb5 complexes are active in regulating DNA replication and Cdk1-Clb2 for regulating mitosis. The Cdk2 homologue is used in DNA replication. By analogy, Cdk2-cyclin E and Cdk2-cyclin A act as yeast Cdk1-Clb5 for DNA replication, whereas Cdk1-cyclin B act as yeast Cdk1- Clb2 for mitosis [22].

Cell cycle regulation of the unstable subunit ensures cell cycle regulation of the kinase activity. With CDKs, other levels of regulation occur including protein inhibitor binding, phosphorylation by other kinases and cyclin subcellular localization [64]. About DDK, the mechanism is simpler because Dbf4 protein is absent in G1 phase due to its proteosomal degradation by the APC (anaphase promoting complex) and as cells enter S phase, Dbf4 is stabilized and the APC is inactivated by CDK phosphorylation.

Much evidence indicates that the MCM2–7 complex is a target of phosphorylation by DDK, and this occurs in several eukaryotes [65]. Studies performed in yeast have identified phosphorylation sites in the N-terminus of MCM4, MCM2 or MCM6 to be important for formation of the pre-IC and for DNA replication [66]. Pre-RC activation due to phosphorylation of the MCM complex by DDK leads to the loading of additional replication factors, such as MCM10, the GINS complex and Cdc45 giving rise to the pre-initiation complex (Pre-IC).

MCM10, which is not a MCM2–7 homologue [67], is needed for the recruitment of the Cdc45 protein after pre-RC formation and for stabilizing the replisome, a mechanism conserved from yeast to humans [68].

The GINS complex, the name of which is based on the numbers 5, 1, 2, and 3 in Japanese (Go, Ichi, Nii, San), is composed of the Sld5, Psf1, Psf2, and Psf3 proteins, and is needed for replication by functioning interdependently with Cdc45 protein in the loading of the replisome. Most of these proteins are conserved in eukaryotic organisms [69]: only yeast Sld3 does not have any homologue in metazoa, whereas Sld2 and Dbp11 are related to mammalian RecQ4L and TopBP1, respectively; the

GINS complex is highly conserved in yeast [70], Xenopus [71] and human [72].
MCM complex phosphorylation also leads to the loading of Cdc45 protein onto origin chromatin in a mechanism which is conserved from yeast [70] to human [69]. Cdc45 protein is needed for loading of the replisome, including DNA polymerases and RPA, the eukaryotic SSB (Figure 1.3), and moves with the replication fork [16].
An interesting point is: how does phosphorylation of MCM2, MCM4, or MCM6 by DDK activate the helicase and allow the replisome loading? One hypothesis is that DDK phosphoryation leads to a conformational change in the MCM5 protein that activates the helicase and represents a signal for the subsequent binding of the Cdc45 protein [22]. In some cases, Cdc45 has been reported to bind the origin earlier in G1 phase, before MCM activation by DDK. In this case, it is possible that Cdc45 protein may be weakly bound to origin chromatin in the G1 phase, and that it is later stabilized by CDK rather than DDK regulation.
The role of CDK in promoting origin activation has also been thoroughly investigated. In yeast, CDK phosphorylates Sld2 and Sld3 [73], causing them to bind Dpb11 (DNA Polymerase B possible subunit, a subunit of DNA polymerase ε holoenzyme, also called Pol2 or PolB), which in turn serves as an anchor for DNA polymerase, RPA and the GINS complex to reach the replisome.
To summarize, a large number of proteins are needed to load the replisome onto the origin (many of them appear in Figure 1.3). These proteins help to activate the MCM helicase. Indeed, the association of MCM2–7, Cdc45, and GINS constitutes a complex named the CMG complex which, when purified from Drosophila embryos, has helicase activity in vitro [74]. Moreover, these proteins bring the DNA polymerases onto the origin, thereby coupling helicase activation and replisome loading. It is also evident that CDK and DDK regulate similar events independently. The DNA is unwound by the helicase and then replicated by the replisome. It is not clear which is the exact role of MCM10 in this model but its requirement for Cdc45 loading and replisome stability is well established.

1.1.4 Additional requirement for DNA replication

Despite several replication origins have been identified to date, no consensus sequence has been reported to predict their localization in metazoans. In addition, ORC, the protein complex that marks all replication origins and is needed for the sequential assembly of the full replicative complex (RC), exhibits little sequence specificity in higher eukaryotes [75] [76]. Recent data highlight that metazoan origins are modular and hence the binding of ORC might be determined by the combination of different elements encompassing both DNA primary structure (e.g. AT-rich sequences and CpG-islands, promoter regions, dinucleotide repeats, matrix attachment regions) and local DNA topology and epigenetics [77] (Figure 1.6). This latter consideration is strongly emerging in last decade due to different data suggesting that chromatin affects the selection and activation of DNA replication origins. In *S.cerevisiae*, ORC is important in nucleosome positioning around ARS1 origin [78] which might help to conserve its epigenetic and autonomous status [13]. Binding of bacteria initiator DnaA is dependent upon negatively (-) supercoiled DNA [79, 80] and a similar mechanism is required for origin function in bacteriophage λ [81].

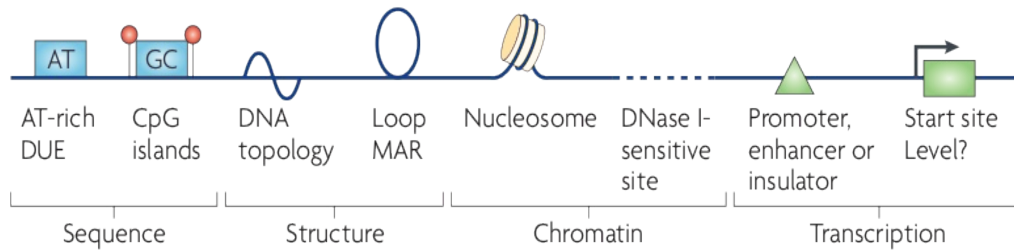


Figure 1.6: Proposed features to determine the selection and activation of replication origin. Many characteristics have been described at metazoan replication origins which can contribute to the selection of a given origin. AT-rich elements, CpG islands, and DNA regions that can be easily unwound (DNA unwinding elements (DUEs) have been reported. At the DNA level, secondary structures, such as cruciform DNA and the formation of loops and nuclear matrix interactions (matrix attachment region or MAR) have been reported. At the chromatin level, nucleosome-free regions, Dnase-sensitive zones as well as histone acetylation have been noticed, but whether these characteristics direct participate in origin definition or are a consequence of chromatin organization for transcription is difficult to conclude. The presence of a possible link between transcription features and replication origin recognition has been described but evidence remains scarce. Adapted from [77].

In Drosophila (-) supercoiled DNA dramatically increases ORC affinity but not specificity to DNA [82], suggesting a common mechanisms for DNA replication initiation across different species. Again in Drosophila, it was shown that Histone 4 (H4) hyperacetylation foci overlap with ORC2 but not with Double parked protein (Dup – a replication protein present in the S phase also in humans) foci, indicating co-localization with pre-RC but not with moving forks [83]. A driving force for origin selection and activation could also be the chromatin re-organization associated with development and/or with cell cycle progression, particularly with the G1/S transition. These features favors a model [83, 84] in which two stages need to be passed in G1 in order to achieve origin specification: the former is the timing decision point (TDP), where early and late replication domains are established, and the latter is the origin decision point (ODP), which selects only a portion of the sites previously licensed to be used in the next S phase [85] (Figure 1.7).

It must be underlined that this view also hints at a role for DNA topology in

establishing an origin of DNA replication. A particular, chromatin structure is maintained by a peculiar DNA conformation and topology [86]. If this model holds true, topoisomerases, which are enzymes able to alter the topology of a DNA region, should play a determining role in origin function, as topology-modifying events may be required for the formation of the pre-RC. Indeed, both topoisomerases I and II were found to interact with the Lamin B2 origin, and to be essential for origin firing in close interaction with ORC [82, 87].

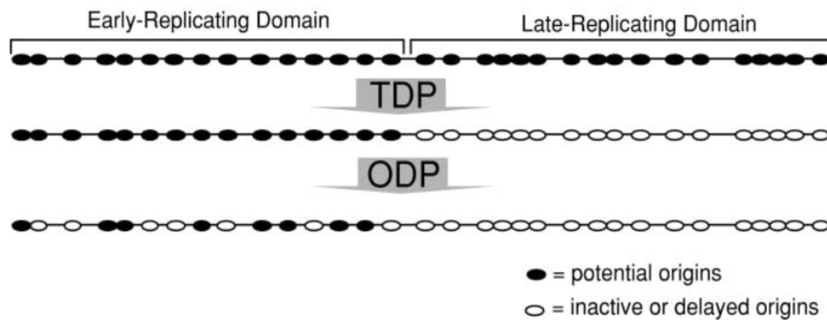


Figure 1.7: Model for the progressive restriction of initiation potential during G1 phase. In early G1 phase, many sites distributed throughout the genome have an equal potential to be used as early replication origins. At the time decision point (TDP), late replicating chromosomal domains become excluded from the pool of potential early replicating origins. At this time, origins within these early replicating domains still have an equal potential for initiation. At the origin decision point (ODP), a subset of these potential origins are chosen for initiation in the upcoming S phase. Adapted from [85].

Together with the DNA topological changes, other elements link replication and transcription. Among these, the involvement of transcription factors (TFs) is the most striking one, due to its strong influence on origin selection itself. Considerable evidence indicating the direct involvement of transcription factors in DNA replication has arisen from the investigation of DNA viruses, including adenovirus, polyomavirus (Py), SV40 and the bovine or human papilloma viruses (BPV and HPV, respectively) [88]. The binding sites of transcription factors are functional elements within or proximal to the replication origins of a variety of viral and eukaryotic

systems [89]. Thus, DNA-bound transcription factors appear to actively recruit initiator proteins to the DNA, indicating their central, albeit auxiliary, role in viral replication initiation [90]. The recruitment of initiators by DNA-bound transcription factors is not limited to virus-encoded initiators. In its latent infection stage, during the replication process, the Epstein Barr virus (EBV) employs the cellular ORC as initiator proteins [88]. Transcription factors aid the recruitment of initiator proteins to origins not only by direct interaction but also by altering local chromatin structure. Chromatin structure generally inhibits DNA replication as well as transcription since it reduces accessibility of proteins to DNA. It has to be underlined that some transcription factors are able to bind their target sites even when these are folded into the nucleosome [91]. In *S. cerevisiae* various transcription factors positively or negatively regulate the ARS1 replication activity [92]. ARS1 chromatin structure appears to be altered and this change seems to correlate with the activity of the origin; in particular, TFs might regulate the initiation step after ORC binding because ORC can also bind the silent origins [62]. In Drosophila, binding sites for Myb and E2F are located in the chorion gene and are required for its amplification [93]. Flies expressing mutants of Myb and E2F1 or of its partner Dp1 show diminished chorion gene amplification and mislocalization of ORC2 [93]. The autonomously replicating monkey sequence *Ors8* contains binding sites for the transcription factor Oct-1 [94].
Interestingly, important transcription factors (and proto-oncogenes) such as c-Myc [95] and the homeotic box family of proteins (HOX) [96, 97] and USF-1[98] were recently found to bind well characterized human replication origin sequences and participate in origin activation, again hinting at transcription events as actors in origin specification.

1.1.5 The still obscure determinants of origin specification

As already mentioned in paragraph 1.1.1, the selection of defined and adequately distributed replication origins seems to represent the safest way to achieve complete genome duplication in eukaryotes. Specific sites named ARS, determined by precise sequence motifs, were found in S. cerevisiae [92]. However, this is not the case for other eukaryotes, in which the sequences directing replication initiation appear to be far less defined. Extreme situations have been reported in Drosophila and Xenopus early embryos, in which replication initiation occurs at random sites along the chromosomes. Strikingly, during embryonic development, in correspondence to remodeling of nuclear structure and chromatin organization, initiation events become restricted to preferred regions [99]. In agreement with the fact that preferred sites of initiation are selected during development, DNA synthesis does not start at random locations in somatic mammalian cells. Also in this context, the mechanism that governs the selection of replication origins in metazoan genomes taking place in the G1 phase of the cell cycle, is still not clear. What makes the understanding of origin specification difficult is mainly the high degree of degeneration of metazoan origin sequences [77]. In human cells, only a few of the overall origins are well characterized and these share no evident sequence similarity [100, 101]. More recently, genome-wide approaches have led to the identification of several origins, but still no consensus sequence has been clearly identified, besides a relative frequency of CpG islands and asymmetric A/T stretches in correspondence to highly active origin sequences, and of the presence of transcription factor binding sites [86, 102] (Figure 1.8). One of the most probable candidates, which could likely contribute to origin specification, is a local chromatin environment ideally suited for the pre-RC assembly. In fact, this transition from sequence-specific to epigenetic specification of replication origins, might have contributed to the plasticity required by a multicellular organism to express a wide variety of genetic programmes from an identical genetic content [13]. In this context, it is clear that not only DNA sequence but also

epigenetic marks have to be correctly transferred to daughter cells. However, chromatin accessibility could not be the only requirement for origin specification, as many specific sequences in a range from 1 to 6 Kb have been described to be capable of maintaining their activity at ectopic positions in the genome [103]. This could result from the combination of several elements, due to the origin modularity (refer to paragraph 1.1.4), such as an open and transcriptionally active chromatin structure, bent DNA structures, close proximity to gene promoters, binding sites for sequence-specific proteins or asymmetric AT-rich stretches. According to this scenario, origins could take advantage of (or "parasitize") regions that are maintained in an accessible conformation for structural reasons or to facilitate transcription, as suggested by the preference of origins to map near promoters in many cases [86].

This "opportunistic" origin specification would remove the selective pressure to maintain each single origin sequence in the genome for its individual contribution to replication. This model is supported by at least two considerations:

i) eukaryotic origins are present in excess through the genome; not all of them fire in every S phase and many remain silent and are inactivated by replication forks passing by during S phase [104];
ii) open chromatin appears to be the underlying feature that is deterministic for ORC binding as revealed by genome-wide approaches [25].

Thus, origin specification likely relies on other factors; these could include proteins, which display a preference for certain origin features (like sequence or other structural properties) and target the RC proteins onto the origin by direct or indirect protein-protein interactions. So far, several proteins have been identified, that can either specify sites of ORC binding, or in any case have a role in DNA replication initiation. Among these it is worth to mention AIF-C, Trf2, Ku80, EBNA1 and HMG1a proteins [105, 106, 107, 108]. In many of these cases, the proposed proteins were shown to function as "ORC-chaperons" in targeting ORC to chromatin regions, thus contributing to origin formation and to a more specified binding of ORC.

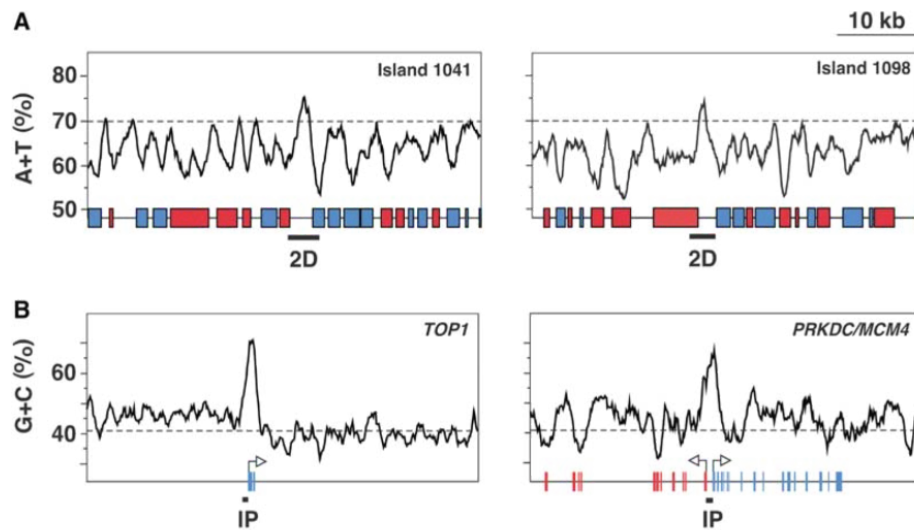


Figure 1.8: A/T-rich and C/G-rich islands at DNA replication origins. A. A/T content through the 1041 and 1098 A/T-rich islands in *S. pombe*. Red and blue boxes represent genes transcribed towards the left and the right, respectively. Genomic regions labeled 2D represent restriction fragments containing an active ORI. Dashed lines indicate the average intergenic A/T content (70%). B. G/C content across the first two exons of the human TOP1 gene and the bidirectionally transcribed PRKDC and MCM4 genes. Arrows indicate the transcription direction. Red and blue bars represent exons. Black boxes labeled IP represent DNA fragments immuno-precipitated by ChIP analysis with anti-human Orc2p antibodies. Dashed lines indicate the human average genomic G/C content (41%). Scale as in (A). Adapted from [86].

In this context, it was very interesting to find that another family of transcription factors, namely the AP-1 proteins, displays an affinity for origin sequences [102, 109]. A possible role for AP-1 proteins in origin decision would be particularly intriguing, because it could represent a basis for the interplay between DNA replication and transcription, as well as explain the proto-oncogenic properties displayed by these proteins.

In order to better understand to what extent AP-1 proteins could be involved in origin function, some knowledge about their structure and function is required; accordingly, the next paragraph will be focused on this family of transcription factors involved in several aspects of the cell life.

1.2 The AP-1 protein family

1.2.1 AP-1 complex protein composition and general features

Activator protein-1 (AP-1) was first identified as a transcription factor that binds an essential cis-element of the human metallothionein lla (hMTIla) promoter and is required for its optimal basal activity both *in vivo* and *in vitro* [110, 111]. Soon after, the AP-1 binding sites were also recognized as taking part in the transcriptional activation of several cellular and viral genes in response to cell treatment with the tumor promoter TPA (12-O-tetradecanoylphorbol 13-acetate), being located at regions thus collectively named TPA-responsive element (TRE) [112].

AP-1 collectively describes a group of diverse nuclear proteins structurally and functionally related, forming dimers and belonging to different sub-families including:

- Jun protein family comprising c-Jun, JunB and JunD;
- Fos protein family composed of c-Fos, FosB, Fra-1 and Fra-2;
- Activating transcription factor (ATF) protein family consisting of ATF-1, ATF-2 (also known as CREB2 or CREBP-2), ATF-3, ATF-4, ATF-5, ATF-6, ATF-7 (also known as ATF-A), B-ATF and CREB1;
- Jun-dimerizing partners (JDP) protein family encompassing JDP-1 and JDP- 2;
- Musculo-aponeurotic fibrosarcoma (Maf) protein family comprising c-Maf, Maf-A, Maf-B, Maf-G, Maf-F, Maf-K and Nrl.

AP-1 activity is induced by many diverse stimuli, both internal and external to the cell, such as growth factors, cytokines, neurotransmitters, polypeptide hormones, cell–matrix interactions, bacterial and viral infections and a variety of physical and chemical stresses [113]. These stimuli activate mitogen activated protein kinase (MAPK) cascades that enhance AP-1 activity through the phosphorylation of distinct

substrates as it will be reviewed in paragraph 1.2.3.

Each of these proteins is differentially expressed and regulated; thus, in every cell type the broad combinatorial possibilities provided by the large number of AP-1 proteins determine the AP-1 dimer binding specificity and affinity and, consequently, the spectrum of regulated genes [114]. A common feature of all these proteins is the presence of an evolutionarily conserved basic DNA-binding domain combined with a leucine zipper, namely the bZIP domain. Dimerization is the fundamental pre-requisite for DNA binding mediated by the basic domain [115], with the composition of the leucine zipper being responsible for dimer specificity and stability. Elucidation on the structure of the AP-1 proteins will be discussed in the next paragraph. Whereas the Jun proteins exist both as homo- and hetero-dimers, the Fos proteins, which cannot homodimerize, form stable heterodimers with Jun and Maf proteins enhancing their DNA-binding activity [116]. Jun-Fos heterodimers bind preferentially the TRE DNA sequence, which is a heptamer consensus sequence (5′-TGA(C/G) TCA-3′), whereas Jun-ATF dimers bind with higher affinity to another consensus sequence known as the cyclic AMP responsive element (CRE) (5′-TGACGTCA-3′) [117]. The Jun-JDP dimer instead, binds with the same affinity to both CRE and TRE sequences. However, the AP-1 binding site exhibits some degree of degeneracy [118]; as a consequence, the sequences to which the AP-1 complex binds may differ upon interaction with structurally unrelated proteins such as NFAT, or proteins from the Ets or Smad families [119] and thus, may also differ in many natural promoters and enhancers of AP-1-regulated genes [117].

The individual Jun and Fos proteins have significantly different transactivation potentials. Whereas c-Jun, c-Fos and FosB are considered strong transactivators, JunB, JunD, Fra-1 and Fra-2 exhibit only weak transactivation potential. In fact, under specific circumstances, the latter might also act as repressors of AP-1 activity as a competitor for binding to AP-1 sites or by forming ‘inactive’ heterodimers with c-Jun, c-Fos or FosB [114].

A wide range of physiological and pathological stimuli regulate AP-1 activity such as cytokines, growth factors, stress signals, infections and oncogenic stimuli [120].

Regulation of the network of AP-1 activity can be achieved at different levels, including changes in AP-1 subunit-encoding gene transcription, control of the stability of their mRNAs, post-translational processing and turnover of pre-existing or newly synthesized AP-1 subunits [117], and specific interactions between AP-1 proteins and other transcription factors and cofactors. This network of molecular regulations will be reviewed in the subsequent paragraphs.

1.2.2 Structure of AP-1 proteins

Dimerization of Jun and Fos proteins occurs between their so called “leucine-zipper” (ZIP) regions via hydrophobic interactions, as originally elucidated used site-directed mutagenesis [121]. The “leucine-zipper” is one of the many structural motifs that characterize proteins able to bind DNA. It consists of an extensive α-helix in which every seventh amino acid a leucine is present. Due to this arrangement, the leucine side chains protrude from one side of the α-helix and form a hydrophobic surface that mediates dimerization [122]. The hydrophobic surfaces of two α-helices wrap around each other as a result of van der Waals interactions, meanwhile closely located hydrophilic amino acids make contact tightening the overall structure in a thermodynamically favorable manner (Figure 1.9). Leucines in the ZIP domain of Jun can be replaced with other hydrophobic residues such as phenylalanines, without any adverse effects on the formation of Jun:Fos heterodimers [123]. In addition, other hydrophobic residues present between the leucines, that together with them form the characteristic 3-4 repeat of α-helices involved in “coiled-coil” interactions [124], are also as important for mediating Jun:Jun and Jun:Fos dimerization [125]. However, hydrophobic interactions alone do not seem to account for the specificity in dimer formation among ZIP proteins. For instance, c-Fos dimerizes with the various Jun proteins but not with GCN4, Myc or another Fos molecule [121, 125]; CREB, which is another ZIP protein, forms homodimers that interact with the cyclic-AMP response

element (CRE) [126] and was not found to interact with either c-Jun, JunB or c-Fos [125]. Surprisingly CREBP-2, which is a protein highly related to CREB, was found to form heterodimers with c-Jun but not with c-Fos and these heterodimers, as well as CREB homodimers, interact with the CRE but not with the TRE sequence [127] [128]. Measurement of the dissociation temperatures, indicates that the increased DNA-binding activity of the Jun:Fos heterodimer is due to its increased thermostability if compared to the Jun:Jun homodimer. Whereas the heterodimer dissociates between 37°C and 42°C, the homodimer dissociates between 25°C and 37°C [125]. The higher thermostability of the heterodimer is responsible for potentiating its DNA-binding activity by increasing the number of molecules present at any given time in the dimeric state. *In vitro* binding experiments of c-Jun:c-Fos heterodimers with the TRE sequence have shown that addition of an excess of unlabeled binding sites to preformed protein-DNA complexes, resulted in a very rapid disappearance of the preformed DNA complex containing c-Jun homodimers, while the protein-DNA complex formed by c-Jun:c-Fos heterodimers was much more stable [129].

c-Fos, on the other hand, does not dimerize even at 4°C [125] and this can be explained by the presence of electrostatic repulsions between negatively-charged side chains, which are abundant in its ZIP region. Thus, once the ZIP domain of c-Fos is replaced with the one of GCN4 or c-Jun, chimera proteins are capable of homodimerization [130]. While the ZIP region mediates the dimerization of these proteins and hence dictates the specificity of complex formation, the interaction with DNA, instead, occurs via a region found immediately upstream of the "leucine-zipper" (Figure 1.10). This region is known as the "basic region" due to the abundance of positively charged residues [131]. Sequence analyses reveal that the "basic region" is highly conserved among all of the Jun and Fos proteins [131] and that it is also conserved in the various CREB and ATF proteins which interact with a sequence similar to the TRE. Site-directed mutagenesis provides a proof that whereas this region is only responsible for DNA-binding, it is not involved in dimerization.

