Scuola Normale Superiore di Pisa

Ph.D. Thesis

Molecular Biology

Localization and dynamics of homeotic oncogenic protein HOXC13 in pre-initiation complex of human DNA replication origins

Laura Comelli

ADVISOR

Prof. Arturo Falaschi

2010

To Prof. Arturo Falaschi

Contents

Publications

Marchetti L; **Comelli L**; D'Innocenzo B; Puzzi L; Luin S; Arosio D; Calvello M; Mendoza-Maldonado R; Peverali F; Trovato F; Riva S; Biamonti G; Abdurashidova G; Beltram F; Falaschi A. "Homeotic proteins participate in the function of human DNA replication origins". Nucleic Acid Research 2010 Ago 6

Comelli L, Marchetti L, Arosio D, Riva S, Abdurashidova G, Beltram F, Falaschi A. "The homeotic protein HOXC13 is a member of human DNA replication complexes". Cell Cycle. 2009 Feb 1;8(3):454-9. Epub 2009 Feb 19

Abstract

In metazoan cells the DNA replication origins are not well defined. Differently from what observed for bacteria cells and for budding yeast, in metazoan the origins does not show a conserved sequence and they appear to be specified by many factors. In order to better understand the mechanisms involved in the origin specification, many studies have been done to identify the proteins involved in the recognition and activation of the origins. From these kind of analysis is emerging that, beside the well-known proteins of the pre replicative complex, also other factors might be involved. Between these, the HOX proteins seem to be able to play a role in the origin activity. One of the first studies of this involvement was done by our group and leads to the identification of three homeotic proteins able to specifically bind *in vitro* the human lamin B2 origin. Thus, in the study conducted during this PhD program, was investigated the involvement of one of these homeotic proteins, namely HOXC13, with human DNA replication origins and with replicative complexes.

We found an interaction of HOXC13 with two crucial factors of the pre Replication Complex (pre-RC), ORC1 and Cdc6 and that HOXC13 binds a good fraction of the origins, in particular the early replicating ones, like the lamin B2 origin and other known human origins. The HOXC13 protein is bound to origin chromatin, at least for the lamin B2 origin, at a precise site within the pre-RC at specific moments of the cell cycle. Interaction with the origin occurs within the area protected by the pre-RC in G1, very close to the start sites of leading strand synthesis and to the binding sites of ORC1, ORC2, Cdc6, topoisomerase (topo) I and topo II. The protein is absent from the origin in M and appears on it at the beginning of G1, reach a peak at G1/S and as synthesis starts, the interaction of HOXC13 with the origin fades, in parallel with the transition from this large pre-RC to a smaller and differently organized post-RC.

Recently also other HOX proteins have been identify as proteins involved in regulation processes of DNA replication, suggesting that the interaction of HOXC13 with the origins might occur in a multi-homeotic proteins complex. Depletion of one of these proteins however is compatible with the continuation of the cell cycle and, according with what observed for the other homeotic proteins, we found that also the depletion of HOXC13 does not alter cell cycle progression or S phase entry. This is probably due to the redundancy of homeotic proteins and indicates a relatively generic function for the HOX proteins.

Among the identified elements influencing the choice and the activity of a sequence as DNA replication origin, much relevance is assumed by the chromatin structure and topology of DNA. Therefore, we analysed the effects of chromatin structure disruption using Tricostatin A, a histone deacetylase inhibitor. The alteration of chromatin caused by this treatment not only sharply reduces origin function, but also disturbs the binding of replication complex members like HOXC13 and the well known Cdc6 to the DNA replication origins, while does not affect the binding of other unrelated proteins like USF1. On the basis of this finding, we infer that an appropriate chromatin organization and DNA topology strongly influence the binding between factors of the pre Replication Complex and DNA replication origins. This influence could be a key element in origin specification.

The described interactions are not restricted to a single origin nor to a single homeotic protein, leading us to conclude that HOX proteins, probably in the context of a multi-protein homeotic effectors, contribute to recruit and stabilize the replicative complexes onto early replicating origins, in presence of specific chromatin and topological configurations.

The relevance of HOXC13 in DNA replication is also underlined by its involvement in oncogenesis, clearly demonstrated in acute myeloid leukaemia when HOXC13 is fused with NUP98 protein.