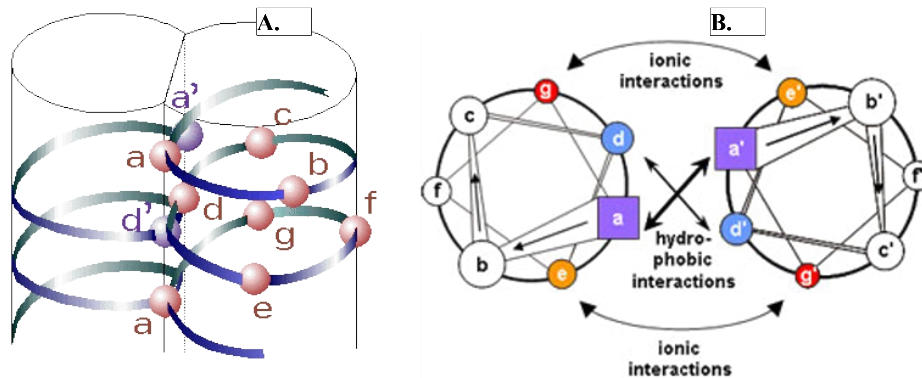


Figure 1.9: Coiled coil protein structure. A coiled-coil protein consists of two identical strands of amino acid sequences that wrap around each other. The first and fourth position (a and d) are generally apolar or hydrophobic amino acids. When the two strands coil around each other, positions a and d are internalized, stabilizing the structure (A), while positions b, c, e, f, and g are exposed on the surface of the protein. Positions e and g are tighten the structure by ionic interactions (B). Adapted from Hitchcock-DeGregori – Tropomyosin 2008.

The spacing between the ZIP and the "basic region" is also very critical. In fact, a duplication of five amino acids that are located between the two regions, and therefore alters their phasing, generates a c-Jun protein variant that can still dimerize with wild type c-Jun or c-Fos but is incapable of binding to DNA [132]. Furthermore, the chimeric dimers formed between the variant protein and either c-Jun or c-Fos also fail to bind DNA and inhibit transactivation [125]. In addition to these domains, a region close to the basic domain required for transcriptional activity of dimers, is the transactivation domain (TAD) (Figure 1.10). Within the TAD N-terminus different phosphorylation sites are present, thus phosphorylation of serine 63 and serine 73 residues of c-Jun by the Jun N-terminal kinase (JNK) family of kinases results in a large increase in its ability to interact with the CBP/p300 family of cofactors and, similarly, in the transcriptional activation potential of the protein [133]. The N-terminus of c-Jun also contains a δ-domain, which is the docking site for JNK and mediates ubiquitin-dependent degradation of the protein [134]. The presence of conserved sequences outside the ZIP and basic regions also in the various Fos

proteins [135] suggests that these regions may be involved in the interaction with various components of the transcriptional machinery. In fact, deletion of the c-Fos C-terminus, strongly reduces its ability to cooperate with c-Jun or JunD [136].

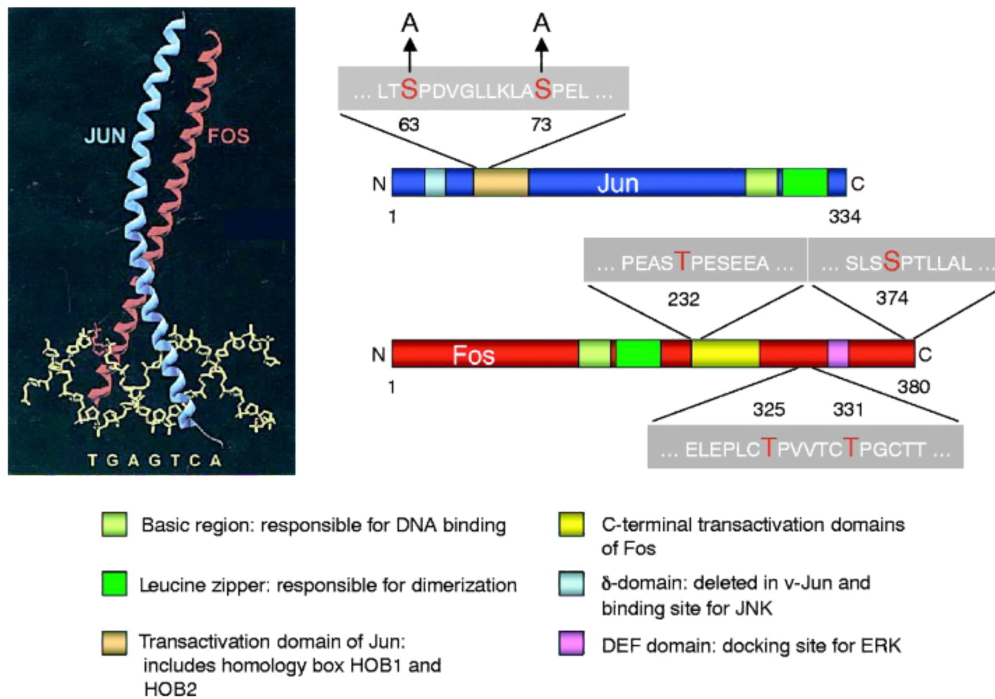


Figure 1.10: The Jun-Fos heterodimer. The c-Jun and c-Fos proteins exhibit several domains, including the so called bZIP domain or leucine zipper plus the basic domain which is required for their interaction with the AP-1 site (TGAGTCA) forming an X-shaped α-helical structure. In addition, transactivation domains and docking sites for several kinases, such as JNK or ERK, are present. These kinases modulate the activity of both proteins phosphorylating two serine and threonine residues. JNK specifically phosphorylates serine residues within the transactivation domain of c-Jun at position 63 and 73 and thereby regulates its transactivation activity; in fact a Jun mutant in which these residues are mutated to alanine (Jun-AA) generates a protein that cannot be activated by JNKs. ERK phosphorylates threonine residues at positions 325 and 331 and a serine residue at position 374 of c-Fos. Additionally, a c-Fos-related kinase phosphorylates a threonine residue at position 232 of c-Fos . Adapted from [114].

1.2.3 AP-1 expression, regulation and activity

The expression of the AP-1 proteins is critical for the decision of the cell fate and thus is regulated at multiple levels, which include control the transcription of their genes, post-translational modifications and dimer composition (refer to paragraph 1.2.1). It has to be underlined that expression of c-Fos and c-Jun in response to many different stimuli causes protein kinase C (PKC) activation, but also their own transcription is induced by PKC activation [137]. While expression of c-Fos is very rapid and highly transient as well as is its turnover as a protein [138], induction of the c-Jun mRNA in response to stimulation is also transient, but the messenger is significantly more stable. [139]. Other agents, like TNF-α or TGF-β [140] lead to a longer lasting induction of c-Jun while their effects do not modify the c-Fos mRNA stability. Three cis-elements have been found to mediate c-Fos induction. The first one is a CRE sequence proximal to the TATA box occupied by ATF or CREB [141]; the second one is a Sis inducible enhancer (SIE) recognized by STAT proteins [142] and the last one is a serum response element (SRE) recognized by a dimer of serum response factor (SRF) and the ternary complex factors (TCF) [143]. Analyses of the promoter of the human c-Jun gene revealed that, upstream of two TATA-like sequences located 24-30 bps upstream a cluster of transcription initiation sites, there is a sequence recognized by the AP-1 complex itself, suggesting that transcription of c-Jun is subjected to a positive autoregulatory loop [144]. Due to the importance of AP-1 proteins for cell proliferation, presence of negative regulations are required for normal cell function. As already mentioned, the c-Fos transcript is subjected to a rapid turnover due to the presence of an RNAse target, AU-rich sequence in the 3' untranslated region (3' UTR); also c-Fos down-regulates its own gene product in a negative autoregulatory manner [145]. The same AU-rich sequence is also present in the c-Jun transcript, but in this case down-regulation of c-Jun is due to the binding of JunB or JunD homodimers or c-Jun:CREB and c-Jun:ATF-2 heterodimers to the AP-1 binding sites upstream the c-Jun gene [126, 128, 146]. The most common post-

translational modification known to regulate/modulate AP-1 activity is phosphorylation carried out by ERK, JNK and p38; all of these proteins belong to the mitogen activated protein kinase (MAPKs) group (Figure 1.11). Furthermore, AP-1 proteins are also regulated by ubiquitination, which targets proteins for proteasome-mediated degradation [147]. The highly conservation degree of the phosphorylation sites throughout evolution of the Jun protein family suggests that they play an important regulatory role [131]. JNK phosphorylation of serine residues 63 and 73 and of threonines 91 and 93 potentiates c-Jun transcriptional capacity and stabilizes the protein increasing its half life [148].

AP-1 proteins are considered to be master regulators of cell life and death due to the wide range of cellular processes they regulate [113]. Control of cell proliferation by AP-1 seems to be mainly due to its ability to regulate the expression and function of several cell cycle regulators such as cyclin D1, cyclin A, cyclin E, p53, $p21^{Cip1}$, $p16^{Ink4a}$ and $p19^{ARF}$ [113, 149]. Analysis of cell culture models reveals that fibroblasts lacking both c-Fos and FosB lose their proliferative capacity, whereas deficiency of only one of these factors does not affect cell proliferation. Instead, cells lacking c-Jun enter premature senescence after the first passage in culture [150] and, even after immortalization, they proliferate slower than wild type cells. Altered expression of individual AP-1 members have revealed unique and crucial roles for each Jun protein, but also some functional redundancy among the Fos proteins. Mice lacking c-Fos are viable but present an osteopetrotic phenotype due to the absence of the cells responsible for resorbing bone, namely the osteoclasts [151]. Furthermore, these mice also present abnormalities in their hematopoietic system. Interestingly c-$Fos^{-/-}$ and $FosB^{-/-}$ double knockout mice are smaller than their wild type counterpart, whereas single knockouts are not [152].

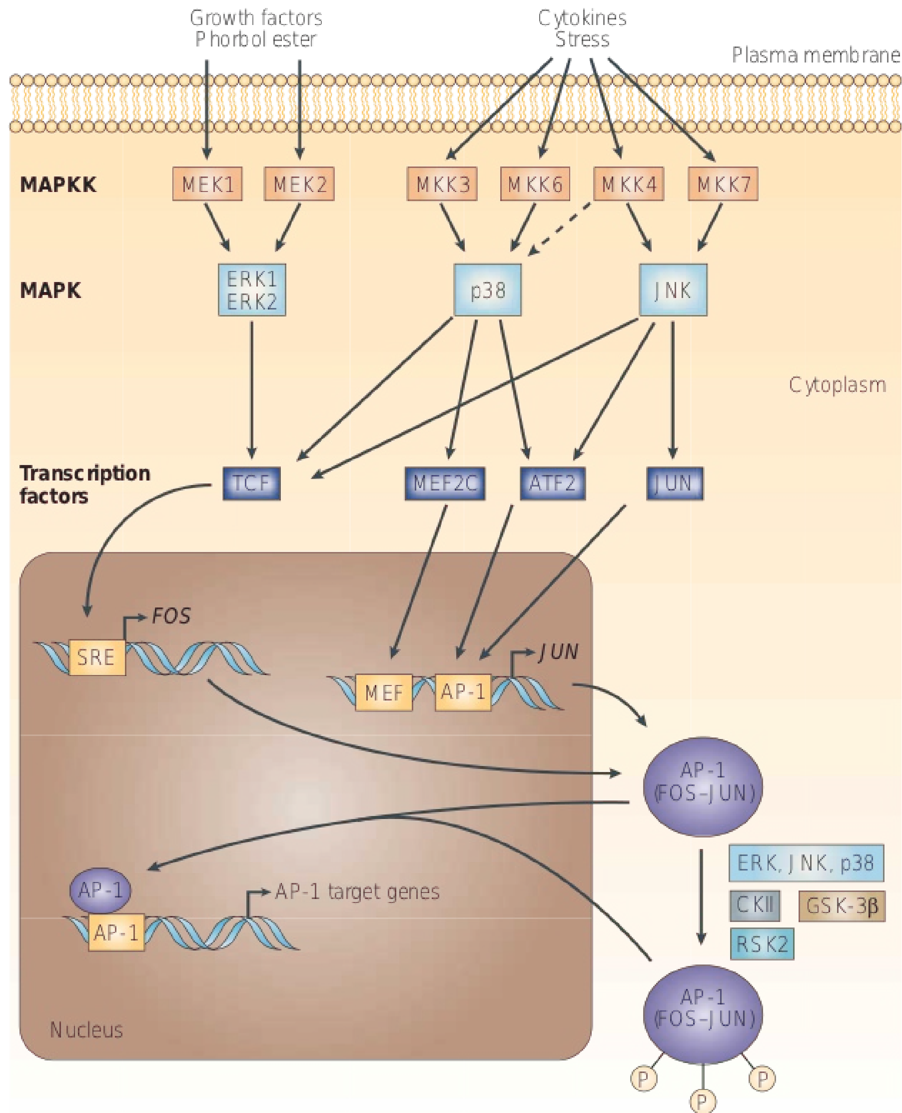


Figure 1.11: Transcriptional and post-translational activation of AP-1. The activity of AP-1 proteins is stimulated by a network of signaling pathways stating from external signals (likewise growth factors) and mitogen-activated protein kinases (MAPKs), p38 and Jun amino-terminal kinase (JNK) families. MAPKs activation generates a cascade of events leading to the recruitment of several transcription factors such as myocyte-enhancer factor 2C (MEF2C), activating transcription factor 2 (ATF2) and Jun ending with the transcription of Fos and Jun genes which activates AP-1 target genes. Post-translational modifications, as phosphorylation, modulates AP-1 activity, resulting in a different transactivating potential, DNA-binding capacity and stability of AP-1 components. Adapted from Eferl et al. Nature Reviews 2003

c-Jun knockout studies revealed an essential role for this protein in mouse embryonal development [153]. In fact c-jun$^{-/-}$ embryos die at mid-to-late gestation due to massive liver hemorrhage and extensive apoptosis of both hematopoietic cells and hepatoblasts [115]. In addition, c-Jun deficient fetuses present malformations in the outflow tract [115]. Furthermore, absence of c-Jun results in impaired formation of the intervertebral disc and in increased apoptosis of notochordal cells [154]. However, knock-in experiments revealed that replacement of c-Jun by JunB wipes out deregulation effects on expression of cyclin D1, p53 and p21^{Cip1} and proliferation defects both *in vitro* and *in vivo* [155]. In the latter case, this substitution can rescue both liver and cardiac defects in a dose dependent manner.

The role of AP-1 in cell fate depends on the cellular context and on the upstream survival or death stimulus, thus resulting in opposite responses as to be pro- and anti-apoptotic. c-Fos was reported to be expressed in cells in which apoptosis naturally occurs due to terminal differentiation [156]. Over-expression of c-Fos led to apoptosis in immature lymphocytes as well as in hepatocytes and in a myeloid leukemia cell line [157]. c-Fos$^{-/-}$ p53$^{-/-}$ double knock out mice develop severe rhabdomyosarcomas; however, re-expression of c-Fos enhances apoptosis [158]. According to its double role, c-Fos has been found not to be essential for apoptosis *in vivo*, since apoptosis also occurs in c-Fos-deficient mice. Furthermore, c-Fos overexpression negatively correlates with increased neuronal cell death after excessive stimulation with kainic acid experiments in diverse brain areas [159, 160]. It has also to be underlined that c-Fos abrogates the c-Jun mediated enhanced transcription of FasL gene, which codes for a well known apoptotic trans-membrane protein [161]. The same intriguing duality characterizes c-Jun. In fact, from one side c-Jun protects neurons against apoptosis induced by withdrawal of neuron growth factor (NGF), whilst if over-expressed, c-Jun induces apoptosis in neurons as well as fibroblasts [162, 163]. In addition, the presence of phosphorylation site mutants of c-Jun (Figure 1.10) confers apoptosis resistance after excessive stimulation with kainic acid [164]. Additionally, this protein prevents apoptosis during mouse hepatogenesis [115, 164] and its absence leads primary embryonic fibroblasts to be more susceptible

to UV-induced cell death [133]. Two pro-apoptotic genes such as FasL and TNF-α contain AP-1 binding sites, as targets of c-Jun [114], as well as p53 does [150, 165].

Despite all these aspects, AP-1 is mostly known to represent a family of proteins which are causally involved, and often expressed, in many highly proliferating tumors. It has to be underlined that involvement of AP-1 proteins in human cancer derives from multi-factorial effects, at the first level the ability to regulate cellular proliferation and survival, as already discussed. Some AP-1 target genes have a role in processes which require degradation/manipulation of extracellular matrix, namely angiogenesis and tumor metastasis. Transfection of a DNAzyme targeting the c-Jun mRNA in endothelial cells was shown to inhibit the capacity to form new blood vessels both *in vitro* and *in vivo*. Decrease of c-Jun is directly correlated with decrease of metalloproteinase-2 (MMP-2) which is produced by endothelial cells and is critical for extracellular remodeling [166]. c-Fos and Fra-1 instead have been shown to regulate MMP1 and MMP3 [167]. In addition, the vascular endothelial growth factor D (VEGFD) is a c-Fos target gene [168], while hypoxia induced factor 1 (HIF-1) is activated by c-Jun [169] and both c-Jun and JunB activate the angiogenic factor proliferin [170]. A detailed list of the AP-1 regulated genes in tumor development is provided in Table 1

Gene product	Activity	Main regulator	
DNMT1	DNA methylation	c-Fos	(upregulates)
EGFR	Stimulates proliferation	c-Jun	(upregulates)
		JunB	(upregulates)
HB-EGF	Stimulates proliferation	c-Jun	(upregulates)
GM-CSF	Stimulates proliferation	c-Jun	(upregulates)
		JunB	(downregulates)
KGF	Stimulates proliferation	c-Jun	(upregulates)
		JunB	(downregulates)
Cyclin D1	Stimulates proliferation	c-Jun	(upregulates)
		JunB	(downregulates)
WAF1	Inhibits proliferation	c-Jun	(downregulates)
p53	Inhibits proliferation	c-Jun	(downregulates)
	Stimulates apoptosis		
ARF	Inhibits proliferation	JunD	(downregulates)
	Stimulates apoptosis		
INK4A	Inhibits proliferation	c-Jun	(downregulates)
	Stimulates apoptosis	JunB	(upregulates)
FASL	Stimulates apoptosis	c-Jun	(upregulates)
		c-Fos	(upregulates)
FAS	Stimulates apoptosis	c-Jun	(downregulates)
BIM	Stimulates apoptosis	c-Jun	(upregulates)
BCL2	Inhibits apoptosis	JunB	(downregulates)
BCL-XL	Inhibits apoptosis	JunB	(downregulates)
BCL3	Inhibits apoptosis	c-Jun	(upregulates)
VEGFD	Angiogenesis	c-Fos	(upregulates)
uPA	Angiogenesis	FRA1	(upregulates)
uPAR	Angiogenesis	FRA1	(upregulates)
Proliferin	Angiogenesis	c-Jun	(upregulates)
		JunB	(upregulates)
MMP1	Invasiveness	c-Fos	(upregulates)
		FRA1	(upregulates)
MMP2	Invasiveness	c-Jun	(upregulates)
MMP3	Invasiveness	c-Fos	(upregulates)
		FRA1	(upregulates)
CD44	Invasiveness	c-Fos	(upregulates)
		c-Jun	(upregulates)
Cathepsin L	Invasiveness	c-Fos	(upregulates)
MTS1	Invasiveness	c-Fos	(upregulates)
KRP1	Invasiveness	c-Fos	(upregulates)
TSC36/FRP	Invasiveness	c-Fos	(upregulates)
Ezrin	Invasiveness	c-Fos	(upregulates)
Tropomyosin 3	Invasiveness	c-Fos	(upregulates)
Tropomyosin 5b	Invasiveness	c-Fos	(upregulates)

Table 1.1: Different AP-1 genes regulation in tumor development and suppression. AP-1 moieties exert different regulation activity towards genes involved in neoplastic progression; in fact Jun proteins mostly regulate genes that are involved in cell proliferation and apoptosis, whilst Fos proteins are required for angiogenesis and tumor invasion by malignant tumors. Adapted from Eferl et al. Nature Review 2003

Chapter

2

Aim of the work

2.1 AP-1 and DNA replication

To date, the mechanisms by which DNA sequences are selected to become replication origins in metazoan (including, human) genomes are still not completely understood (see paragraph 1.1.5). The findings that: i) AP-1 proteins are required for re-entering the cell cycle after starvation; ii) constitute a control point for G1 progression and are required for initiation of DNA synthesis in response to serum [171]; iii) enhance replication of the polyomavirus genome in the presence of low amounts of the fundamental protein for viral replication, namely large T antigen (LT) [172], makes these proteins interesting candidates for a role in origin specification. An involvement

in the regulation of origin activation of these proteins would not be surprising, in light of their ability to influence cell cycle and proto-oncogenic activity (see paragraph 1.2.3). Nevertheless, this consideration is reinforced by the presence of transcription factor binding sites close to origin sequences (see paragraph 1.1.3) and by the assumption that AP-1 binding sites could represent one element for origins specification (see paragraph 1.1.4) [102]. Thus, a number of questions concerning the possible general role played by these proteins in the context of DNA replication regulation can be raised: does the affinity for origin sequences of AP-1 proteins correspond to an actual origin binding in vivo? Does origin binding by AP-1 proteins correlate temporally and spatially with the sites of DNA synthesis, in the context of cell cycle progression? Can AP-1 be considered as part of the multi-protein complex responsible for the DNA replication, at the origin?

The aim of this work is to try to answer these questions. To do so, I concentrated on the most important moieties of the AP-1 family, human c-Fos and c-Jun.

Chapter

3

Results

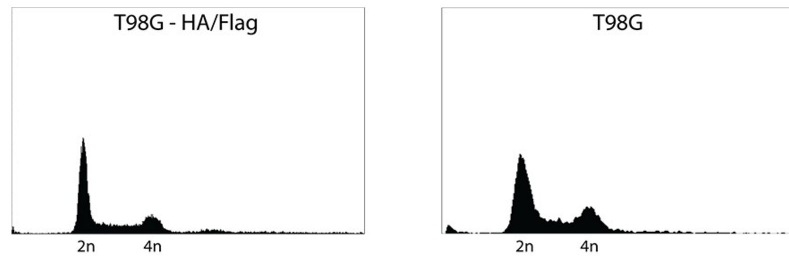
3.1 Characterization of T98G cell clone stably expressing HA/Flag HOXC13

The starting point of this work was the recent discovery of the involvement of transcription factors containing a homeobox domain, namely Homeotic proteins (HOX), in the initiation of DNA replication. Several components of this family were shown to be involved in this process by binding origin sequences in a one hybrid system or interacting with known replication factors in pull-down experiments. These include HOXA7, HOXA10, HOXA11, HOXA13, HOXB4, HOXB7, HOXC8, HOXC9, HOXC10, HOXC13, HOXD10, HOXD11 and HOXD13 [19, 96, 97, 173]

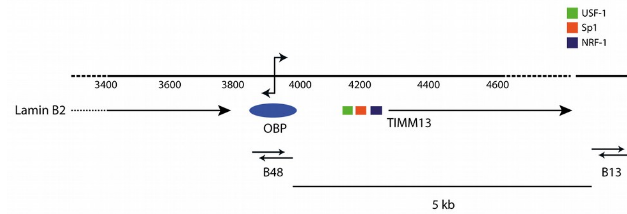
[70, 174, 175]. Linkage between Homeotic and AP-1 proteins came from the discovery that the latter are mediators of proliferative effects induced by HOX proteins [176, 177].

To further investigate this interaction, we took the advantage of having a T98G (glioblastoma ATCC CRL-1690) cell clone which stably overexpress a HOXC13 protein carrying a double tag, HA and FLAG, at its N-terminus, which we called HOXC13-HF. We decided to use this peculiar cell line in order to compare all the results to the ones previously published obtained using the wild type cell line [96, 97]. We first proceeded to evaluate the cell cycle profile, functionality of the LaminB2 origin of DNA replication and presence of the double tagged HOXC13 at the origin start site. Concerning the cell cycle, we did not notice any deregulation, compared to the wild type cell line, in asynchronous growing cells by FACS analysis after propidium iodide (PI) incorporation (Figure 3.1A). We then analyzed the activation of the LaminB2 origin by quantifying the levels of nascent DNA by competitive PCR [178]. We found a reproducible enrichment of a DNA region encompassing Lamin B2 start site (B48) over a control region located 5 Kbps downstream from it (B13) [178] (Figure 3.1B). This confirmed that this replication origin is active and that the presence of HOXC13-HA/Flag does not affect its functionality (Figure 3.1C).

A.



B.



C.

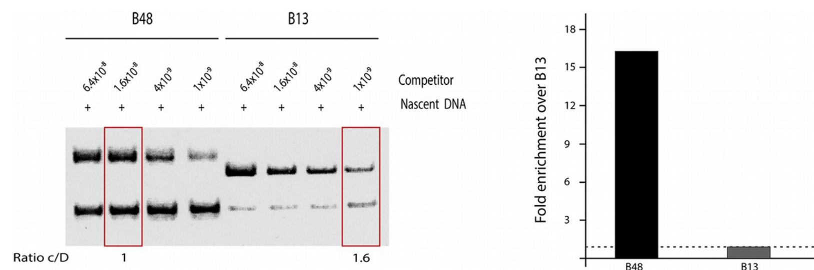


Figure 3.1: Flow cytometry and nascent DNA analyses on the T98G-HA/Flag stable clone. A. Comparison between the T98G HA/Flag stable clone and the wild type cell line performed with propidium iodide incorporation by flow cytometry analysis, showed no apparent cell cycle deregulation. B. Schematic representation of the position of origin (B48) and non-origin (B13) regions. The TIMM13 gene promoter region is located 250 bps downstream the LaminB2 start site and contains the binding sites for USF-1 (green box), SP-1 (orange box) and NRF-1 (violet box) transcription factors. C. The relative abundance of the sequences corresponding to the LaminB2 origin of DNA replication and to the B13 non-origin control was evaluated by competitive PCR probing a constant amount of the nascent DNA purified from the T98G HA/Flag clone (indicated as D), with sets of four serial 4-fold dilutions of a B48/B13 competitor plasmid (indicated as c) and using specific primers for the two sequences, as published [178]. The histograms report the relative enrichment of the B48 sequence over B13 as derived from the analysis of the PCR reactions shown in parallel.

Next we wanted to assess whether the position of the HOXC13 HA/Flag protein was also maintained onto the origin region as for the endogenous protein. For this purpose, we applied a chromatin immuno-precipitation (ChIP) approach using an anti-HA specific antibody; after DNA purification, competitive PCR was performed to estimate again the relative enrichment of B48 towards B13. As shown in figure 3.2, competitive PCR performed on DNA deriving from chromatin immuno-precipitated using an anti-HA specific antibody in the T98G HA/Flag clone resulted in an enrichment for the origin start site region (B48) in comparison to the control region, in agreement with previously published data for the endogenous protein [96].

As already mentioned in section 1.1.4, origin topology represents one of the main elements regulating origin selection and activation. Thus, alterations induced by HDAC class I and II or topoisomerase I and II inhibitors causes loss of binding of replication factors as previously demonstrated for both Cdc6 and HOXC13 [97].

Etoposide (VP16) and camptothecin (CPT) cause a non-reversible block of topoisomerases II and I on DNA, respectively, forming the so-called “cleavage complex” [87]; for this reason, these compounds are considered poisons as opposed to the reversible inhibitors, which are considered drugs. Thus, to detect possible modifications or alterations at the LaminB2 origin topology due to the presence of HOXC13-HA/Flag, *in vivo* Topoisomerase I and II interaction analyses by both ligation mediated (LM) [179] and terminal-transferase domain (TD) [180] PCRs were performed comparing HeLa, wild type T98G and the T98G HA/Flag clone. As shown in figure 3.3, the cleavage sites of both Topoisomerases were the same in each cell line used, leading us to conclude that the presence of the double tagged protein does not alter LaminB2 origin topology.

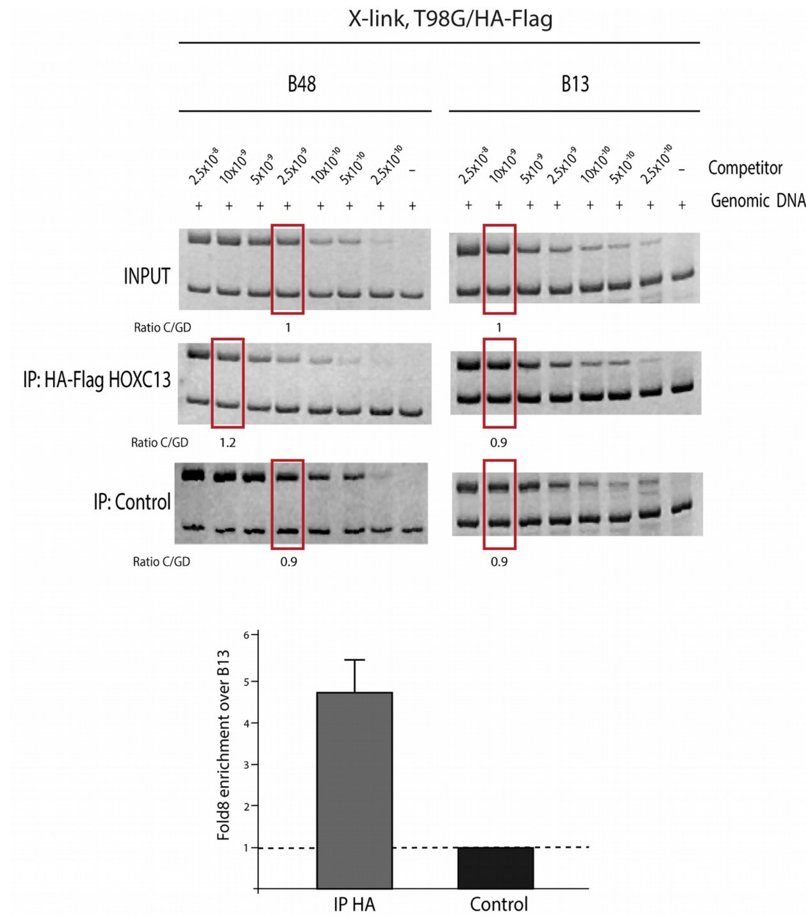


Figure 3.2: X-linked HA ChIP analysis on T98G HA/Flag stable clone. Competitive PCR results, aimed at assessing the relative enrichment of the origin (B48) in comparison to non origin DNA (B13), in chromatin immunoprecipitated using an anti HA antibody, showed that HA/FLAG-HOXC13 leads to an enrichment for the LaminB2 start site region in the T98G HA/Flag stable cells clone, as previously found for the endogenous HOXC13 [96]. These data, together with the normal flow cytometry profile, the LaminB2 origin activity and the enrichment for origin DNA by HOXC13-HF ChIP, suggest that the stable clone expressing double tagged HOXC13 behaves similarly to its wild type counterpart and that it can be used as a useful tool to detect proteins interacting with HOXC13 on the LaminB2 replication origin.

3.2 HOXC13-HA/Flag Co-immunoprecipitation analysis

We then wanted to identify the potential interacting partners of HOXC13 using the cell clone stably expressing the protein. We performed chromatin immunoprecipitation experiments using a specific anti HA antibody to analyze the material deriving from HA-ChIP (X-linked ChIP), as described in paragraph 5.4.1 of the Materials and Methods chapter. Briefly, at the end of the ChIP procedure, one third of the immuno-precipitated DNA:protein complexes was treated in order to retrieve DNA for competitive PCR, while the remaining material was boiled in Laemli buffer and the resulting proteins analyzed by western blotting. As a positive control, we used whole cell lysate (WCL) of both wild type T98G and the T98G HA/Flag cell clone before immunoprecipitation, whereas the material coming from anti-HA ChIP using wild type cells was used as a negative control.

As reported in figure 3.4, the material derived from ChIP using a specific anti HA antibody under native conditions (Native ChIP or NchIP) was also investigated; the results obtained using the anti-HOXC13 and anti-FLAG antibodies were consistent with those detected by X-linked ChIP, thus confirming the reliability of the latter procedure. A positive result without the usage of crosslinking agents, such as formaldehyde, obtained in NChIP corresponds to very strong protein:protein and protein:DNA interactions [181]. It is worth mentioning that the results obtained in native conditions, here used as a control, confirm already published results [97].

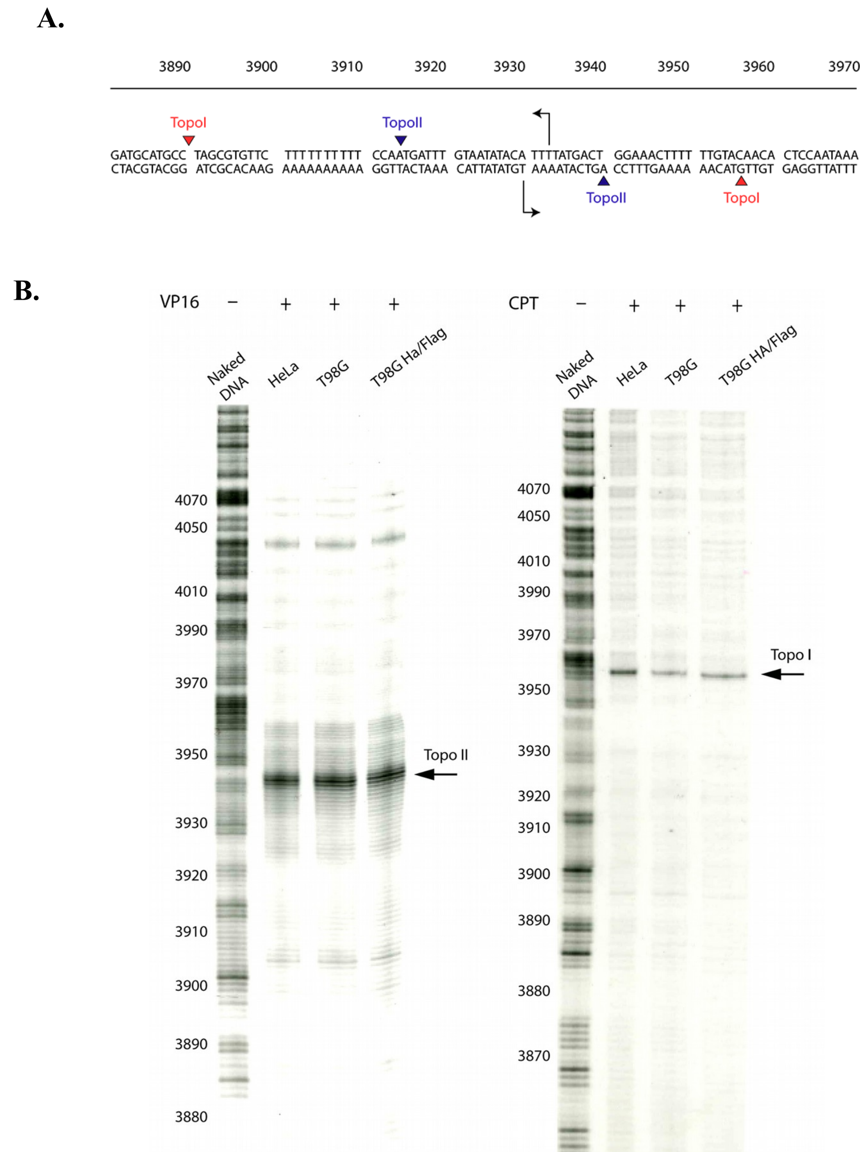


Figure 3.3: *in vivo* cleavage analyses of Topoisomerase I and II. A. Summary of topo I and II *in vivo* cleavage sites at the lamin B2 origin area involved in the replicative complexes as previously reported [87]. B. LM (Topoisomerase I) and TD (Topoisomerase II) PCR analyses were performed to detect the two enzymes cleavage sites, across LaminB2 origin region after incubation with etoposide (left side) and campotecin (right side) respectively. The cleavage sites obtained comparing T98G HA/Flag stable clone with HeLa and T98G cell lines, are the same confirming that the LaminB2 origin topology is not altered by the presence of the HOXC13 HA/Flag protein.

The presence of ORC4, a component of the heteroexameric ORC complex (a hallmark of the replication origins) [182], as a partner of HOXC13 HA/Flag, is consistent with the observation that this protein has role in DNA replication (paragraph 3.1). The presence of USF-1, a transcription factor known to have a binding site within the TIMM13 promoter region located ~250 bps downstream the replicative complex, was also investigated [98]. USF-1 is known to bind its target DNA region throughout the cell cycle, and was thus included as a positive control in the LaminB2 pre-RC ChIP analyses [96, 97]. Also this factor was found to interact with HOXC13. This observation is more intriguing, since the material used in our analysis had an average DNA fragment size of 150-200 bps, thus significantly shorter than the distance between the established USF-1 and HOXC13 binding sites. This observation is consistent with the conclusion that the two proteins might directly interact, as it will be further discussed later.

We next assayed the presence of the AP-1 proteins c-Fos and c-Jun and indeed detected an interaction between these two proteins and HOXC13 HA/Flag.

Taken together, these results show that c-Fos and c-Jun interact with a recent discovered DNA replication protein, namely HOXC13, and that this interaction also involves the pre-RC protein ORC4 and USF-1. These interactions are sufficiently strong to be detected also in the absence of any cross-linking agent.

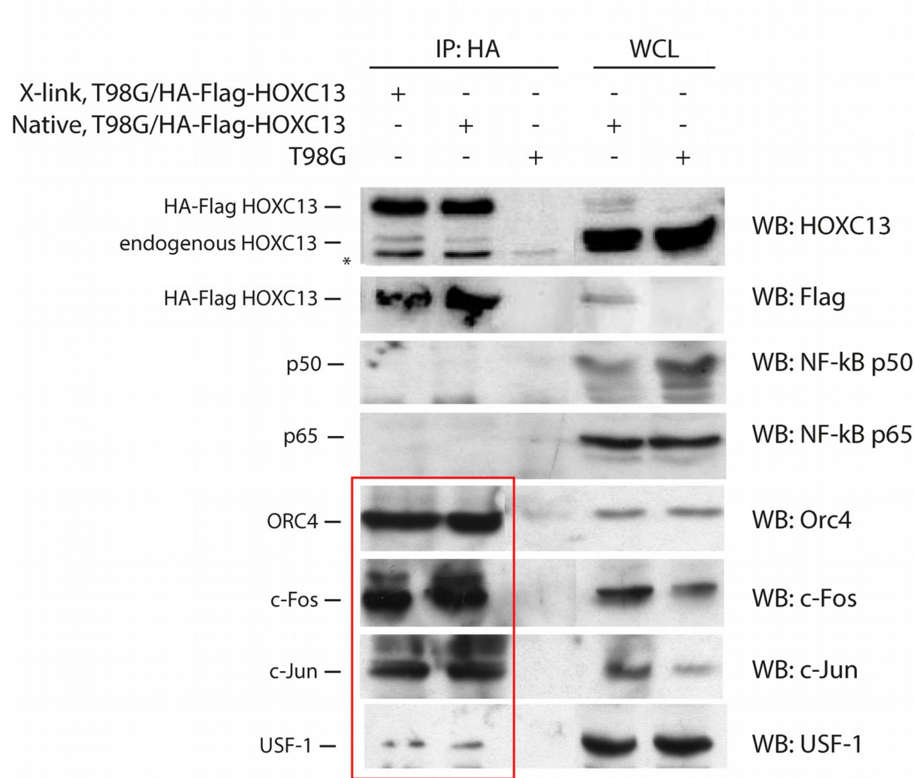
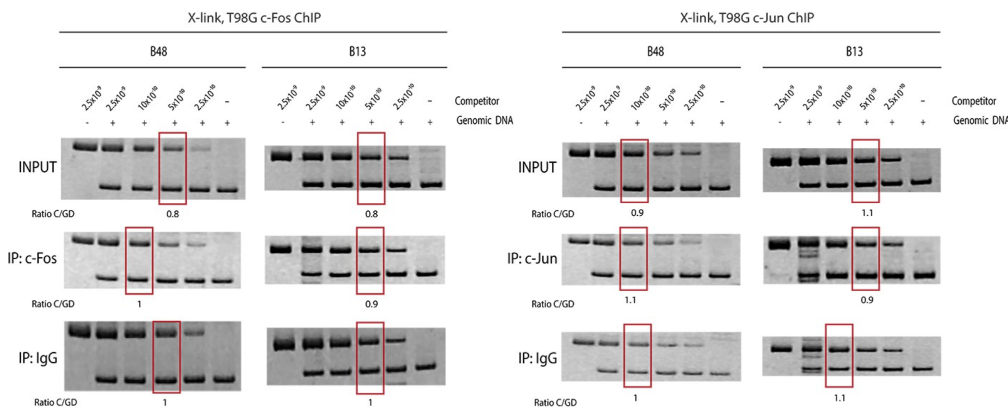


Figure 3.4: Co-IP partners of HOXC13-HA after X-linked ChIP and Native ChIP. Material deriving from chromatin immuno-precipitation using a specific anti HA antibody, under both cross-linked and native conditions, was investigated by western blot to search for possible HOXC13 HA/Flag partners (red box). An antibody against HOXC13 revealed the presence of both endogenous (lower band) and tagged (upper band) proteins. The asterisk indicates a non specific signal detected by the anti HOXC13 antibody. The AP-1 proteins c-Fos and c-Jun, and ORC4 were successfully immuno-precipitated with HOXC13 HA/Flag in the presence or absence of formaldehyde cross-link. USF-1 was also detected as a possible partner of the same multi-protein complex. ChIP performed using a specific anti HA antibody in wild type T98G cells was used as negative control. Antibodies against p50 and p65 NF-kB subunits were used as negative controls, whereas cellular lysates of both cell lines were used as positive controls.

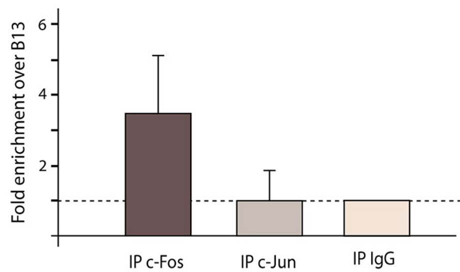
3.3 Analysis of AP-1 binding onto the LaminB2 origin

To understand whether the protein:protein interaction between AP-1 proteins and HOXC13 HA/Flag occurs on the LaminB2 origin of DNA replication, we performed chromatin immunoprecipitations (X-linked ChIP) using specific anti c-Fos and anti c-Jun antibodies. The DNA:protein immunocomplexes were isolated and, after DNA purification, competitive PCR analyzing the relative enrichment of the origin region with respect to B13 was performed [178].

A.



B.



C.

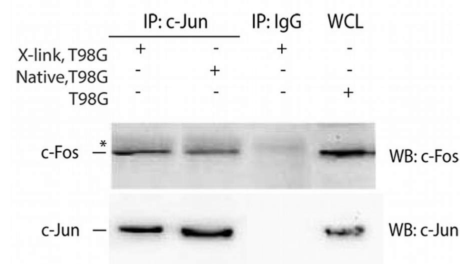


Figure 3.5: ChIP analysis for AP-1 proteins on the LaminB2 origin. To investigate the presence of AP-1 on the LaminB2 origin, X-linked ChIP using anti c-Fos and c-Jun specific antibodies was performed and the resulting DNA was subjected to competitive PCR analysis. A. X-linked ChIP analyses allowed to detect the interaction of c-Fos and not of c-Jun to the LaminB2 origin. B. Quantification of the results. Shown are the mean and s.d. of three independent experiments. C. Part of

the material deriving from c-Jun ChIP under both X-linked and native conditions, was analyzed in western blot indicating c-Fos as co-IP partner.

c-Fos, but not c-Jun, was found to interact with the B48 region (Figure 3.5A and 3.4B), in disagreement with the results obtained. We decided to analyze the material deriving from c-Jun ChIP to search for the presence of c-Fos by western blotting. The results indeed show that c-Fos is indeed as a partner of c-Jun, indicating that this interaction probably occurs somewhere else in the nucleus (Figure 3.5C).

3.4 High resolution analysis of the Lamin B2 origin of DNA replication

To investigate more precisely the binding site for c-Fos across the LaminB2 replication origin and to investigate on the negative results instead obtained for c-Jun, we decided to increase the resolution of the X-linked ChIP analysis. We focused on a larger DNA region of 1.1 Kbps, including the 150 bps B48 fragment and the TIMM13 promoter genomic areas. This region was divided it into 12 overlapping fragments, as shown in figure 3.6. Purified DNA deriving from chromatin immuno-precipitation was amplified by real time PCR to determine the relative abundance of each fragment. The distant B13 genomic region served as a control.

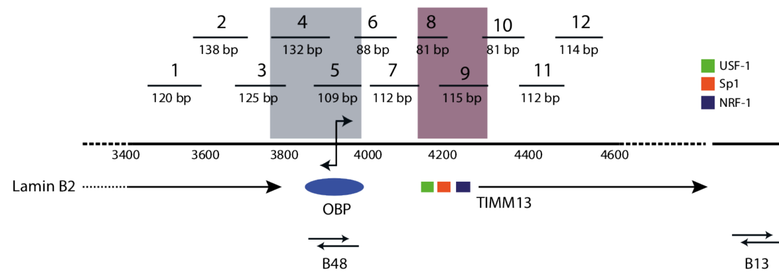


Figure 3.6: Schematic representation of the position of the 12 overlapping PCR segments. Scheme of the amplicons used for ChIP analysis, across the 1.1 kbp DNA region containing the start site of the Lamin B2 origin of DNA replication (highlighted in gray) and the TIMM13 promoter (purple) regions. B48 and B13 exemplify origin and the non-origin regions. Purified DNA coming from chromatin immuno-precipitation was hence assayed by real time PCR looking for the relative enrichment of each fragment compared to the B13 control. The DNA sequence covered by the origin binding complex corresponds to fragments 4 and 5; the TIMM13 promoter region to the fragments 8 and 9.

To the best of our knowledge no similar analyses have been reported to study the chromatin profile of a specific DNA region and to map protein positioning with an average resolution of 110 bp across that region. The choice of primers to be used was made by analyzing the 1.1 Kbp DNA region using Amplify 3 (freeware) and Vector NTI© (Invitrogen) software; each primer set was then empirically tested by real time PCR, as described in the next paragraphs. A primer list is reported in the Materials and Methods.

3.4.1 Validation of the high resolution ChIP technology

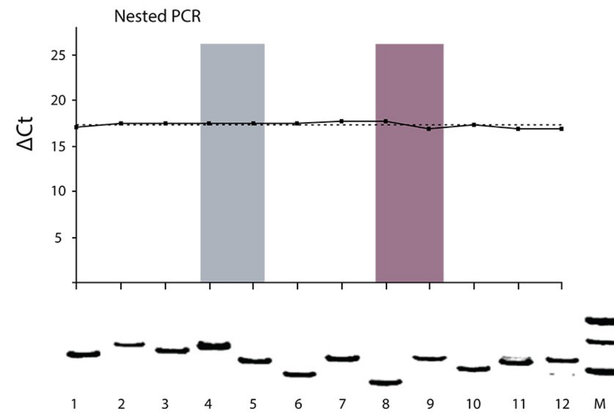
We first wanted to verify the reliability of our high resolution ChIP technology. We started by performing a series of PCRs to verify whether all the sets of primers were able to amplify their respective DNA target segments at equivalent efficiency. The

starting material for these real-time amplifications was a 1.1 kbp fragment obtained from genomic DNA using the 1L-12U primers. As shown in figure 3.7A, the ΔCt values for all the primer sets was comparable, as it was the amount of amplified DNA after polyacrylamide gel electrophoresis and ethidium bromide staining.

The same analysis was then repeated using genomic DNA extracted from starved T98G (G0). The minor deviations in the ΔCt values detected in this case (shown in figure 3.7B) were taken into account for the calculations reported in the subsequent experiments.

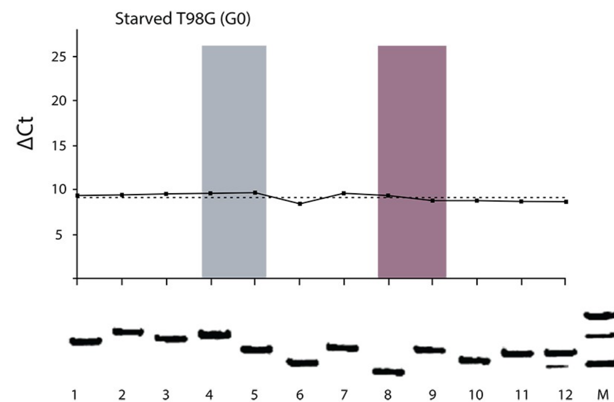
The amplification efficiency for each set of primers was calculated by using the standard curve method after performing real time PCR amplification of four scalar DNA amount (25 ng, 12.5 ng, 6.25 ng and 3.125 ng) from G0 cells; the mean of three independent experiments per DNA concentration and per fragment is plotted in figure 3.8. Correlation analysis indicated that the R^2 value was in a range between 0.94 and 0.99.

A.



Fragment no.	1	2	3	4	5	6	7	8	9	10	11	12
ΔCt Mean	17,81	17,37	17,40	17,34	17,42	17,46	17,16	17,12	18,20	17,57	18,02	18,29
Std Dev	0,15	0,16	0,14	0,18	0,12	0,10	0,31	0,34	0,43	0,02	0,30	0,49

B.



Fragment no.	1	2	3	4	5	6	7	8	9	10	11	12
ΔCt Mean	9.34	9.38	9.81	9.88	9.68	8.29	9.61	9.33	8.77	8.75	8.58	8.59
Std Dev	0.03	0.01	0.18	0.05	0.03	0.04	0.10	0.06	0.04	0.06	0.69	0.30

Figure 3.7: Primer analysis after amplification of a DNA amplicon obtained using the 1L-12U primers or total genomic DNA from starved (G0) T98G cells. Real time profile and ethidium bromide polyacrylamide gel stainings after PCR amplification for both pre-amplified 1L-12U DNA (A) and total genomic DNA from starved T98G cells (B). The ΔCt value corresponds to the difference between total PCR cycles and the Ct obtained. For each real time PCR graph, the dotted line represent the mean of the ΔCts. The LaminB2 origin and the TIMM13 promoter regions are highlighted in gray and purple respectively.

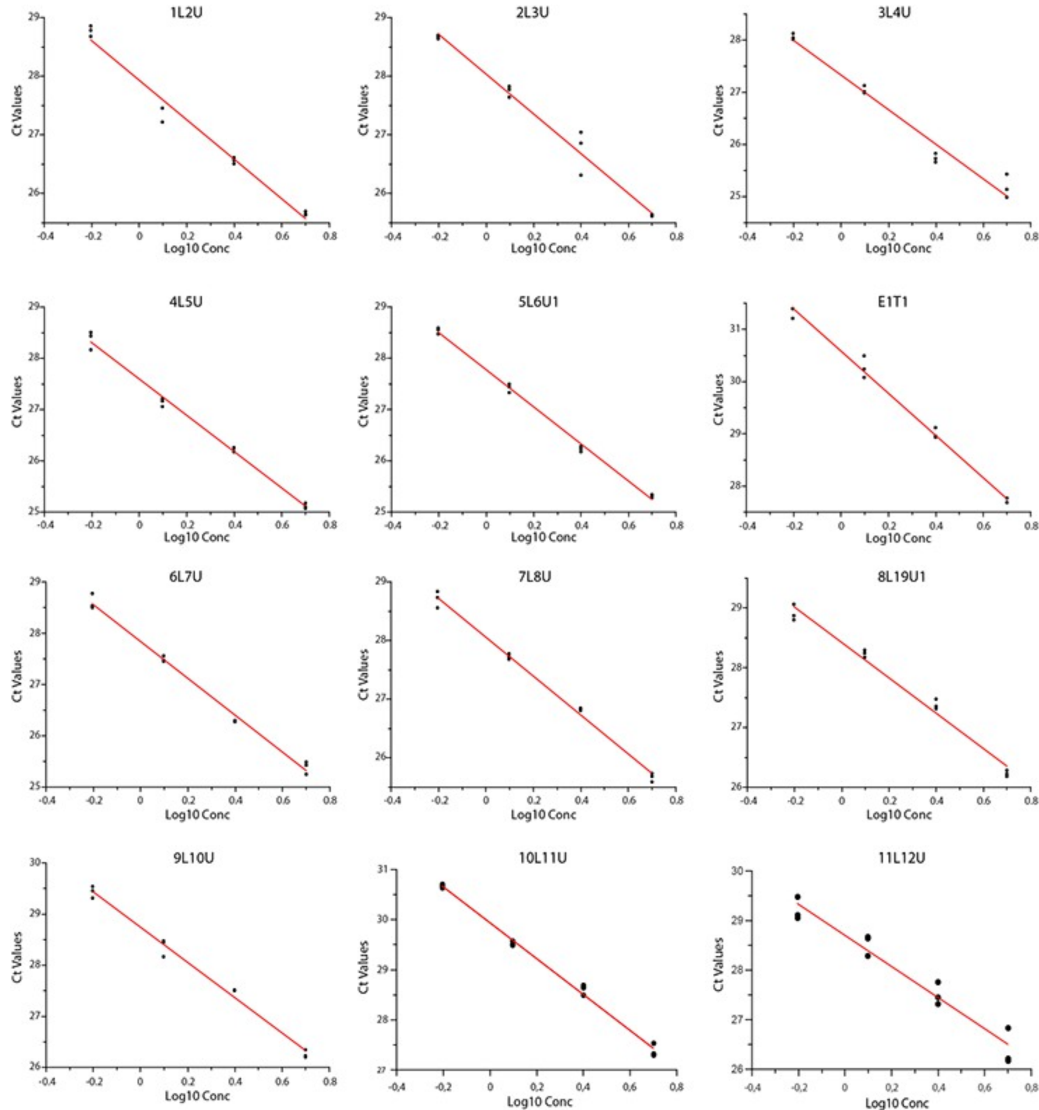


Figure 3.8: Primer efficiency analyses performed using G0 DNA. The means of four scalar DNA concentrations derived from three different G0 DNA extractions for each of the twelve fragments are reported. For each analysis, the slope of the fitting curve was used to define primers efficiency with the standard curve approach.

Finally, a similar analysis was applied for the amplification of chromatin obtained from asynchronous growing cells after digestion with micrococcal nuclease, which digests genomic DNA to mononucleosomal size, equivalent to the starting material for the subsequence high-resolution ChIP analyses. As shown in the real time graphs

in figure 3.9, samples obtained in both native and X-linked conditions revealed two sites that were more sensitive to the enzymatic digestion, namely the LaminB2 start site (corresponding to fragments 4 and 5) and the TIMM13 promoter region (corresponding to fragments 8 and 9) regions. This result is in agreement with the information that both replication origins and promoters are free from nucleosomes and hence less resistant to enzymatic digestion. It has to be underlined that, while sensitivity toward micrococcal nuclease remains similar in the TIMM13 region, it varies significantly at the origin region after formaldehhyde crosslink, as it is particularly evident from the analysis of the amount of amplified material detected after ethiudium bromide staining (bottom parts of figures 3.9A and 3.9B). Fragment 2 seemed also relatively protected in the presence of formaldehyde.

Taken together, these results set up the basis for the subsequent high resolution analyses of the protein:DNA interactions occurring at the Lamin B2 origin

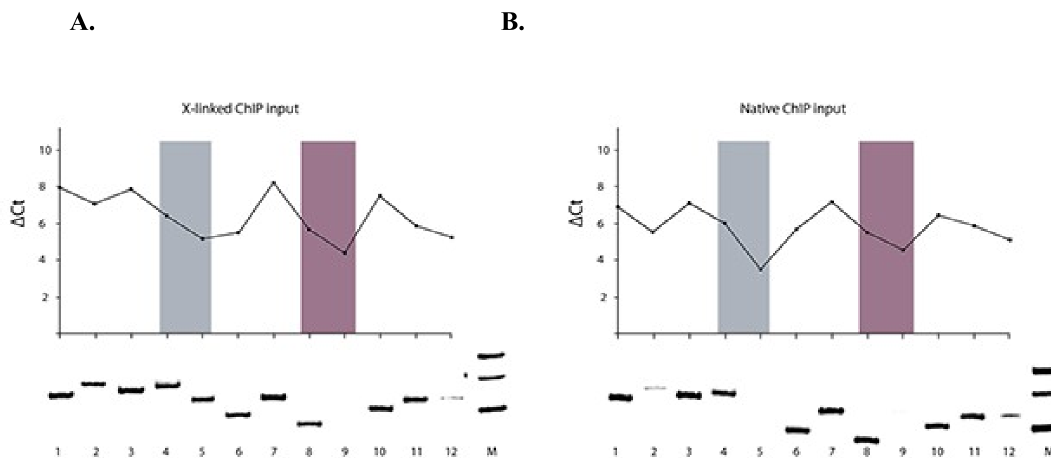


Figure 3.9: Input chromatin profile in the presence and absence of formaldehyde crosslink. Real time PCR profile (upper part) and ethidium bromide gel staining (lower) from PCR amplifications of both cross-linked (A) and not cross-linked (B) chromatin input materials. The DNA region encompassing the start site and hence the origin binding proteins (fragments 4 and 5 highlighted in gray) are more sensitive to micrococcal nuclease digestion under native conditions, likely due to the absence of protein:protein and DNA:protein crosslinks, whereas the TIMM13 promoter region (fragments 8 and 9, highlighted in purple) show a similar sensitivity towards enzymatic digestion in both conditions.

3.4.2 Analysis of the Lamin B2 origin of DNA replication by high resolution ChIP

To further validate our experimental approach, we moved to map the position of some known proteins across the DNA segment of interest in asynchronous growing T98G cells. Binding of H2B was reproducibly detected throughout the investigated area, with the exception of the start site region (fragments 4 and 5) and the nucleosome-free region of the TIMM13 promoter, which is known to associate with transcription factors USF-1, Sp1 and NRF-1 (fragments 8 and 9, figure 3.10 left panel). Histone acetylation levels was then analyzed by performing NChIP using a specific antibody against acetylated H3(K14). Binding of this protein was mostly detected immediately downstream the promoter region (fragments 10 and 11).

We next decided to apply high-resolution ChIP analysis to some proteins known to bind different portions of the origin area. We performed X-linked ChIP using specific antibodies against ORC4 as a marker of the start site region [182] and against USF-1 as a marker of the TIMM13 promoter region [98]; we also wanted to assay the position of HOXC13 to further confirm its interaction with the LaminB2 origin as previously published [96, 97] (Figure 3.11).

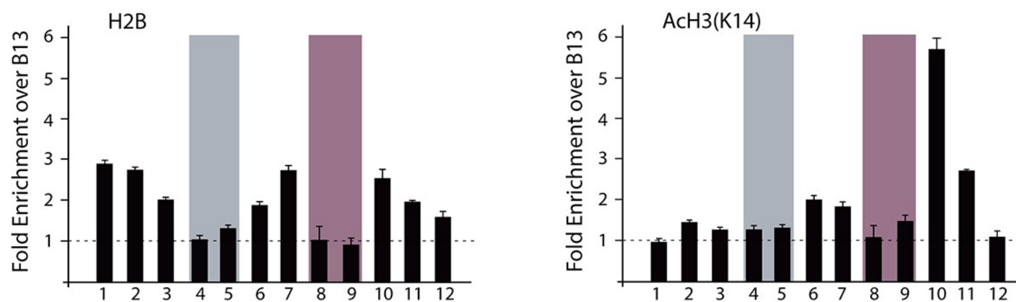


Figure 3.10: H2B and acetylated H3(K14) NChIP analysis. Real time PCR chromatin binding profiles of H2B (left) and lysine 14-acetylated H3 (right). LaminB2 origin and TIMM13 promoter region are highlighted in gray and purple respectively. Shown are the means and s.d. of at least three independent experiments.

ORC4 was enriched at fragment 5, encompassing the DNA replication start site while USF-1 was found to bind the TIMM13 gene promoter, as expected. Quite surprisingly, however, we also found a reproducible peak of enrichment for USF1 in correspondence to the origin region, encompassing fragments 4 and 5. HOXC13 was found to bind both the replication start site and the downstream promoter region. No binding to any of the investigated fragments was detected for the subunit 65 of the NF-kB protein, used here as a negative control.

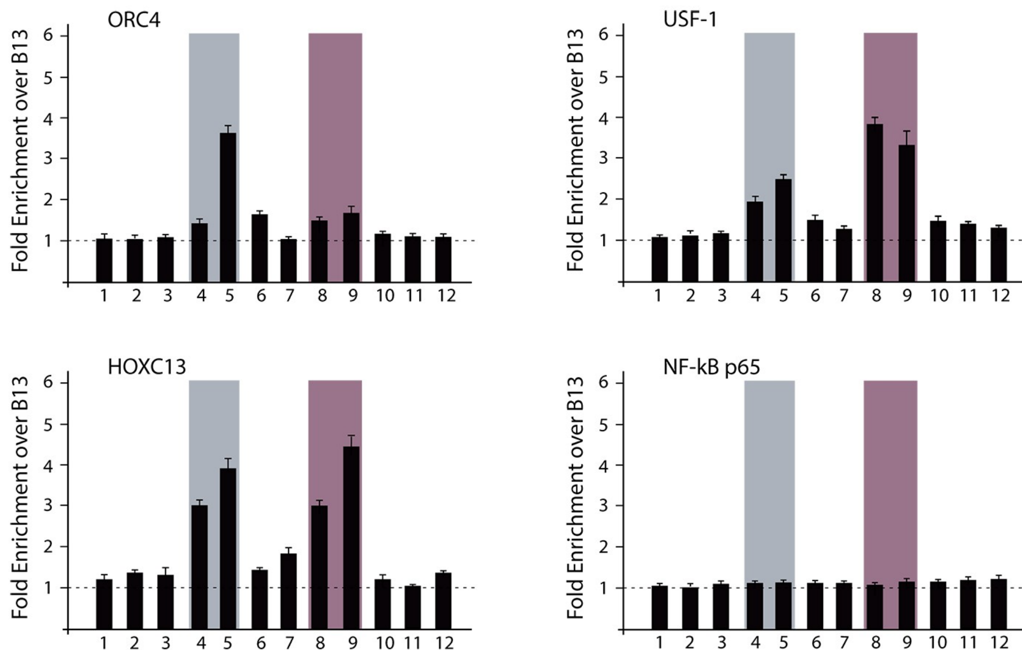


Figure 3.11: Pre-RC proteins binding profiles across LaminB2 origin. Real time PCR chromatin binding profiles of different proteins across the LaminB2 genomic region. The LaminB2 origin and TIMM13 promoter regions are highlighted in gray and purple respectively.

We therefore proceeded to analyze the binding position of the c-Fos and c-Jun proteins. As shown in figure 3.12, we performed this analysis in both T98G and HeLa cell line by performing a formaldehyde-mediated ChIP and, in T98G cells, also in native condition. We found a reproducible enrichment for c-Fos at the start site region (fragments 4 and 5), whereas c-Jun was found to bind the TIMM13 area (fragments 8

and 9). Similar results were also detected by X-ChIP in HeLa cells and by performing the experiments in the absence of crosslinking agents.

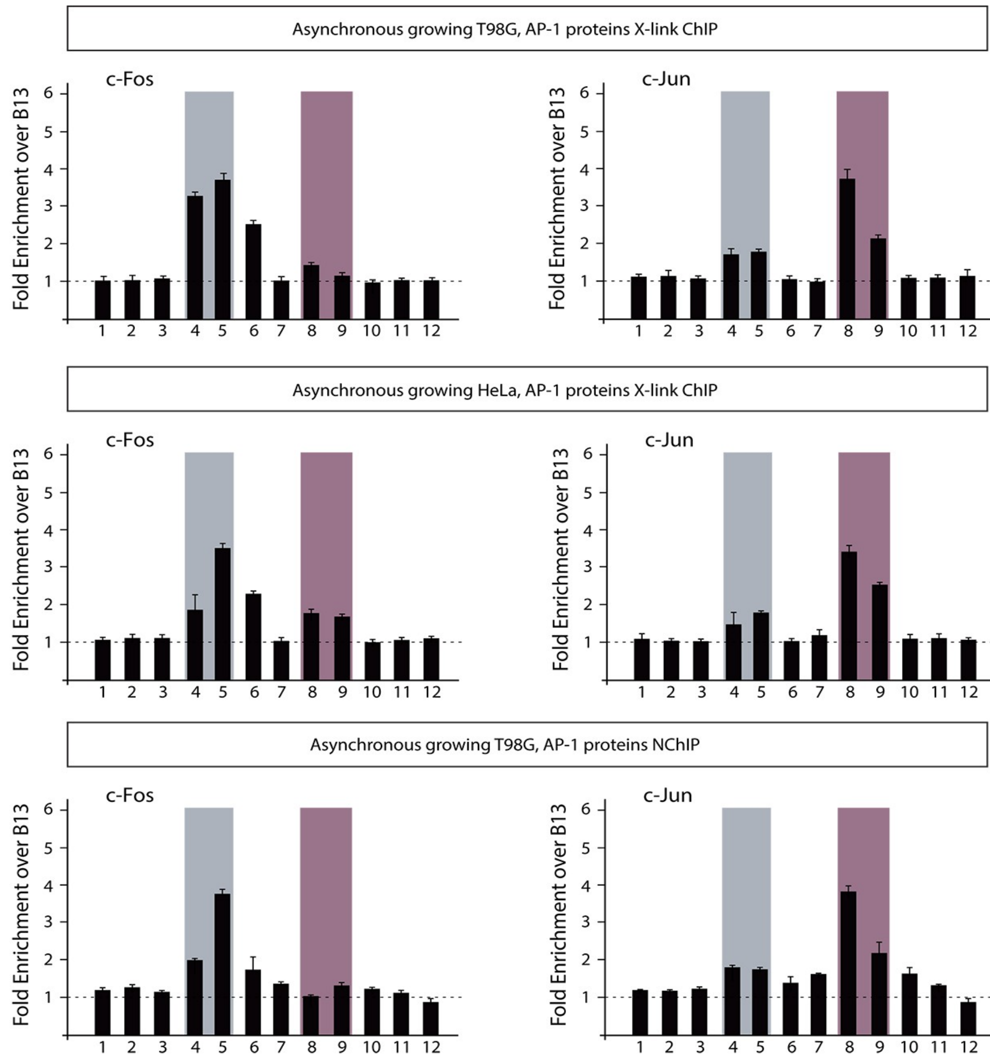


Figure 3.12: c-Fos and c-Jun ChIP analysis. Chromatin binding profiles of AP-1 proteins across the LaminB2 origin area in X-linked chromatin from T98G cells (upper graphs), in X-linked chromatin from HeLa cells (mid graphs) and in T98G chromatin under native conditions (lower graphs). The LaminB2 origin and TIMM13 promoter regions are highlighted in gray and purple respectively.

Taken these results together, the investigated proteins that bind the Lamin B2 origin region define two major areas, one corresponding to the actual start site for DNA

replication and the second to the TIMM13 promoter. Quite unexpectedly, interaction with USF1 and HOXC13 was detected for both regions, while the two AP-1 components selectively bound either the start site (c-Fos) or the promoter (c-Jun).

To detect other possible interactors with the origin region, we performed chromatin immuno-precipitation searching for the relative enrichment of the origin DNA using antibodies specifically recognizing other proteins (Figure 3.13). In particular, we investigated the possible presence of the Pre-B-cell leukemia transcription factors 1, 2 and 3 (better known as PBX1, PBX2 and PBX3) which are members of the three amino acids loop extension (TALE) protein family, known to interact with HOX proteins and to increase their DNA binding specificity [183]. Moreover, we also investigated the presence of another member of the Jun family, JunD, which also interacts with c-Fos [184]. Finally, we looked for the presence of the nuclear factor of activated T cells 1 or NFAT1 (known also as NFATc2 or NFATp), a transcription factor able to interact with the Fos:Jun heterodimers forming a ternary complex which stabilizes AP-1:DNA binding [185, 186]. None of these proteins was however found to bind with the investigated origin DNA region. These results also confirmed the specificity of the previously detected interactions.

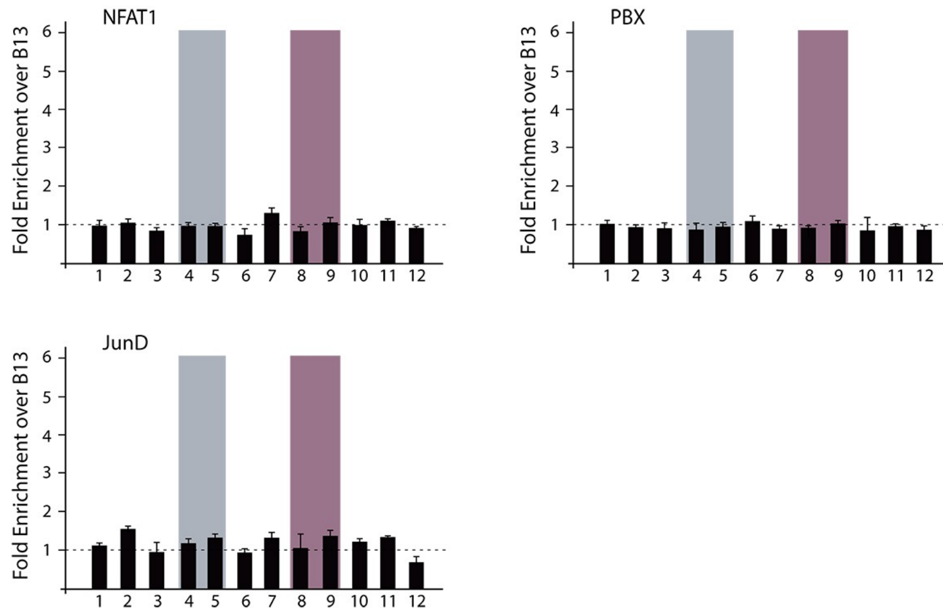


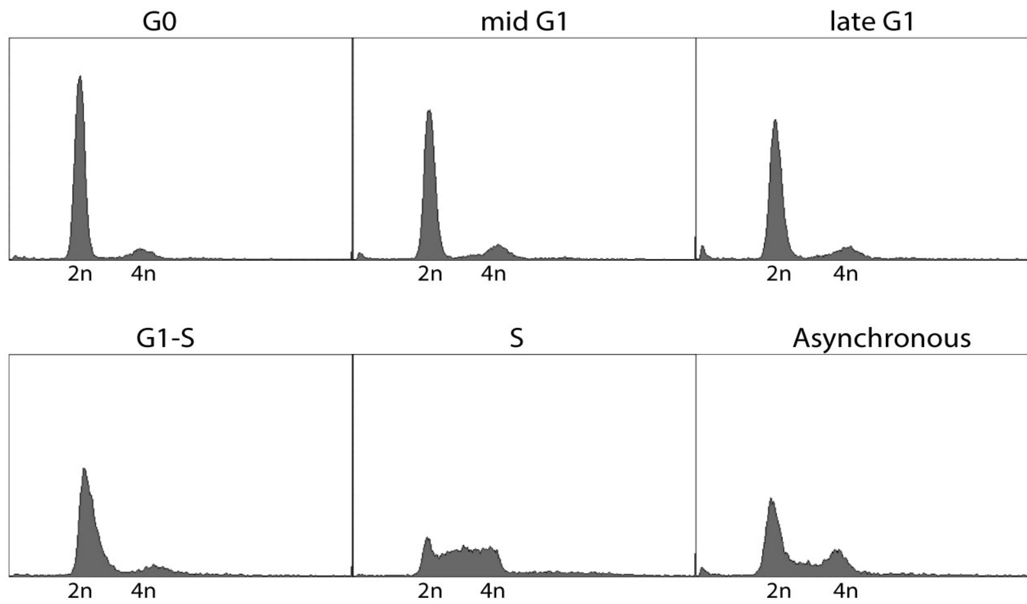
Figure 3.13: ChIP analysis to detect binding, across the LaminB2 origin, of other potential protein candidates. The investigated proteins included NFAT1, PBX1/2/3 and JunD, as indicated. The LaminB2 origin and TIMM13 promoter regions are highlighted in gray and purple respectively.

3.5 Spatial and temporal analysis of AP-1 binding onto the LaminB2 origin

To evaluate a possible correlation between the association of the AP-1 proteins, c-Fos and c-Jun, with the 1.1 Kbps DNA region analyzed containing both the LaminB2 replication origin and TIMM13 promoter, we performed chromatin immuno-precipitation analyses at different time point during the cell cycle. For this purpose, T98G cells were synchronized in G0 by culture in serum free medium for 72 hours; then, cells were allowed to re-enter the cell cycle by adding fresh complete medium. Since c-Fos and c-Jun are involved in cell proliferation [187, 188] and are required for entry into the S phase by interacting with cyclin D1 and cyclin E [149], we decided to follow the chromatin binding profile in the G0, mid G1, late G1, G1-S

border and S phases of the cell cycle, corresponding to 0, 9, 14, 16 and 20 hours after serum addiction, as already reported [189]. At each time point, propidium iodide incorporation was monitored by flow cytometry to confirm the synchronization of the cell stage.

A.



B.

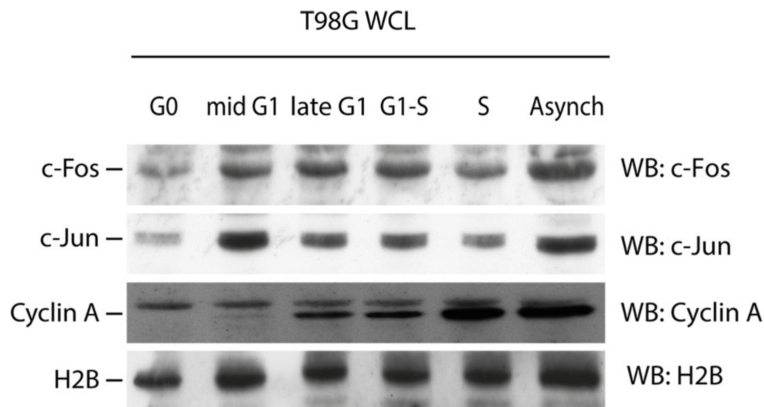


Fig. 3.14: Flow cytometry profile and expression levels of AP-1 proteins during the cell cycle. A. Flow cytometry analysis of the cellular DNA content during the phases of the cell cycle under examination (0, 9, 14, 16 h and 20 h for G0, mid G1, late G1, G1-S border and early S respectively) compared to asynchronous growing cells. B. Expression levels of c-Fos, c-Jun and Cyclin A during the analyzed cell cycle phases.

Cellular lysates also were analyzed by western blotting to detect c-Fos, c-Jun and cyclin A levels in the phases of the cell cycle under examination. Figure 3.14A shows the flow cytometry profiles of cellular DNA content at different times after serum addition. Analysis of the levels of expression of c-Fos and c-Jun at the same time points by western blotting indicated that both proteins were significantly induced in mid- and late G1 compared to starved cells (Figure 3.14B). The levels of Cyclin A, which start to be present in mid G1 and reaches its maximum level of expression during S phase, confirmed our synchronization.

Each AP-1 ChIP was confirmed by western blot analysis looking for the immuno-precipitated protein (Figure 3.15). As reported in figure 3.16, analyses performed with starved cells, in G0, demonstrate that neither c-Fos nor c-Jun are bound to the genomic DNA segment under investigation, even if the proteins were expressed in the cells, as concluded from the western blotting analysis. Similar results were obtained for the positive control ORC4 protein. These results are in agreement with the knowledge that, similar to other transcription factors linked to the cell cycle (such as E2F), AP-1 proteins are expressed during G0 as they are required for the cell cycle entry upon growth factors stimulation [188].

Several hours after serum re-addiction, during mid G1, both AP-1 proteins were found to bind the TIMM13 promoter region. Binding of ORC4 instead, was restricted to the LaminB2 origin start site area, as expected.

Few hours before DNA replication, when the pre-RC complex is loaded but not yet licensed, in the late G1 phase of the cell cycle, the AP-1 proteins chromatin binding profile became the same as observed in asynchronous growing cells. Hence, c-Jun was found to bind the TIMM13 promoter region (fragments 8 and 9), whereas c-Fos "shifted" from the promoter to bind the start site region (fragments 4 and 5).

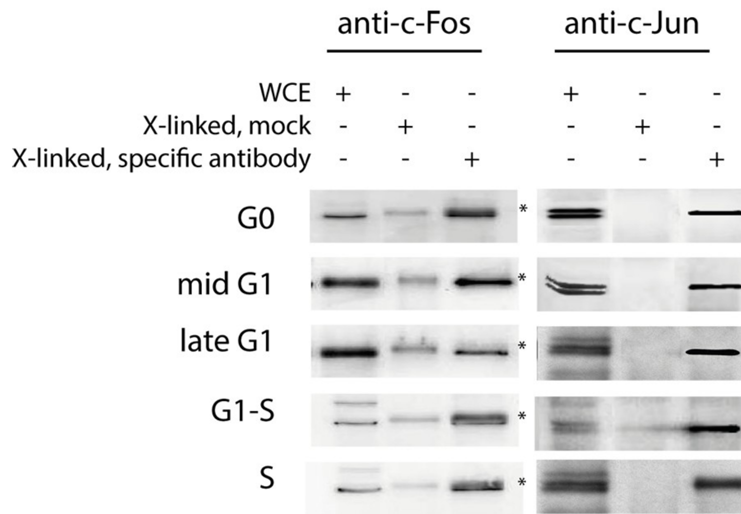


Fig. 3.15: Western Blot analyses of AP-1 proteins in different phases of the cell cycle. Part of the material coming from c-Fos and c-Jun ChIPs was analysed by western blotting using antibodies against the respective proteins. A T98G cells whole cell lysate (WCE) served as a positive control, while cross-linked extracts processed with pre-immune serum served as a negative control.

Binding of ORC4 was restricted at the start site region, as in asynchronous growing cells (fragments 4 and 5).

After the pre-RC licensing, when the pre-IC is assembling, during the G1-S transition, none of the two AP-1 proteins was found to interact with either the LaminB2 origin or the TIMM13 promoter regions; only ORC4 continued to be present at the origin region. Finally, in the early stages of the S phase, while ORC4 binding was still detectable, c-Fos disappeared from the whole analyzed region, while c-Jun continued to bind the TIMM13 promoter.

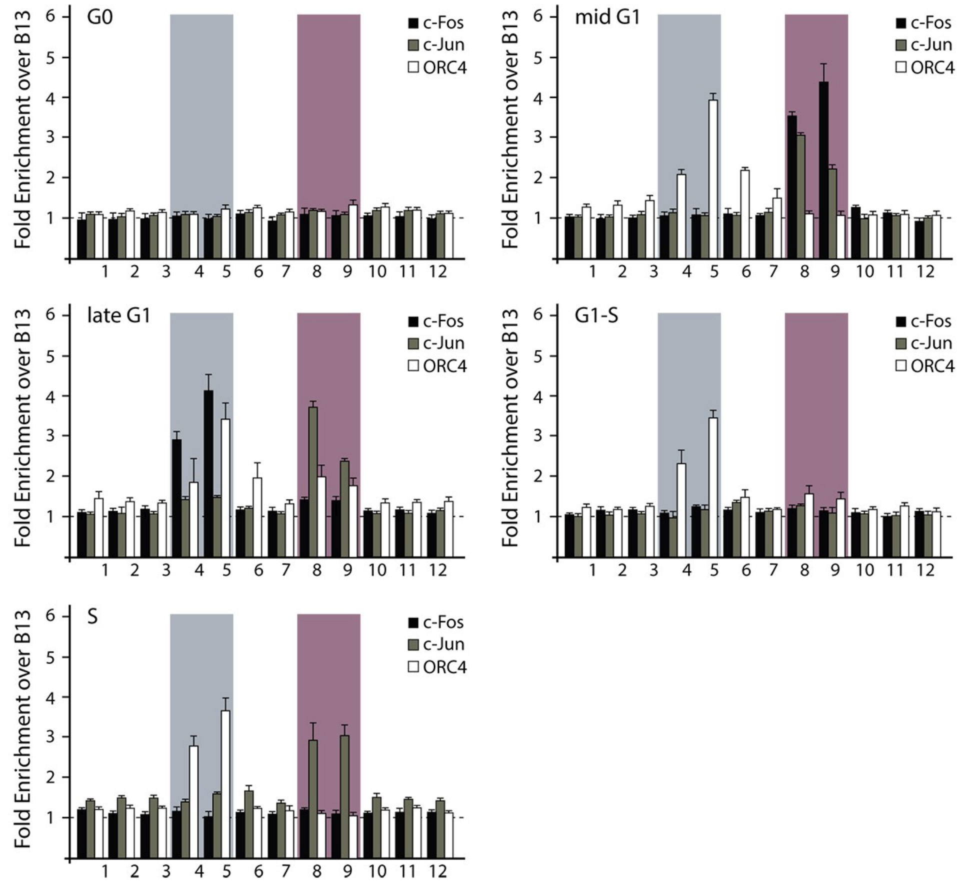


Fig. 3.16: ChIP analysis for AP-1 and ORC4 proteins through the cell cycle. Result of chromatin IP during the subsequent phases of the cell cycles using the indicated antibodies. The LaminB2 origin and TIMM13 promoter regions are highlighted in gray and purple respectively.

We then aimed to understand whether any possible protein:protein interaction between the two AP-1 proteins and ORC4 could be detected during all the previously investigated phases of the cell cycle. For this purpose, we decided to analyze part of the material deriving from the c-Fos and c-Jun chromatin immuno-precipitations by western blotting, looking for the presence of ORC4.

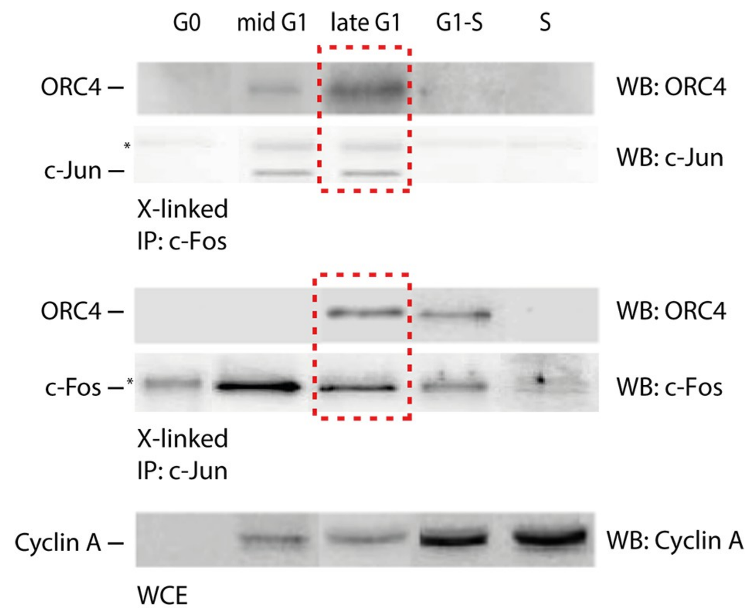


Fig. 3.17: Co-IP of ORC4 in AP-1-ChIP samples across the cell cycle. To evaluate the possible interplay of the two investigated AP-1 proteins and ORC4, part of the material deriving from the c-Fos and c-Jun ChIPs performed in different phases of the cell cycle was analyzed by western blotting using an anti-ORC4 antibody. The red boxes highlight the phase of the cell cycle in which all the three proteins are present together, namely late G1.

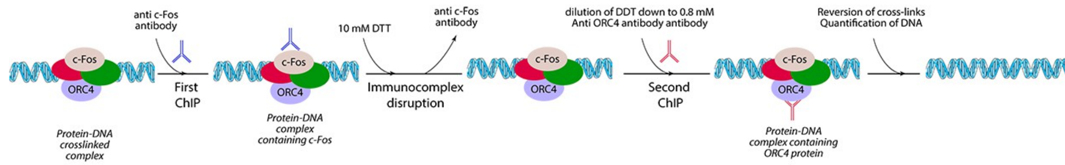
As reported in figure 3.17, we obtained a positive result for the presence of ORC4 in the c-Fos and c-Jun chromatin immunoprecipitates obtained in late G1. This result is consistent with the possibility that these three proteins are located at and interact each other and with the lamin B2 origin at this time point of the cell cycle. Additional interactions between specific AP-1 proteins and ORC4 were detected in other phases of the cell cycle, including mid G1 for c-Fos-ORC4 and G1-S border for c-Jun-ORC4.

To determine whether these interactions occur onto the 1.1 Kbps DNA region containing both the LaminB2 origin and the TIMM13 promoter regions, or elsewhere inside the nucleus, we applied the so called sequential chromatin immunoprecipitation analysis (Re-ChIP) protocol [190]. Briefly, Re-ChIP, which is applied to crosslinked, immunoprecipitated chromatin, allows one to detect the

presence of a specific protein in a complex formed at a chosen genomic region. In brief, crosslinked chromatin, which was immunoprecipitated with anti-c-Fos or c-Jun antibodies, was subjected to a mild DTT treatment and re-probed using an anti ORC4 antibody. Finally, the immunocomplex was treated to reverse the crosslink and DNA analyzed by real time qPCR, as schematically shown in figure 3.18A. As a control of the validity of the procedure, part of the material obtained from the Re-ChIP was recovered and analyzed by western blotting to look for the presence of all the three proteins under examination (figure 3.18B).

Presence of the DNA segments corresponding to the origin start site (fragment 5), the TIMM13 promoter region (fragments 8 and 9), and the two fragments immunoprecipitated with anti-histone antibodies (fragments 7 and 10) was analyzed by real-time PCR. The results of these Re-ChIP experiments revealed the presence of ORC4 in both the c-Fos and c-Jun immunoprecipitates exclusively obtained from late G1 cells, both at the DNA replication and TIMM13 promoter areas (figure 3.19). These results indicate that an interaction between ORC4 and the two AP-1 proteins occurs in vivo at the Lamin B2 in a specific window of the cell cycle that precedes the G1-S transition.

A.



B.

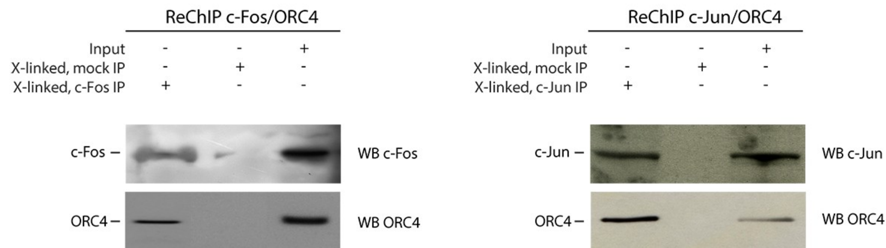


Figure 3.18: Development of the Re-ChIP procedure. (A) Schematic representation of the Re-ChIP procedure. Crosslinked chromatin, which was immunoprecipitated with anti-c-Fos or c-Jun antibodies, was subjected to a mild DTT treatment and re-probed using an anti ORC4 antibody. The immunocomplex was then treated to reverse the crosslink and DNA was analyzed by real time qPCR. (B) Part of the material of each re-ChIP was subjected to immunoblot before qPCR to confirm the co-presence of c-Fos/ORC4 and c-Jun/ORC4.

These results indicate that an interaction between ORC4 and the two AP-1 proteins occurs in vivo at the Lamin B2 in a specific window of the cell cycle that precedes the G1-S transition.

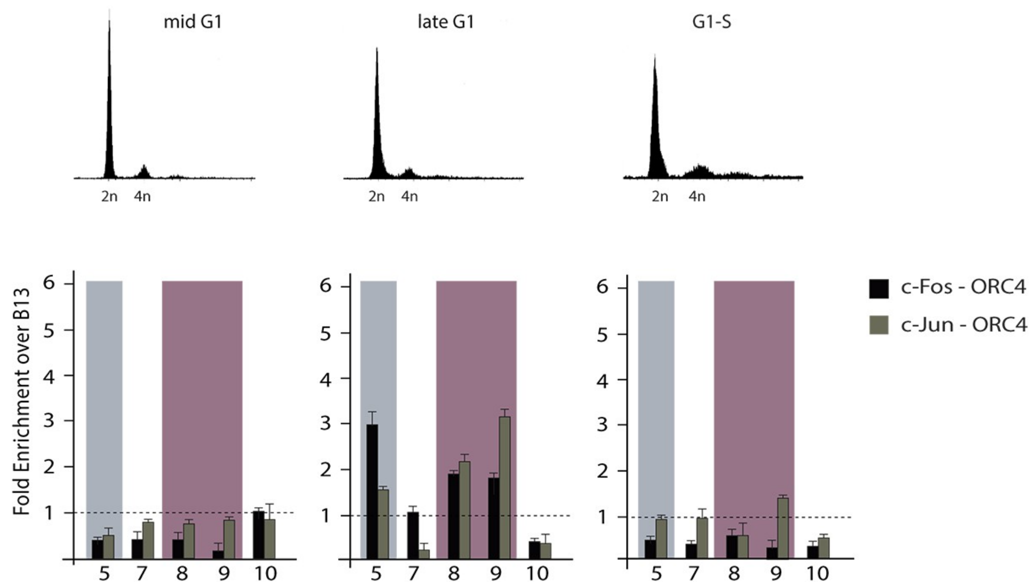


Figure 3.19: AP-1/ORC4 chromatin binding profiles by re-ChIP analyses. Upper part: flow cytometry profiles of T98G cells synchronized at different time points corresponding to different stages of the cell cycle, and subjected to Re-ChIP. Lower part: results of real time qPCR analysis, showing that interaction between c-Fos, c-Jun and ORC4 occurs on the Lamin B2 origin only during the late G1 phase of the cell cycle.

3.6 Effects of the disruption of topological structure on protein binding across the LaminB2 origin region

AP-1 proteins are known to bind DNA producing double helix distortions, resulting in a thermodynamically favorable state which favors binding of other proteins to DNA. To understand whether this might alto be the case for the AP-1:LaminB2 interaction, we wanted to chemically perturb DNA topology and hence induce replicative stress, followed by the analysis of the protein:lamin B2 origin interactions.

Different drugs can be used to perturb the topological state of DNA. Since Topoisomerase II was demonstrated to interact with the LaminB2 origin of DNA replication [87], we decided to study the effects of 5-[N-phenylcarboxamido]-2-thiobarbituric, acid known also as merbarone [191], which binds Topoisomerase II and prevents its interaction with DNA, thus impeding relaxation of the torsional stress induced by DNA unwinding. In contrast to etoposide, which freezes Topoisomerases II in an irreversible complex, merbarone binding to this enzyme is reversible [191]. We initially studied the effects of the drug on DNA synthesis by analyzing BrdU incorporation in merbarone-treated cells. As shown in figure 3.20, although cells subjected to drug treatment displayed a normal propidium iodide incorporation profile compared to control or DMSO-treated cells, DNA synthesis was nevertheless drastically decreased.

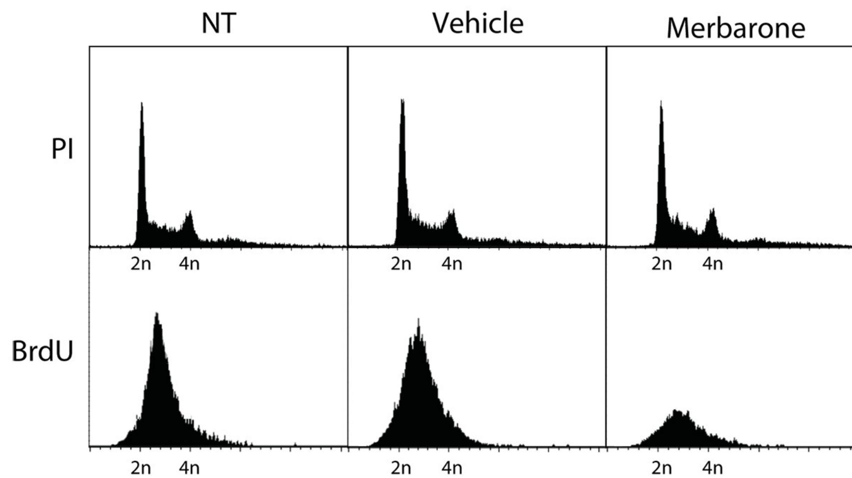


Fig. 3.20: Effect of the topoisomerase II inhibitor merbarone on DNA synthesis. Flow cytometry analysis of T98G cells treated with merbarone. The upper panels show propidium iodide (PI) staining; the lower panels show BrdU incorporation.

Previous published data have shown that topoisomerase II joins the pre-RC during the mid G1 phase of the cell cycle [87]. To monitor the chromatin binding profile of the proteins analyzed so far across the Lamin B2 origin region, we performed chromatin immuno-precipitation after 2 hours of merbarone treatment and compared these results with those obtained from cells treated with only DMSO. ORC4 binding to the start site region, an event occurring early in G1, was not affect by drug treatment, as it was not binding of USF1 to the TIMM13 promoter. In contrast, merbarone completely blocked origin interaction with HOXC13, a protein that we previously showed to bind the DNA replication start site in late G1 , in addition to USF1 binding with the start site region (Figure 5B). Of interest, we found that the binding profile of c-Fos was drastically altered by the drug, since the protein was found associated to the TIMM13 promoter instead of the start site, thus with a profile similar to the one observed in the mid G1 phase of the cell cycle. Taken together, these results indicate that passage from the early to the late phases of the cell cycle is concomitant with a shift in the position of c-Fos binding and that this event requires a topological modification in the origin region.

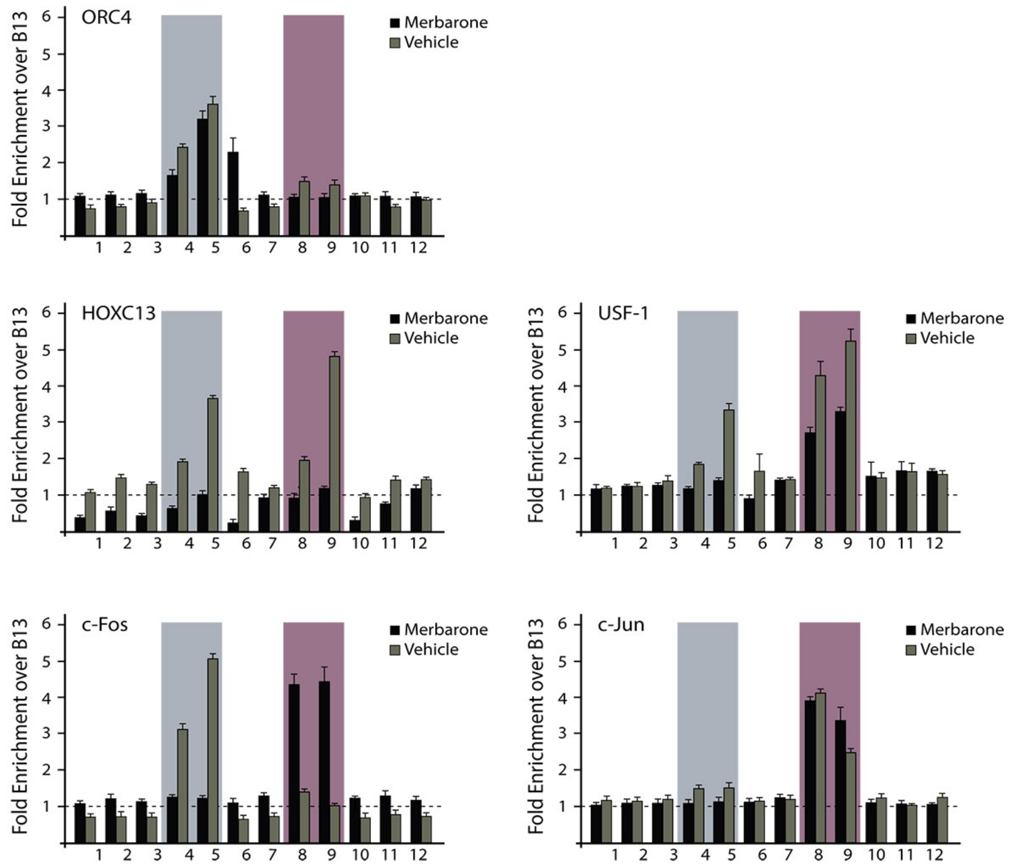


Fig. 3.21: Effects of merbarone on protein:DNA interactions across the Lamin B2 origin region. The graphs show the results of high resolution ChIP experiments in T98G cells treated with merbarone.

Chapter

4

Discussion

In this work we have explored the spatial and temporal dynamics of the interaction of the AP-1 proteins c-Fos and c-Jun, with the Lamin B2 origin of DNA replication. The results indicate that both proteins interact with the member of the pre-RC complex ORC4 in coincidence with origin activation and that this interaction appears regulated by the topological chromatin environment. These results are consistent with the possibility that both c-Fos and c-Jun participate in Lamin B2 origin specification.

4.1 c-Fos and c-Jun are involved in the process of DNA replication

This work provides evidence that both c-Fos and c-Jun interact with the LaminB2 origin of DNA replication and might be part of the origin-associated replicative machinery. This statement is supported by a series of experimental observations. First, we found that these transcription factors can be immuno-precipitated together with HA-Flag HOXC13, which we previously demonstrated to be a component of the replicative machinery [96, 97]. Second, both AP-1 factors also interact with ORC4, a well known DNA replication origin-binding factor. Third, the interaction with ORC4 occurs at different phases of the cell cycle. Fourth, binding of c-Fos and c-Jun with the origin DNA is sufficiently strong to be detected also by ChIP in native, non-crosslinked conditions. Based on these considerations, c-Fos and c-Jun appear to join the list of transcriptional activators also known to act as co-factors in the process of DNA replication. This list also includes Myb, E2F1 [93], Oct-1 [94], c-Myc [95] and various homeotic proteins [19, 96, 97, 173, 70, 174, 175].

4.2 c-Fos and c-Jun interact with the replicative machinery during late G1

Our results demonstrate that the interaction between c-Fos and c-Jun and the origin of DNA replication is regulated during the cell cycle and mainly occurs during the late G1 phase, when the pre-replicative complex assembly is completed. This conclusion is drawn by the results of two complementary approaches involving DNA:protein and protein:protein analyses. In the first case, using high resolution ChIP, we observed that, during the late G1 phase of the cell cycle, whereas chromatin probed with anti c-Fos antibody was enriched for the DNA fragment encompassing the start site of the

Lamin B2 origin, c-Jun was instead associated with the TIMM13 promoter area. It is certainly not surprising to find transcription factors associated with early-replicating chromatin, such as that of the Lamin B2, especially since early DNA replication has been reported to occur at open, transcriptionally-active chromatin [25, 192, 193]. Thus, it might well be envisaged that c-Fos and c-Jun are located in the promoter region of the TIMM13 gene to play a role in transcriptional activity. It is instead more intriguing that the AP-1 factors might play a role in determining origin function. Indeed, this possibility is consistent with bioinformatic data suggesting that the presence of c-Fos and c-Jun binding sites can be used as a discriminant approach for human origin discovery [102]. Moreover, our results are in agreement with those reported for the c-Myc transcription factor, which was also found to bind early replication origins and be involved in origin activity in human cells by promoting G1-S transition and favoring S phase entry [95], similarly to what reported for c-Fos and c-Jun [149, 188].

4.3 c-Fos and c-Jun interaction with LaminB2 origin relies on chromatin topology

As previously mentioned (refer to paragraphs 1.1.3 and 1.1.4), several authors now consider the determinants of mammalian ORC binding as dictated by chromatin structure and epigenetic modification, rather than primary DNA sequence [36, 82, 194]. Given that ORC binding may occur in the context of chromatin, the question then arises whether chromatin remodeling has a role during DNA replication initiation. In this work we addressed this question by analyzing, at high resolution, the chromatin binding profile of several proteins in a 1.1 Kbps DNA region encompassing both the Lamin B2 start site region and the TIMM13 promoter. Then, we also subjected cells to treatment with merbarone, a topoisomerase II reversible inhibitor which, as already mentioned, establishes a critical time point for the pre-

replicative complex assembly, preventing cell exit from the mid G1 phase of the cell cycle [195]. In cells treated with this drug, we did not find any alteration for the ORC4 binding to the Lamin B2 origin start site region. This result was not unexpected, due to the well establish ability of this protein to interact with the DNA replication start site during the M to G1 transition [14, 196]. USF-1, a transcription factor known to bind the TIMM13 promoter region through all the phases of the cell cycle [98], was instead affected by drug treatment. Indeed, we found that while its binding to the promoter region was not drug sensitive, the enrichment found for the start site region disappeared. The topological interference of merbarone was also confirmed by the absence, from both regions, of the transcription factor HOXC13, which was previously demonstrated to join the pre-replicative complex during the late G1 phase of the cell cycle [97]. This results is consistent with the conclusion that a topological change of chromatin might occur during the transition from mid to late-G1. Further corroborating this possibility, we found that, while c-Jun enrichment for the TIMM13 region is maintained similar to both asynchronous and synchronized cells, merbarone treatment instead specifically affected the c-Fos binding profile. The finding that, under these conditions, c-Fos protein did not bind to the start site of the Lamin B2 origin of DNA replication as found in asynchronously growing cells and in late G1 cells, showing instead a binding profile similar to that of mid G1 cells, confirms that merbarone critically acts at a time point between mid to late G1 when a transition in chromatin conformation occurs at the origin.

4.4 A speculative model for origin activation

In a recent study, the dyad symmetry (DS) region of origin of plasmid replication (OriP) of the Epstein-Barr virus (EBV) was found to be flanked by nucleosomes that undergo chromatin remodeling at the G1-S border of the cell cycle. These changes,

which correlate with host MCM3 binding, suggest that cell cycle changes in chromatin topology are coordinated with replication licensing at OriP [118]. Linkage between chromatin conformation and DNA replication was also found for the bacteria. For example, studies performed on plasmid R6K [197, 198] have demonstrated induction of DNA double helix distortion during the process of DNA replication. While R6K loop formation occurs at a distance, transmission electron microscopy (TEM) data obtained studying HPV-11 replication [199] showed a local DNA distortion involving ~500 bps with the generation of both small and large DNA:protein particles and looped species. A similar feature was obtained analyzing *in vitro* the LaminB2 origin [200]. In this work, the authors detected two particular sequences allowing this origin to deviate from the canonical structure and giving rise to triplex helix formation. Two specific DNA sequences were reported as responsible for these so called "non-canonical structures". When presence of both these sequences was tested within the analyzed 1.1 Kbs origin sequence, one was found in correspondence of the start site region (fragment 4 and 5) while the other close to the USF-1 binding site (fragment 8), as shown in pale blue in Figure 4.1. This finding is fully consistent with the overall conclusion that presence of non canonical DNA structures accompany the process of origin activation. Furthermore, the same authors, a few years later, also described that human ORC4 protein binds *in vitro* not double-, but triple-stranded DNA [201], once again pointing out that non B-DNA structures, rather than primary sequence, might affect ORC association with DNA. This conclusion appears to be in agreement with our established knowledge on the characteristics of both Drosophila ORC (*Dm*ORC) and Schizosaccharomice pombe ORC (*Sp*ORC), which display different sequence specificity but common strong preference for negatively supercoiled DNA.

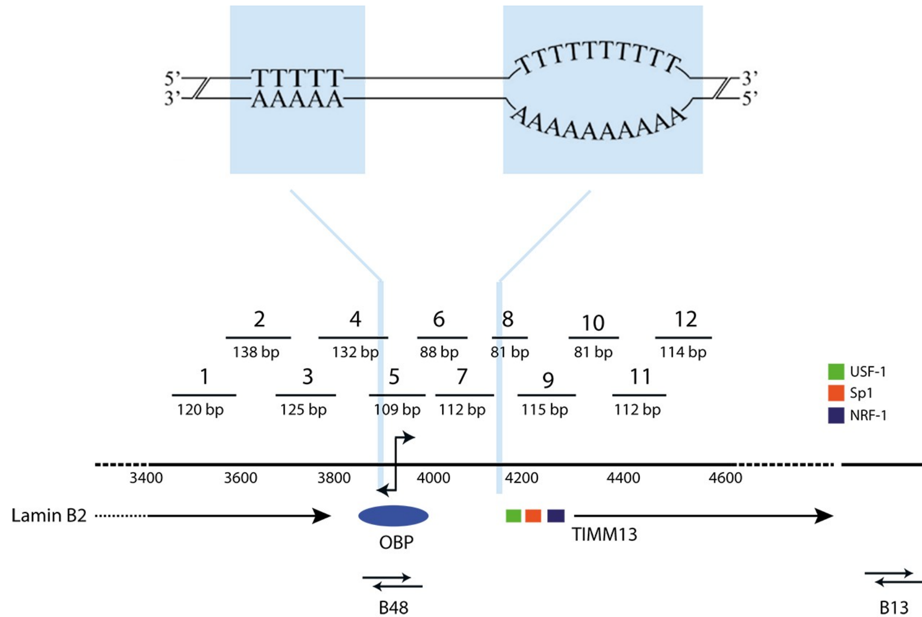


Figure 4.1: Alignment of the "non-canonical sequences" to the 1.1 Kbps region analyzed. The DNA sequences described as "non-canonical" [200] (highlighted in pale blue squares) were superimposed to the 1.1 Kbps DNA used for our high resolution chromatin binding profile investigations.

A possible model, which could explain our results on AP-1 and USF-1 binding and is consistent with the literature information is shown in figure 4.2. According to this model, during the transition from mid to late G1 phase of the cell cycle, there is a local chromatin topological change which brings the proteins studied close to each other. This model, however, cannot explain by itself the opposite enrichments observed for c-Fos and c-Jun. It might thus be possible that early replication origins, such as LaminB2, are prone to form a four-stranded DNA structure known as G-quadruplex or G4. Because mammalian ORC has a strong preference for binding to supercoiled DNA [82], and supercoiled DNA promotes the formation of G4 structures in the appropriate DNA sequences, this sequence may be a major determinant of origin recognition in human cells [202, 203, 204, 205].

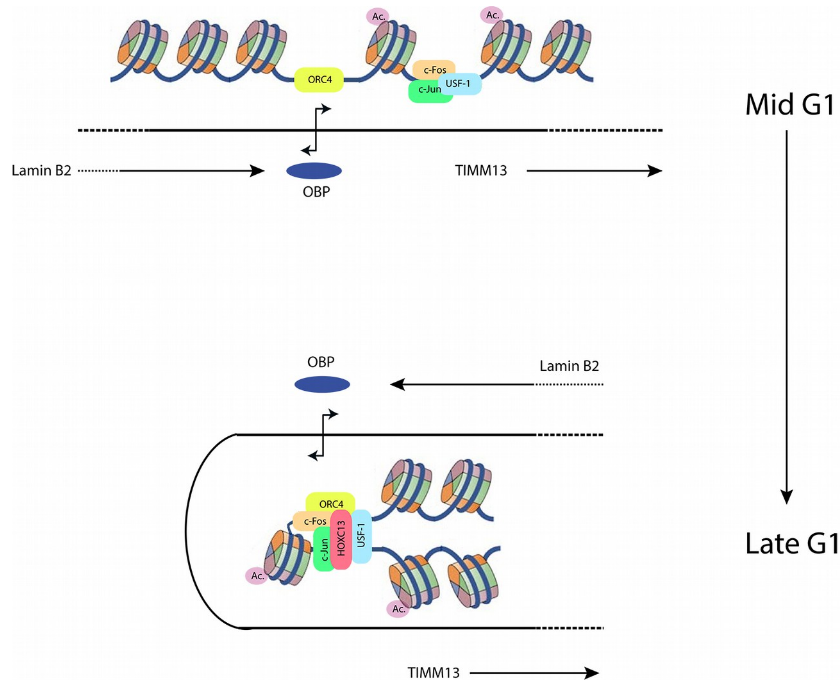


Figure 4.2: Proposed model. Combining our findings and data from literature, we speculate on a possible model which might explain the enrichment found for both origin and the TIMM13 promoter. In this model, during the transition from mid to late G1 phase of the cell cycle, the local chromatin environment changes, allowing these two sequences to be close each other supporting the formation of the complete pre-RC. ORC4 is represented in yellow, USF-1 in pale blue, HOXC13 in pink, acetylated histone 3 in purple, c-Fos and c-Jun in orange and green respectively.

Further experiments are clearly needed to further explore this possibility. In the context of these experiments, it will also be important to ascertain whether the process of activation of the Lamin B2 origin (and of other origins more in general) might also involve the presence of non-coding RNAs [206, 207]. Recent information indeed shows that G-quadruplex RNA structures play an essential role in Epstein-Barr virus DNA replication by directly interacting with EBNA1, a protein known to be critical for both replication and genome maintenance during latency in proliferating cells. In particular, EBNA1 binding to human ORC1 [208, 209] appears to be mediated mediated by RNA G-quadruplex; moreover this process is stimulated

by the presence of human Cdc6 [209]. Finally, a recent article has pointed out the strong affinity of human ORC1 protein to G-quadruplex RNA and single stranded DNA rather than double stranded DNA [210]. Thus, it might well be envisaged that structured RNAs might also broadly participate in origin definition and activation as hypothesized by the drastic genome re-organization after RNAse treatment inside the nucleus [211]. The lamin B2 origin might represent a very suitable model to explore and further dissect these molecular events.

Chapter 5

Materials and methods

5.1 Cell culture, synchronization and Merbarone treatment

T98G (ATCC CRL-1690), HeLa (ATCC CCL-2) cells as well as a stable clone of T98G cells overexpressing HA/FLAG HOXC13 (kind gift of Dr. Ramiro Mendoza-Maldonado from ICGEB in Trieste) were cultured in Dulbecco's modified Eagle medium (D-MEM) containing 1 g/ml glucose, 1 mM pyruvate, Glutamax (Life Technologies) and supplemented with 10 U/L penicillin, 10 µg/L streptomycin and 10% fetal calf serum (Gibco-Life Technologies).

5.1.1 Cell synchronization and FACS analysis

Subconfluent T98G cells (~70% confluence) were synchronized at G0 phase of the cell cycle by 72 hours of culture in serum free medium.. The proliferation block was than released by adding back the complete medium; cells were then cultured for further 9, 14, 16 and 20 hours before cold ethanol fixation, in order to obtain synchronized cell populations at mid G1, late G1, G1/S border respectively according to a published procedure [189].

After three washes with PBS, cells were stained with propidium iodide by resuspending the obtained pellet in 500µl of a solution containing 375µl Sodium citrate 0.1% w/v, 125µl 1mg/ml propidium iodide (Sigma P4864), 6.25µl NP-40 0.1% v/v and 0.625µl 10mg/ml RNAse A (Roche); following incubation of 15 minutes at room temperature in the dark, the cells were analyzed with FACScalibur flow cytometer (BD Bioscences).

5.1.2 Merbarone treatment

Topoisomerase II inhibitor merbarone (5-(N-Phenylcarbamoyl)-2-thiobarbituric acid, NSC-336628) was purchased from Sigma (M2070) and dissolved in DMSO to achieve a concentration of 2mM. T98G cells were treated with merbarone at the final concentration of 100µM in complete medium for 1 hour at 37°C and 5%CO_2 as published [191]. Untreated or diluent (DMSO)-treated cells were used as negative controls in chromatin immuno-precipitation experiments.

5.2 Antibodies

The following primary antibodies were used for either ChIP or Western Blot (WB)

experiments: mouse monoclonal anti FLAG M2 HRP-conjugated, A8592, from Sigma (WB 1:1000), rabbit polyclonal anti HOXC13 gently gifted by Dr. Ramiro Mendoza-Maldonado from ICGEB in Trieste (WB 1:500), mouse monoclonal anti HA agarose-conjugated, Clone HA-7, A2095, from Sigma (ChIP 0.5μg/mg proteins), rabbit polyclonal anti c-Fos, sc-52, from Santa Cruz (WB 1:400, ChIP 1μg/mg proteins), rabbit polyclonal anti c-Jun, sc-45, from Santa cruz (WB 1:500, ChIP 1μg/mg proteins), rabbit polyclonal anti NF-kB p50, sc-7178, from Santa Cruz (WB 1:500), rabbit polyclonal anti NF-kB p65, sc-372, from Santa Cruz (ChIP 1μg/mg proteins), mouse monoclonal anti NF-kB p65, MAB3026, Upstate (WB 1:500), mouse monoclonal anti ORC4, 611170, from BD Biosciences (WB 1:1000), rabbit polyclonal anti ORC4, sc-20634, from Santa Cruz (ChIP 1μg/mg proteins), rabbit polyclonal anti USF-1, sc-229, from Santa Cruz (WB 1:500, ChIP 1μg/mg proteins), rabbit polyclonal anti H2B, sc-10808, (WB 1:500, ChIP 1μg/mg proteins), rabbit polyclonal anti acetyl-Histone 3 (K14) Millipore, 06-599, (WB 1:1000, ChIP 1μg/mg proteins), rabbit polyclonal anti PBX 1-2-3, sc-888, from Santa Cruz (WB 1:500, ChIP 1μg/mg proteins), rabbit polyclonal anti JunD, sc-74, from Santa Cruz (WB 1:500, ChIP 1μg/mg proteins) rabbit polyclonal anti NFAT1, sc-13034, (WB 1:500, ChIP 1μg/mg proteins), mouse monoclonal anti Cyclin A, sc-239, from Santa Cruz (WB1:500). As negative controls we employed pre-immune serum form rabbit and mouse from Sigma (R9133 and M5905 respectively). We used monoclonal mouse anti rabbit light chain 211-032-171 and goat anti mouse light chain 115-035-174 secondary antibodies (Jackson immune Research) in Western Blot experiments .

5.3 in vivo topoisomerase I and II mapping

Topoisomerases I and II mapping in T98G, T98G-HF and HeLa cells was performed according to an already published protocol [87]. Briefly, cells were incubated with either 1 mM CPT or 10 nM VP16 in complete medium for 1 min, washed twice with

PBS containing the same amount of drug and lysed in 250 mM Tris–HCl, pH 8, 25 mM EDTA, 5 mM NaCl, 0.5% SDS and 800 mg/ml proteinase K. DNA was isolated by phenol:chloroform:isoamyl alcohol extraction followed by ethanol precipitation and resuspended in TE buffer (1 mM EDTA and 10 mM Tris–HCl pH 8). The topoisomerase I cleavage sites was detected by LM-PCR as previously described [179, 212, 213]. The topoisomerase II cleavage site was detected by TD-PCR [180]. The primer sets used for LM- and TD-PCR are the same already described [100].

5.4 Chromatin immuno precipitation analyses (ChIP)

5.4.1 X-ChIP

Chromatin preparations were obtained by biochemical fractionation as described in [14] with modifications introduced to avoid non specific signals and to optimize the analysis of proteins tightly bound to the DNA as well as for Native ChIP [181]. Briefly, cells were lysed by Dounce homogenization in hypothonic buffer (10mM Hepes pH 7.9, 1.5mM $MgCl_2$, 10mM KCl, 5mM Sodium Butyrate). Nuclei were collected by centrifugation, resuspended in buffer S1 (20mM Hepes pH 7.9, 250mM Sucrose, 10mM $MgCl_2$), gently overlayed on buffer S2 (20mM Hepes pH 7.9, 350mM Sucrose, 0.5mM $MgCl_2$) and recovered by centrifugation. Purified nuclei were washed with buffer B (20mM Hepes pH 7.9, 1.5mM $MgCl_2$, 20mM Kcl, 5mM Sodium Butyrate, 10% glycerol) and incubated at 4°C in lysis buffer (20mM Hepes pH 7.9, 750mM ε-aminocaproic acid, 0.5% Triton-X 100). Chromatin was recovered by centrifugation and resuspended in crosslink buffer (1% formaldehyde in 20mM Hepes pH 7.9); after 10 minutes the reaction was quenched by adding 125mM Glycine. Unbound material was removed by one washing step in NaCl 1M in buffer C (same as B without protease inhibitors), two more washing steps in buffer C and finally chromatin was resuspended in buffer D (20mM Hepes pH 7.9, 3mM $CaCl_2$)

and subjected to Micrococcal nuclease (Roche) digestion. Reaction was stopped by adding 4mM EDTA; the obtained material was sonicated 3 times in ice for 5 seconds in order to further shear the DNA (~200bps) and then, after quantification by Bradford assay, subjected to immunoprecipitation. The immunocomplexes were washed in low salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl, pH 7.9, 150 mM NaCl), high salt buffer (500 mM NaCl), LiCl buffer (250 mM LiCl, 1% IGEPAL, 1% deoxycholic acid, 1 mM EDTA, 10 mM Tris-HCl, pH 8) and TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8). The washed precipitates were divided for western blot and DNA extraction. Laemli buffer 5x was added to the material for western blot, after 10 minutes at 95°C the material was ran in a 10% denaturing acrylamide gel and transferred to a PVDF 0.45μm (Amersham) using a Hoefer Semaphor semi-dry blotting apparatus. DNA was recovered from the immunocomplexes by phenol:chloroform:isoamyl alcohol (24:25:1) extraction after RNAseA (Roche) and Proteinase K digestion. Finally, DNA was precipitated with absolute ethanol and washed once with 70% ethanol and quantified.

5.4.2 Native chromatin immuno precipitation analyses (NChIP)

The procedure followed for NChIP is the same as described above with the exception that we omitted not only formaldehyde treatment but also the sonication step in order to avoid disruption of the DNA:protein and protein:protein complexes.

5.4.3 Sequential chromatin immunoprecipitation analyses (Re-ChIP)

In order to identify interacting or closely proximal proteins simultaneously binding to the 1.1 Kbps genomic region, encompassing both the LaminB2 ORI and the TIMM13 promoter, we employed the sequential chromatin immunoprecipitations technique (Re-ChIP) [190] with two different antibodies. A brief introduction to this protocol

was previously described in the Results chapter (paragraph 3.5). Briefly, the purified chromatin was used to immunoprecipitate the DNA:protein complexes with the first antibody, as previously described. Immuno-complexes were released from the agarose beads by two sequential incubation with an equal volume (~50 µl) of 10mM DTT for 30 minutes at 37°C. The eluates of these two steps were combined and diluted 100 times; a volume corresponding to 10% of the total sample was kept as input control while the remaining amount was immunoprecipitated with the second antibody. Immuno-precipitated material was washed and the DNA was purified as previously described

5.5 PCR analyses

5.5.1 Competitive PCR

Quantification of the relative abundance of LaminB2 origin was performed by competitive PCR following a published protocol [178]. Briefly known scalar dilution of competitor DNA was mixed to 25ng of immunoprecipitated DNA (using anti-HA, anti c-Fos, anti c-Jun and pre immune serum) and from input chromatin, followed by amplification with the primers specific for origin and not origin regions, B48 and B13 respectively, for 35 cycles. PCR products were resolved on a 10% acrylamide gel, stained with ethidium bromide and visualized under UV light. Intensity of the bands corresponding to the B48 or B13 (target or T) and competitor (C) products, were quantified using ImageJ software.

The principle of competitive PCR is based on the concept that the levels of template and competitor maintain their ratio throughout the amplification reaction and that the ratio of the final products is a linear function of the input levels of the competitor. Therefore, when the T:C ratio is equal to 1, the amount of target DNA exactly corresponds to the amount of the competitor.

5.5.2 High resolution LaminB2 PCR

As mentioned in paragraph 3.4 of the Results chapter, taking advantage of Amplify 3 and Vector NTI © softwares, we designed twelve pairs of primers on the 1.1 Kbps region encompassing both the start site of the LaminB2 ORI and the TIMM13 promoter. These primers amplify partially overlapping regions of 80-120bps and have a similar annealing temperature with the exception of primers targeted to the AT reach replication start site region, which annealing temperature is 60°C as reported at the end of this paragraph. The efficiency of all the primer sets was analyzed in real time PCR using the slope analysis approach, as reported in paragraph 3.4.1 and in figure 3.7 of the Results chapter, by amplifying four scalar concentrations of genomic DNA, obtained from three independent purifications. The relative enrichment of all the amplified fragments obtained in real-time PCR was calculated using the ΔΔCt method. Briefly, for each fragment and for B13, used as a control, the Ct value of the amplified DNAs obtained from IP with both specific antibody (Ab) and with pre-immune serum (IS) was subtracted from the input Ct value (ΔCt).

$$\Delta Ct_{fragmX}(Ab) = Ct_{input} - Ct_{fragmX}(Ab) \Delta Ct_{fragmX}(IS) = Ct_{input} - Ct_{fragmX}(IS)$$

Then ΔCt (IS) value was subtracted from the ΔCt (Ab).

$$\Delta\Delta Ct_{fragmX} = \Delta Ct_{fragmX}(Ab) - \Delta Ct_{fragmX}(IS)$$

$$\Delta\Delta Ct_{B13} = \Delta Ct_{B13}(Ab) - \Delta Ct_{B13}(IS)$$

Each ΔΔCt value was normalized for B13 and enrichment resulted as follows:

$$Enrichment_{fragmX} = 2^{(\Delta\Delta Ct_{fragmX} - \Delta\Delta Ct_{B13})}$$

Fragment Number		Name	Sequence	Start /End	Amplicon Lenght	Annealing
1	FW	1L	GTGACTCGAAGCCTAGCGCCCTCCCTGC	3480	120bps	70°C
	REV	2U	GCTTCAGCGTTAACCTGGCGCTGTGCA	3600		
2	FW	2L	GCGCCAGGTTAACGCTGAAGCCTGCC	3580	138bps	70°C
	REV	3U	AACTGCCGCGTGCAGGCTTCAGACCAA	3718		
3	FW	3L	GGTTGGTCTGAAGCCTGCACGCGGC	3690	125bps	70°C
	REV	4U	TTGCAGGTTGTGCTGTGACGCTCGCTG	3815		
4	FW	4L	CGTCCAGCGAGCGTCACAGCACAACC	3785	132bps	60°C
	REV	5U	TCATTGGAAAAAAAAAAAGAACACGCTAGGCATGCA	3917		
5	FW	5L	TGCATGCCTAGCGTGTTCTTTTTTTTTTTCCAATGA	3883	109bps	60°C
	REV	6U	GGCAGAACCTAAAATCAAAATGTTTATTGGAGTG	3992		
6	FW	T1	GGTTCTGCCTCTGAGTTTATTCCTGAGG	3985	88bps	60°C
	REV	E1	GGGGTGGAGGGATCTTTCTTAGACA	4073		
7	FW	6L	GGGCCTCTGCCCTAATGAAGCGGATGTCT	4025	112bps	70°C
	REV	7U	GGGTCCCATGCATCGCCTGGGTCC	4137		
8	FW	7L	GGACCCAGGCGATGCATGGGACCC	4114	81bps	70°C
	REV	8U	CGTGACGAAGAGTCAGCTTGTGCAACAGCGT	4195		
9	FW	8L	TCTTCGTCACGTGATGCGACCGGCTC	4185	115bps	70°C
	REV	9U	CACGCTCTGCCTCCAGCTCGTCCC	4300		
10	FW	9L	GGCAGAGCGTGAGTACAAAGTGATCGGCCTC	4290	81bps	70°C
	REV	10U	GGTTGCGACTCCGCGGGAAGAGGGA	4388		
11	FW	10L	CAACCACGGGTAGCTCGTGTAGGTAACGGCA	4384	112bps	70°C
	REV	11U	TACAACTCCCACACGACCGCGCGC	4496		
12	FW	11L	GCGCGGTCGTGTGGGAGTTGTAGTCCTC	4475	114bps	70°C
	REV	12U	GGCCGCGTGCGCCGACTCGTAACT	4589		

Table 5.1: Primers used for high resolution protein binding analysis at the LaminB2 replication origin. The primers set detected and used for the origin dissection are reported.

References

1. Jacob, F. and S. Brenner, *[On the regulation of DNA synthesis in bacteria: the hypothesis of the replicon].* C R Hebd Seances Acad Sci, 1963. **256**: p. 298-300.
2. Kornberg, A., *Control of initiation of the Escherichia coli chromosome.* Cold Spring Harb Symp Quant Biol, 1991. **56**: p. 275-8.
3. Huberman, J.A. and A.D. Riggs, *Autoradiography of chromosomal DNA fibers from Chinese hamster cells.* Proc Natl Acad Sci U S A, 1966. **55**(3): p. 599-606.
4. DePamphilis, M.L., *Eukaryotic DNA replication: anatomy of an origin.* Annu Rev Biochem, 1993. **62**: p. 29-63.
5. Huberman, J.A., *Cell cycle. A licence to replicate.* Nature, 1995. **375**(6530): p. 360-1.
6. Stillman, B., *Cell cycle control of DNA replication.* Science, 1996. **274**(5293): p. 1659-64.
7. Ibarra, A., E. Schwob, and J. Mendez, *Excess MCM proteins protect human cells from replicative stress by licensing backup origins of replication.* Proc Natl Acad Sci U S A, 2008. **105**(26): p. 8956-61.
8. Bell, S.P. and A. Dutta, *DNA replication in eukaryotic cells.* Annu Rev Biochem, 2002. **71**: p. 333-74.
9. Conti, C., J.A. Seiler, and Y. Pommier, *The mammalian DNA replication elongation checkpoint: implication of Chk1 and relationship with origin firing as determined by single DNA molecule and single cell analyses.* Cell Cycle, 2007. **6**(22): p. 2760-7.
10. Duggin, I.G., et al., *The replication fork trap and termination of chromosome replication.* Mol Microbiol, 2008. **70**(6): p. 1323-33.

11. Kong, D. and M.L. DePamphilis, *Site-specific DNA binding of the Schizosaccharomyces pombe origin recognition complex is determined by the Orc4 subunit.* Mol Cell Biol, 2001. **21**(23): p. 8095-103.
12. Gilbert, D.M., *In search of the holy replicator.* Nat Rev Mol Cell Biol, 2004. **5**(10): p. 848-55.
13. Mechali, M., *DNA replication origins: from sequence specificity to epigenetics.* Nat Rev Genet, 2001. **2**(8): p. 640-5.
14. Mendez, J. and B. Stillman, *Chromatin association of human origin recognition complex, cdc6, and minichromosome maintenance proteins during the cell cycle: assembly of prereplication complexes in late mitosis.* Mol Cell Biol, 2000. **20**(22): p. 8602-12.
15. Thomson, A.M., P.J. Gillespie, and J.J. Blow, *Replication factory activation can be decoupled from the replication timing program by modulating Cdk levels.* J Cell Biol, 2010. **188**(2): p. 209-21.
16. Aparicio, O.M., A.M. Stout, and S.P. Bell, *Differential assembly of Cdc45p and DNA polymerases at early and late origins of DNA replication.* Proc Natl Acad Sci U S A, 1999. **96**(16): p. 9130-5.
17. Liu, P., et al., *The Chk1-mediated S-phase checkpoint targets initiation factor Cdc45 via a Cdc25A/Cdk2-independent mechanism.* J Biol Chem, 2006. **281**(41): p. 30631-44.
18. Diffley, J.F., *Regulation of early events in chromosome replication.* Curr Biol, 2004. **14**(18): p. R778-86.
19. Wohlschlegel, J.A., et al., *Inhibition of eukaryotic DNA replication by geminin binding to Cdt1.* Science, 2000. **290**(5500): p. 2309-12.
20. Zou, L. and B. Stillman, *Formation of a preinitiation complex by S-phase cyclin CDK-dependent loading of Cdc45p onto chromatin.* Science, 1998. **280**(5363): p. 593-6.
21. Noton, E. and J.F. Diffley, *CDK inactivation is the only essential function of the APC/C and the mitotic exit network proteins for origin resetting during mitosis.* Mol Cell, 2000. **5**(1): p. 85-95.

22. Sclafani, R.A. and T.M. Holzen, *Cell cycle regulation of DNA replication.* Annu Rev Genet, 2007. **41**: p. 237-80.
23. Schepers, A. and P. Papior, *Why are we where we are? Understanding replication origins and initiation sites in eukaryotes using ChIP-approaches.* Chromosome Res, 2010. **18**(1): p. 63-77.
24. Karnani, N., et al., *Genomic study of replication initiation in human chromosomes reveals the influence of transcription regulation and chromatin structure on origin selection.* Mol Biol Cell, 2010. **21**(3): p. 393-404.
25. MacAlpine, H.K., et al., *Drosophila ORC localizes to open chromatin and marks sites of cohesin complex loading.* Genome Res, 2010. **20**(2): p. 201-11.
26. Norio, P., et al., *Progressive activation of DNA replication initiation in large domains of the immunoglobulin heavy chain locus during B cell development.* Mol Cell, 2005. **20**(4): p. 575-87.
27. Hiratani, I., et al., *Global reorganization of replication domains during embryonic stem cell differentiation.* PLoS Biol, 2008. **6**(10): p. e245.
28. Bartek, J., C. Lukas, and J. Lukas, *Checking on DNA damage in S phase.* Nat Rev Mol Cell Biol, 2004. **5**(10): p. 792-804.
29. Nyberg, K.A., et al., *Toward maintaining the genome: DNA damage and replication checkpoints.* Annu Rev Genet, 2002. **36**: p. 617-56.
30. Santocanale, C. and J.F. Diffley, *A Mec1- and Rad53-dependent checkpoint controls late-firing origins of DNA replication.* Nature, 1998. **395**(6702): p. 615-8.
31. Shirahige, K., et al., *Regulation of DNA-replication origins during cell-cycle progression.* Nature, 1998. **395**(6702): p. 618-21.
32. Mimura, S., et al., *Phosphorylation-dependent binding of mitotic cyclins to Cdc6 contributes to DNA replication control.* Nature, 2004. **431**(7012): p. 1118-23.
33. Lau, E., et al., *The functional role of Cdc6 in S-G2/M in mammalian cells.* EMBO Rep, 2006. **7**(4): p. 425-30.
34. Oehlmann, M., A.J. Score, and J.J. Blow, *The role of Cdc6 in ensuring*

complete genome licensing and S phase checkpoint activation. J Cell Biol, 2004. **165**(2): p. 181-90.

35. Bell, S.P. and B. Stillman, *ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex.* Nature, 1992. **357**(6374): p. 128-34.
36. Vashee, S., et al., *Sequence-independent DNA binding and replication initiation by the human origin recognition complex.* Genes Dev, 2003. **17**(15): p. 1894-908.
37. Duncker, B.P., I.N. Chesnokov, and B.J. McConkey, *The origin recognition complex protein family.* Genome Biol, 2009. **10**(3): p. 214.
38. Speck, C., et al., *ATPase-dependent cooperative binding of ORC and Cdc6 to origin DNA.* Nat Struct Mol Biol, 2005. **12**(11): p. 965-71.
39. Balasov, M., R.P. Huijbregts, and I. Chesnokov, *Role of the Orc6 protein in origin recognition complex-dependent DNA binding and replication in Drosophila melanogaster.* Mol Cell Biol, 2007. **27**(8): p. 3143-53.
40. Balasov, M., R.P. Huijbregts, and I. Chesnokov, *Functional analysis of an Orc6 mutant in Drosophila.* Proc Natl Acad Sci U S A, 2009. **106**(26): p. 10672-7.
41. DePamphilis, M.L., *The 'ORC cycle': a novel pathway for regulating eukaryotic DNA replication.* Gene, 2003. **310**: p. 1-15.
42. Tatsumi, Y., et al., *The ORC1 cycle in human cells: I. cell cycle-regulated oscillation of human ORC1.* J Biol Chem, 2003. **278**(42): p. 41528-34.
43. Li, C.J., A. Vassilev, and M.L. DePamphilis, *Role for Cdk1 (Cdc2)/cyclin A in preventing the mammalian origin recognition complex's largest subunit (Orc1) from binding to chromatin during mitosis.* Mol Cell Biol, 2004. **24**(13): p. 5875-86.

44, Sasaki, T. and D.M. Gilbert, *The many faces of the origin recognition complex.* Curr Opin Cell Biol, 2007. **19**(3): p. 337-43.

45. Zhang, Z., et al., *Structure and function of the BAH-containing domain of Orc1p in epigenetic silencing.* EMBO J, 2002. **21**(17): p. 4600-11.

46. Pak, D.T., et al., *Association of the origin recognition complex with heterochromatin and HP1 in higher eukaryotes.* Cell, 1997. **91**(3): p. 311-23.
47. Lidonnici, M.R., et al., *Subnuclear distribution of the largest subunit of the human origin recognition complex during the cell cycle.* J Cell Sci, 2004. **117**(Pt 22): p. 5221-31.
48 Prasanth, S.G., et al., *Human Orc2 localizes to centrosomes, centromeres and heterochromatin during chromosome inheritance.* EMBO J, 2004. **23**(13): p. 2651-63.
49. Prasanth, S.G., et al., *Human origin recognition complex is essential for HP1 binding to chromatin and heterochromatin organization.* Proc Natl Acad Sci U S A, 2010. **107**(34): p. 15093-8.
50. Prasanth, S.G., K.V. Prasanth, and B. Stillman, *Orc6 involved in DNA replication, chromosome segregation, and cytokinesis.* Science, 2002. **297**(5583): p. 1026-31.
51. Liu, S., et al., *Structural analysis of human Orc6 protein reveals a homology with transcription factor TFIIB.* Proc Natl Acad Sci U S A, 2011. **108**(18): p. 7373-8.
52 Prasanth, S.G., et al., *Dynamics of pre-replication complex proteins during the cell division cycle.* Philos Trans R Soc Lond B Biol Sci, 2004. **359**(1441): p. 7-16.
53. Hemerly, A.S., et al., *Orc1 controls centriole and centrosome copy number in human cells.* Science, 2009. **323**(5915): p. 789-93.
54. Blow, J.J. and B. Hodgson, *Replication licensing--defining the proliferative state?* Trends Cell Biol, 2002. **12**(2): p. 72-8.
55. Randell, J.C., et al., *Sequential ATP hydrolysis by Cdc6 and ORC directs loading of the Mcm2-7 helicase.* Mol Cell, 2006. **21**(1): p. 29-39.
56. Remus, D., et al., *Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing.* Cell, 2009. **139**(4): p. 719-30.
57. Nguyen, V.Q., C. Co, and J.J. Li, *Cyclin-dependent kinases prevent DNA re-replication through multiple mechanisms.* Nature, 2001. **411**(6841): p. 1068-

73.

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

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

60. Diffley, J.F., *DNA replication: building the perfect switch.* Curr Biol, 2001. **11**(9): p. R367-70.

61. Adachi, Y., J. Usukura, and M. Yanagida, *A globular complex formation by Nda1 and the other five members of the MCM protein family in fission yeast.* Genes Cells, 1997. **2**(7): p. 467-79.

62. Wyrick, J.J., et al., *Genome-wide distribution of ORC and MCM proteins in S. cerevisiae: high-resolution mapping of replication origins.* Science, 2001. **294**(5550): p. 2357-60.

63. Montagnoli, A., et al., *Drf1, a novel regulatory subunit for human Cdc7 kinase.* EMBO J, 2002. **21**(12): p. 3171-81.

64. Morgan, D.O., *The Cell Cycle: principles of control*. 2007 ed2007: Oxford University Press/New Science Press

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

66. Masai, H., et al., *Phosphorylation of MCM4 by Cdc7 kinase facilitates its interaction with Cdc45 on the chromatin.* J Biol Chem, 2006. **281**(51): p. 39249-61.

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

68. Izumi, M., et al., *The human homolog of Saccharomyces cerevisiae Mcm10 interacts with replication factors and dissociates from nuclease-resistant nuclear structures in G(2) phase.* Nucleic Acids Res, 2000. **28**(23): p. 4769-77.

69. Aparicio, T., et al., *The human GINS complex associates with Cdc45 and MCM and is essential for DNA replication.* Nucleic Acids Res, 2009. **37**(7): p. 2087-95.

70. Yabuuchi, H., et al., *Ordered assembly of Sld3, GINS and Cdc45 is distinctly regulated by DDK and CDK for activation of replication origins.* EMBO J, 2006. **25**(19): p. 4663-74.

71. Kubota, Y., et al., *A novel ring-like complex of Xenopus proteins essential for the initiation of DNA replication.* Genes Dev, 2003. **17**(9): p. 1141-52.

72. Boskovic, J., et al., *Molecular architecture of the human GINS complex.* EMBO Rep, 2007. **8**(7): p. 678-84.

73. Zegerman, P. and J.F. Diffley, *Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast.* Nature, 2007. **445**(7125): p. 281-5.

74. Moyer, S.E., P.W. Lewis, and M.R. Botchan, *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, 2006. **103**(27): p. 10236-41.

75. Gilbert, D.M., *Making sense of eukaryotic DNA replication origins.* Science, 2001. **294**(5540): p. 96-100.

76. Cvetic, C. and J.C. Walter, *Eukaryotic origins of DNA replication: could you please be more specific?* Semin Cell Dev Biol, 2005. **16**(3): p. 343-53.

77. Mechali, M., *Eukaryotic DNA replication origins: many choices for appropriate answers.* Nat Rev Mol Cell Biol, 2010. **11**(10): p. 728-38.

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

79. Bramhill, D. and A. Kornberg, *Duplex opening by dnaA protein at novel sequences in initiation of replication at the origin of the E. coli chromosome.* Cell, 1988. **52**(5): p. 743-55.

80. Fuller, R.S. and A. Kornberg, *Purified dnaA protein in initiation of replication at the Escherichia coli chromosomal origin of replication.* Proc Natl Acad Sci U S A, 1983. **80**(19): p. 5817-21.

81. Davey, M.J., et al., *Motors and switches: AAA+ machines within the replisome.* Nat Rev Mol Cell Biol, 2002. **3**(11): p. 826-35.
82. Remus, D., E.L. Beall, and M.R. Botchan, *DNA topology, not DNA sequence, is a critical determinant for Drosophila ORC-DNA binding.* EMBO J, 2004. **23**(4): p. 897-907.
83. Aggarwal, B.D. and B.R. Calvi, *Chromatin regulates origin activity in Drosophila follicle cells.* Nature, 2004. **430**(6997): p. 372-6.
84. Gilbert, D.M., *Cell fate transitions and the replication timing decision point.* J Cell Biol, 2010. **191**(5): p. 899-903.
85. Li, F., et al., *Spatial distribution and specification of mammalian replication origins during G1 phase.* J Cell Biol, 2003. **161**(2): p. 257-66.
86. Antequera, F., *Genomic specification and epigenetic regulation of eukaryotic DNA replication origins.* EMBO J, 2004. **23**(22): p. 4365-70.
87. Abdurashidova, G., et al., *Functional interactions of DNA topoisomerases with a human replication origin.* EMBO J, 2007. **26**(4): p. 998-1009.
88. Murakami, Y. and Y. Ito, *Transcription factors in DNA replication.* Front Biosci, 1999. **4**: p. D824-33.
89. Turner, R. and R. Tjian, *Leucine repeats and an adjacent DNA binding domain mediate the formation of functional cFos-cJun heterodimers.* Science, 1989. **243**(4899): p. 1689-94.
90. Kohzaki, H. and Y. Murakami, *Transcription factors and DNA replication origin selection.* Bioessays, 2005. **27**(11): p. 1107-16.
91. Tyler, J.K. and J.T. Kadonaga, *The "dark side" of chromatin remodeling: repressive effects on transcription.* Cell, 1999. **99**(5): p. 443-6.
92. Marahrens, Y. and B. Stillman, *A yeast chromosomal origin of DNA replication defined by multiple functional elements.* Science, 1992. **255**(5046): p. 817-23.
93. Beall, E.L., et al., *Role for a Drosophila Myb-containing protein complex in site-specific DNA replication.* Nature, 2002. **420**(6917): p. 833-7.
94. Todd, A., et al., *Deletion analysis of minimal sequence requirements for*

autonomous replication of ors8, a monkey early-replicating DNA sequence. J Cell Biochem, 1995. **57**(2): p. 280-9.

95. Dominguez-Sola, D., et al., *Non-transcriptional control of DNA replication by c-Myc.* Nature, 2007. **448**(7152): p. 445-51.
96. Comelli, L., et al., *The homeotic protein HOXC13 is a member of human DNA replication complexes.* Cell Cycle, 2009. **8**(3): p. 454-9.
97. Marchetti, L., et al., *Homeotic proteins participate in the function of human-DNA replication origins.* Nucleic Acids Res, 2010. **38**(22): p. 8105-19.
98. Toth, E.C., et al., *Interactions of USF and Ku antigen with a human DNA region containing a replication origin.* Nucleic Acids Res, 1993. **21**(14): p. 3257-63.
99. Hyrien, O., C. Maric, and M. Mechali, *Transition in specification of embryonic metazoan DNA replication origins.* Science, 1995. **270**(5238): p. 994-7.
100. Abdurashidova, G., et al., *Start sites of bidirectional DNA synthesis at the human lamin B2 origin.* Science, 2000. **287**(5460): p. 2023-6.
101. Ladenburger, E.M., C. Keller, and R. Knippers, *Identification of a binding region for human origin recognition complex proteins 1 and 2 that coincides with an origin of DNA replication.* Mol Cell Biol, 2002. **22**(4): p. 1036-48.
102. Cadoret, J.C., et al., *Genome-wide studies highlight indirect links between human replication origins and gene regulation.* Proc Natl Acad Sci U S A, 2008. **105**(41): p. 15837-42.
103. DePamphilis, M.L., et al., *Regulating the licensing of DNA replication origins in metazoa.* Curr Opin Cell Biol, 2006. **18**(3): p. 231-9.
104. Bielinsky, A.K., *Replication origins: why do we need so many?* Cell Cycle, 2003. **2**(4): p. 307-9.
105. Minami, H., et al., *Binding of AlF-C, an Orc1-binding transcriptional regulator, enhances replicator activity of the rat aldolase B origin.* Mol Cell Biol, 2006. **26**(23): p. 8770-80.
106. Atanasiu, C., et al., *ORC binding to TRF2 stimulates OriP replication.* EMBO

Rep, 2006. **7**(7): p. 716-21.

107. Schepers, A., et al., *Human origin recognition complex binds to the region of the latent origin of DNA replication of Epstein-Barr virus.* EMBO J, 2001. **20**(16): p. 4588-602.

108. Thomae, A.W., et al., *Interaction between HMGA1a and the origin recognition complex creates site-specific replication origins.* Proc Natl Acad Sci U S A, 2008. **105**(5): p. 1692-7.

109. Meisch, F. and M.N. Prioleau, *Genomic approaches to the initiation of DNA replication and chromatin structure reveal a complex relationship.* Brief Funct Genomics, 2011. **10**(1): p. 30-6.

110. Scholer, H., et al., *In vivo competition between a metallothionein regulatory element and the SV40 enhancer.* Science, 1986. **232**(4746): p. 76-80.

111. Lee, W., et al., *Activation of transcription by two factors that bind promoter and enhancer sequences of the human metallothionein gene and SV40.* Nature, 1987. **325**(6102): p. 368-72.

112. Angel, P., et al., *Phorbol ester-inducible genes contain a common cis element recognized by a TPA-modulated trans-acting factor.* Cell, 1987. **49**(6): p. 729-39.

113. Shaulian, E. and M. Karin, *AP-1 as a regulator of cell life and death.* Nat Cell Biol, 2002. **4**(5): p. E131-6.

114. Hess, J., P. Angel, and M. Schorpp-Kistner, *AP-1 subunits: quarrel and harmony among siblings.* J Cell Sci, 2004. **117**(Pt 25): p. 5965-73.

115. Eferl, R., et al., *Functions of c-Jun in liver and heart development.* J Cell Biol, 1999. **145**(5): p. 1049-61.

116. Wagner, E.F., *AP-1--Introductory remarks.* Oncogene, 2001. **20**(19): p. 2334-5.

117. Chinenov, Y. and T.K. Kerppola, *Close encounters of many kinds: Fos-Jun interactions that mediate transcription regulatory specificity.* Oncogene, 2001. **20**(19): p. 2438-52.

118. Zhou, H., et al., *Frequency and distribution of AP-1 sites in the human*

genome. DNA Res, 2005. **12**(2): p. 139-50.

119. Verrecchia, F., et al., *Induction of the AP-1 members c-Jun and JunB by TGF-beta/Smad suppresses early Smad-driven gene activation.* Oncogene, 2001. **20**(18): p. 2205-11.

120. Angel, P., A. Szabowski, and M. Schorpp-Kistner, *Function and regulation of AP-1 subunits in skin physiology and pathology.* Oncogene, 2001. **20**(19): p. 2413-23.

121. Gentz, R., et al., *Parallel association of Fos and Jun leucine zippers juxtaposes DNA binding domains.* Science, 1989. **243**(4899): p. 1695-9.

122. Landschulz, W.H., P.F. Johnson, and S.L. McKnight, *The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins.* Science, 1988. **240**(4860): p. 1759-64.

123. Schuermann, M., et al., *Non-leucine residues in the leucine repeats of Fos and Jun contribute to the stability and determine the specificity of dimerization.* Nucleic Acids Res, 1991. **19**(4): p. 739-46.

124. O'Shea, E.K., et al., *Preferential heterodimer formation by isolated leucine zippers from fos and jun.* Science, 1989. **245**(4918): p. 646-8.

125. Smeal, T., et al., *Different requirements for formation of Jun: Jun and Jun: Fos complexes.* Genes Dev, 1989. **3**(12B): p. 2091-100.

126. Gonzalez, G.A., et al., *A cluster of phosphorylation sites on the cyclic AMP-regulated nuclear factor CREB predicted by its sequence.* Nature, 1989. **337**(6209): p. 749-52.

127. Benbrook, D.M. and N.C. Jones, *Heterodimer formation between CREB and JUN proteins.* Oncogene, 1990. **5**(3): p. 295-302.

128. Macgregor, P.F., C. Abate, and T. Curran, *Direct cloning of leucine zipper proteins: Jun binds cooperatively to the CRE with CRE-BP1.* Oncogene, 1990. **5**(4): p. 451-8.

129. Rauscher, F.J., 3rd, et al., *Common DNA binding site for Fos protein complexes and transcription factor AP-1.* Cell, 1988. **52**(3): p. 471-80.

130. Neuberg, M., et al., *Two functionally different regions in Fos are required for*

the sequence-specific DNA interaction of the Fos/Jun protein complex. Nature, 1989. **338**(6216): p. 589-90.

131. Vogt, P.K. and T.J. Bos, *jun: oncogene and transcription factor.* Adv Cancer Res, 1990. **55**: p. 1-35.
132. Angel, P. and M. Karin, *The role of Jun, Fos and the AP-1 complex in cell-proliferation and transformation.* Biochim Biophys Acta, 1991. **1072**(2-3): p. 129-57.
133. Wisdom, R., *AP-1: one switch for many signals.* Exp Cell Res, 1999. **253**(1): p. 180-5.
134. Treier, M., L.M. Staszewski, and D. Bohmann, *Ubiquitin-dependent c-Jun degradation in vivo is mediated by the delta domain.* Cell, 1994. **78**(5): p. 787-98.
135. Matsui, M., et al., *Isolation of human fos-related genes and their expression during monocyte-macrophage differentiation.* Oncogene, 1990. **5**(3): p. 249-55.
136. Hirai, S., B. Bourachot, and M. Yaniv, *Both Jun and Fos contribute to transcription activation by the heterodimer.* Oncogene, 1990. **5**(1): p. 39-46.
137. Greenberg, M.E. and E.B. Ziff, *Stimulation of 3T3 cells induces transcription of the c-fos proto-oncogene.* Nature, 1984. **311**(5985): p. 433-8.
138. Vesely, P.W., et al., *Translational regulation mechanisms of AP-1 proteins.* Mutat Res, 2009. **682**(1): p. 7-12.
139. Angel, P., et al., *Oncogene jun encodes a sequence-specific trans-activator similar to AP-1.* Nature, 1988. **332**(6160): p. 166-71.
140. Kim, S.J., et al., *Autoinduction of transforming growth factor beta 1 is mediated by the AP-1 complex.* Mol Cell Biol, 1990. **10**(4): p. 1492-7.
141. Sheng, M., M.A. Thompson, and M.E. Greenberg, *CREB: a Ca(2+)-regulated transcription factor phosphorylated by calmodulin-dependent kinases.* Science, 1991. **252**(5011): p. 1427-30.
142. Darnell, J.E., Jr., I.M. Kerr, and G.R. Stark, *Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular*

signaling proteins. Science, 1994. **264**(5164): p. 1415-21.

143. Treisman, R., *The serum response element.* Trends Biochem Sci, 1992. **17**(10): p. 423-6.

144. Angel, P., et al., *The jun proto-oncogene is positively autoregulated by its product, Jun/AP-1.* Cell, 1988. **55**(5): p. 875-85.

145. Schonthal, A., et al., *The Fos and Jun/AP-1 proteins are involved in the downregulation of Fos transcription.* Oncogene, 1989. **4**(5): p. 629-36.

146. Hirai, S.I., et al., *Characterization of junD: a new member of the jun proto-oncogene family.* EMBO J, 1989. **8**(5): p. 1433-9.

147. Karin, M., Z. Liu, and E. Zandi, *AP-1 function and regulation.* Curr Opin Cell Biol, 1997. **9**(2): p. 240-6.

148. Musti, A.M., M. Treier, and D. Bohmann, *Reduced ubiquitin-dependent degradation of c-Jun after phosphorylation by MAP kinases.* Science, 1997. **275**(5298): p. 400-2.

149. Shaulian, E. and M. Karin, *AP-1 in cell proliferation and survival.* Oncogene, 2001. **20**(19): p. 2390-400.

150. Schreiber, M., et al., *Control of cell cycle progression by c-Jun is p53 dependent.* Genes Dev, 1999. **13**(5): p. 607-19.

151. Johnson, R.S., B.M. Spiegelman, and V. Papaioannou, *Pleiotropic effects of a null mutation in the c-fos proto-oncogene.* Cell, 1992. **71**(4): p. 577-86.

152. Brown, J.R., et al., *Fos family members induce cell cycle entry by activating cyclin D1.* Mol Cell Biol, 1998. **18**(9): p. 5609-19.

153. Johnson, R.S., et al., *A null mutation at the c-jun locus causes embryonic lethality and retarded cell growth in culture.* Genes Dev, 1993. **7**(7B): p. 1309-17.

154. Behrens, A., et al., *Impaired intervertebral disc formation in the absence of Jun.* Development, 2003. **130**(1): p. 103-9.

155. Passegue, E., et al., *JunB can substitute for Jun in mouse development and cell proliferation.* Nat Genet, 2002. **30**(2): p. 158-66.

156. Smeyne, R.J., T. Curran, and J.I. Morgan, *Temporal and spatial expression of*

a fos-lacZ transgene in the developing nervous system. Brain Res Mol Brain Res, 1992. **16**(1-2): p. 158-62.

157. Mikula, M., et al., *The proto-oncoprotein c-Fos negatively regulates hepatocellular tumorigenesis.* Oncogene, 2003. **22**(43): p. 6725-38.
158. Fleischmann, A., et al., *Rhabdomyosarcoma development in mice lacking Trp53 and Fos: tumor suppression by the Fos protooncogene.* Cancer Cell, 2003. **4**(6): p. 477-82.
159. Zhang, J., et al., *c-fos regulates neuronal excitability and survival.* Nat Genet, 2002. **30**(4): p. 416-20.
160. Hu, X.T., et al., *Fos regulates neuronal activity in the nucleus accumbens.* Neurosci Lett, 2008. **448**(1): p. 157-60.
161. Lasham, A., et al., *Regulation of the human fas promoter by YB-1, Puralpha and AP-1 transcription factors.* Gene, 2000. **252**(1-2): p. 1-13.
162. Ham, J., et al., *A c-Jun dominant negative mutant protects sympathetic neurons against programmed cell death.* Neuron, 1995. **14**(5): p. 927-39.
163. Bossy-Wetzel, E., L. Bakiri, and M. Yaniv, *Induction of apoptosis by the transcription factor c-Jun.* EMBO J, 1997. **16**(7): p. 1695-709.
164. Behrens, A., M. Sibilia, and E.F. Wagner, *Amino-terminal phosphorylation of c-Jun regulates stress-induced apoptosis and cellular proliferation.* Nat Genet, 1999. **21**(3): p. 326-9
165. .Shaulian, E., et al., *The mammalian UV response: c-Jun induction is required for exit from p53-imposed growth arrest.* Cell, 2000. **103**(6): p. 897-907.
166. Zhang, G., et al., *Effect of deoxyribozymes targeting c-Jun on solid tumor growth and angiogenesis in rodents.* J Natl Cancer Inst, 2004. **96**(9): p. 683-96.
167. Hu, E., et al., *Targeted disruption of the c-fos gene demonstrates c-fos-dependent and -independent pathways for gene expression stimulated by growth factors or oncogenes.* EMBO J, 1994. **13**(13): p. 3094-103.
168. Marconcini, L., et al., *c-fos-induced growth factor/vascular endothelial growth factor D induces angiogenesis in vivo and in vitro.* Proc Natl Acad Sci

U S A, 1999. **96**(17): p. 9671-6.

169. Alfranca, A., et al., *c-Jun and hypoxia-inducible factor 1 functionally cooperate in hypoxia-induced gene transcription.* Mol Cell Biol, 2002. **22**(1): p. 12-22.

170. Toft, D.J., et al., *Reactivation of proliferin gene expression is associated with increased angiogenesis in a cell culture model of fibrosarcoma tumor progression.* Proc Natl Acad Sci U S A, 2001. **98**(23): p. 13055-9.

171. Riabowol, K., J. Schiff, and M.Z. Gilman, *Transcription factor AP-1 activity is required for initiation of DNA synthesis and is lost during cellular aging.* Proc Natl Acad Sci U S A, 1992. **89**(1): p. 157-61.

172. Ito, K., et al., *c-Jun stimulates origin-dependent DNA unwinding by polyomavirus large Tantigen.* EMBO J, 1996. **15**(20): p. 5636-46.

173. Luo, L., et al., *Regulation of geminin functions by cell cycle-dependent nuclear-cytoplasmic shuttling.* Mol Cell Biol, 2007. **27**(13): p. 4737-44.

174. de Stanchina, E., et al., *Selection of homeotic proteins for binding to a human DNA replication origin.* J Mol Biol, 2000. **299**(3): p. 667-80.

175. Salsi, V., et al., *HOXD13 binds DNA replication origins to promote origin licensing and is inhibited by geminin.* Mol Cell Biol, 2009. **29**(21): p. 5775-88.

176. Krosl, J. and G. Sauvageau, *AP-1 complex is effector of Hox-induced cellular proliferation and transformation.* Oncogene, 2000. **19**(45): p. 5134-41.

177. Kataoka, K., et al., *A set of Hox proteins interact with the Maf oncoprotein to inhibit its DNA binding, transactivation, and transforming activities.* J Biol Chem, 2001. **276**(1): p. 819-26.

178. Diviacco, S., et al., *A novel procedure for quantitative polymerase chain reaction by coamplification of competitive templates.* Gene, 1992. **122**(2): p. 313-20.

179. Dai, S.M., et al., *Ligation-mediated PCR for quantitative in vivo footprinting.* Nat Biotechnol, 2000. **18**(10): p. 1108-11.

180. Komura, J. and A.D. Riggs, *Terminal transferase-dependent PCR: a versatile*

and sensitive method for in vivo footprinting and detection of DNA adducts. Nucleic Acids Res, 1998. **26**(7): p. 1807-11.

181. O'Neill, L.P. and B.M. Turner, *Immunoprecipitation of native chromatin: NChIP.* Methods, 2003. **31**(1): p. 76-82.
182. Quintana, D.G., et al., *Identification of HsORC4, a member of the human origin of replication recognition complex.* J Biol Chem, 1997. **272**(45): p. 28247-51.
183. Laurent, A., et al., *PBX proteins: much more than Hox cofactors.* Int J Dev Biol, 2008. **52**(1): p. 9-20.
184. Wang, W.M., A.Y. Lee, and C.M. Chiang, *One-step affinity tag purification of full-length recombinant human AP-1 complexes from bacterial inclusion bodies using a polycistronic expression system.* Protein Expr Purif, 2008. **59**(1): p. 144-52.
185. Ramirez-Carrozzi, V.R. and T.K. Kerppola, *Dynamics of Fos-Jun-NFAT1 complexes.* Proc Natl Acad Sci U S A, 2001. **98**(9): p. 4893-8.
186. Macian, F., C. Lopez-Rodriguez, and A. Rao, *Partners in transcription: NFAT and AP-1.* Oncogene, 2001. **20**(19): p. 2476-89.
187. Kovary, K. and R. Bravo, *Existence of different Fos/Jun complexes during the G0-to-G1 transition and during exponential growth in mouse fibroblasts: differential role of Fos proteins.* Mol Cell Biol, 1992. **12**(11): p. 5015-23
188. Kovary, K. and R. Bravo, *The jun and fos protein families are both required for cell cycle progression in fibroblasts.* Mol Cell Biol, 1991. **11**(9): p. 4466-72.
189. Paolinelli, R., et al., *Acetylation by GCN5 regulates CDC6 phosphorylation in the S phase of the cell cycle.* Nat Struct Mol Biol, 2009. **16**(4): p. 412-20.
190. Furlan-Magaril, M., H. Rincon-Arano, and F. Recillas-Targa, *Sequential chromatin immunoprecipitation protocol: ChIP-reChIP.* Methods Mol Biol, 2009. **543**: p. 253-66.
191. Drake, F.H., et al., *In vitro and intracellular inhibition of topoisomerase II by the antitumor agent merbarone.* Cancer Res, 1989. **49**(10): p. 2578-83.

192. Lubelsky, Y., et al., *DNA replication and transcription programs respond to the same chromatin cues.* Genome Res, 2014. **24**(7): p. 1102-14.

193. Picard, F., et al., *The spatiotemporal program of DNA replication is associated with specific combinations of chromatin marks in human cells.* PLoS Genet, 2014. **10**(5): p. e1004282.

194, Demeret, C., Y. Vassetzky, and M. Mechali, *Chromatin remodelling and DNA replication: from nucleosomes to loop domains.* Oncogene, 2001. **20**(24): p. 3086-93.

195. Rampakakis, E. and M. Zannis-Hadjopoulos, *Transient dsDNA breaks during pre-replication complex assembly.* Nucleic Acids Res, 2009. **37**(17): p. 5714-24.

196, Abdurashidova, G., et al., *Localization of proteins bound to a replication origin of human DNA along the cell cycle.* EMBO J, 2003. **22**(16): p. 4294-303.

197. Flashner, Y. and A. Shafferman, *Alpha and beta replication origins of plasmid R6K show similar distortions of the DNA helix in vivo.* Proc Natl Acad Sci U S A, 1990. **87**(23): p. 9123-7.

198. Saxena, M., M. Abhyankar, and D. Bastia, *Replication initiation at a distance: determination of the cis- and trans-acting elements of replication origin alpha of plasmid R6K.* J Biol Chem, 2010. **285**(8): p. 5705-12.

199. Sim, J., et al., *Remodeling of the human papillomavirus type 11 replication origin into discrete nucleoprotein particles and looped structures by the E2 protein.* J Mol Biol, 2008. **375**(4): p. 1165-77.

200. Kusic, J., et al., *Noncanonical DNA elements in the lamin B2 origin of DNA replication.* J Biol Chem, 2005. **280**(11): p. 9848-54.

201. Kusic, J., et al., *Human initiation protein ORC4 prefers triple stranded DNA.* Mol Biol Rep, 2009

202. Papior, P., et al., *Open chromatin structures regulate the efficiencies of pre-RC formation and replication initiation in Epstein-Barr virus.* J Cell Biol, 2012. **198**(4): p. 509-28.

203. Gilbert, D.M., *Replication origins run (ultra) deep.* Nat Struct Mol Biol, 2012. **19**(8): p. 740-2.
204. Besnard, E., et al., *Unraveling cell type-specific and reprogrammable human replication origin signatures associated with G-quadruplex consensus motifs.* Nat Struct Mol Biol, 2012. **19**(8): p. 837-44.
205. Valton, A.L., et al., *G4 motifs affect origin positioning and efficiency in two vertebrate replicators.* EMBO J, 2014. **33**(7): p. 732-46.
206. Norseen, J., et al., *RNA-dependent recruitment of the origin recognition complex.* EMBO J, 2008. **27**(22): p. 3024-35.
207. Christov, C.P., et al., *Functional requirement of noncoding Y RNAs for human chromosomal DNA replication.* Mol Cell Biol, 2006. **26**(18): p. 6993-7004.
208. Norseen, J., F.B. Johnson, and P.M. Lieberman, *Role for G-quadruplex RNA binding by Epstein-Barr virus nuclear antigen 1 in DNA replication and metaphase chromosome attachment.* J Virol, 2009. **83**(20): p. 10336-46.
209. Moriyama, K., et al., *Epstein-Barr nuclear antigen 1 (EBNA1)-dependent recruitment of origin recognition complex (Orc) on oriP of Epstein-Barr virus with purified proteins: stimulation by Cdc6 through its direct interaction with EBNA1.* J Biol Chem, 2012. **287**(28): p. 23977-94.
210. Hoshina, S., et al., *Human Origin Recognition Complex Binds Preferentially to G-quadruplex-preferable RNA and Single-stranded DNA.* J Biol Chem, 2013. **288**(42):p. 30161-71
211. Caudron-Herger, M., et al., *Coding RNAs with a non-coding function: maintenance of open chromatin structure.* Nucleus, 2011. **2**(5): p. 410-24.
212. Strumberg, D., et al., *Conversion of topoisomerase I cleavage complexes on the leading strand of ribosomal DNA into 5'-phosphorylated DNA double-strand breaks by replication runoff.* Mol Cell Biol, 2000. **20**(11): p. 3977-87.
213. Mueller, P.R., B. Wold, and P.A. Garrity, *Ligation-mediated PCR for genomic sequencing and footprinting.* Curr Protoc Mol Biol, 2001. **Chapter 15**: p. Unit 15 3.

Acknowledgments

I would to thank first my family and Roberta who encouraged me every time to go further any difficulties I encountered during this period.
Special thanks to Professor Mauro Giacca who gave me the possibility to come to the ICGEB to finish this work and supported me since 2010.
Many thanks to all the guys and lab mates I've met during this period both in Pisa and here, in Trieste, for each beautiful moment spent together, especially to Federico Cremisi.

Finally, a huge thank to Professor Arturo Falaschi, a master of life and science. His interesting in all issues, ranging from science to literature, his humanity and humility were a life style to follow. I will always keep all his qualities in my mind as model to follow.
As I know Professor Falaschi loved to tell “La Divina Commedia” as a bedtime fairy tale to his sons; for this reason I would like to end in this way:

“Ma quello ingrato popolo maligno
che discese di Fiesole ab antico,
e tiene ancor del monte e del macigno,

ti si farà, per tuo ben far, nimico:
ed è ragion, ché tra li lazzi sorbi
si disconvien fruttare al dolce fico.

Vecchia fama nel mondo li chiama orbi;
gent'è avara, invidiosa e superba:
dai lor costumi fa che tu ti forbi.

La tua fortuna tanto onor ti serba,
che l'una parte e l'altra avranno fame
di te; ma lungi fia dal becco l'erba.”

Dante Alighieri “La Divina Commedia” – Inferno Canto XV