1.Introduction

1.1. DNA replication

1.1.1. Eukaryotic DNA replication

Cells begin DNA replication from specifically selected chromosomal sites termed replication initiation sites or replication origins. The chromosomal DNA is subdivided into a number of tandemly organized replicons, ranging in number from ~400 in yeast to ~30,000 in humans. Each replicon contains a replication origin, a sequence that is the final target of specific proteins that lead to the recognition of the specific origin, unwinding of DNA, formation of replication forks and synthesis of new strands. The two opposite moving replication forks progress through the replication until they merge with those issued from the adjacent ones. In metazoans DNA replication is a tightly regulated process, ensuring that the genome is duplicated only once each cell cycle, before chromosome segregation and cytokinesis.

The initial event in DNA replication is constituted by the origin selection mediated by the binding of the Origin Recognition Complex (ORC) into origin DNA sequences. The ORC complex provides a "landing-pad" for other proteins including Cdc6, Cdt1, DNA helicase complex Mcm2-7 and others. All together these proteins form the so called pre-Replication Complex (pre-RC) that is formed in G1 phase of the cell cycle and marks the potential sites for the initiation of DNA replication. Each replication origin spaced apart from 50 to 250 kb, depending on the development stage, growth conditions and cell transformation status.

1.1.2. The Replicon model

The main principles characterizing the events that initiate the DNA replication are conserved in evolution and were described in replicon model, formulated in 1963 by Jacob and Brenner[1]. This model postulates the presence of specific cis-acting sequence elements, termed replicator, that genetically determine replication-initiation sites on DNA molecules. The replicators interact with trans-acting regulatory factors, termed initiators, able to recognize the replicators on the genome in response to appropriate

cellular signals (Figure1). The first event in the initiation of DNA synthesis is the chromatin decondensation and local opening of the duplex strands to provide the access for the initiators.

The replicon model was proposed to explain the regulation of the *Escherichia coli* DNA replication and provided a foundation to understand how the initiation of DNA replication occurs in all organisms. With this model two new elements were introduced: the existence of specific site where the double helix of DNA is opened and the involvement of proteins other than polymerases in DNA replication.

In bacteria the initiation of DNA replication at *oriC* requires origin binding by the initiator protein dnaA. This binding alters the structure of the origin and provides an associative platform for targeting additional proteins as dnaB, an helicase involved in the progression of the replication fork. DnaB also allows the loading of an other replication-fork enzyme called primase, which is a replication-priming RNA polymerase[2]. Together, the helicase and primase form the core of the so called "primosome"[3]. During the entire course of replication, the dnaB helicase catalyzes the unwinding of genomic DNA in an ATP dependent reaction[4]. This model was validated in numerous prokaryotic and viral systems.

The unwinding of the double stranded genomic DNA is energetically unfavorable, thus the cells have evolved helicases enzymes, that couple the energy of NTP binding and hydrolysis to the unwinding[5]. Families of helicases share several sequences and structural motifs, implying that there is a common unwinding mechanism used by these enzymes.

The first extension of this model from prokaryotes to eukaryotic chromosomes was the identification of Autonomously Replicating Sequences (ARS) in budding yeast. Analysis of the ARS revealed that these are about 100bp sequences and consist of a 17bp consensus A-domain region with an 11bp ARS Consensus Sequences (ACS) that is AT rich and flanked by poorly conserved B domains. The A and the B1 domains are binding sites for the proteins involved in DNA replication, while others B elements act as enhancers for the origin efficiency. The identification of consensus ARS elements has permitted the isolation of the Origin Recognition Complex ORC, the protein complex that binds to the origin sequences and allows the DNA replication in all eukaryotes.

Actually, in many eukaryotes the consensus sequence conservation at replicons may not be completely faithful because of the very little sequence specificity among

origins, differently from what observed in budding yeast. AT richness is generally retained feature and is assumed to be important for facilitate the opening of DNA strands at the origin. However, in the other eukaryotes as in fission yeast, the origins are generally longer, without any identifiable conserved consensus sequences analogous to ACS and in many cases, in contrast to the budding yeast, inefficient origins, that fire randomly, are spread in the genome[6]. Thus, *S. cerevisiae* appears to be an exception in the eukaryotes. The eukaryotic origins seem to be defined by different combinations of elements depending on the context that license them to be recognized by "initiators" including several features like DNA sequence, but there is not a single consensus combination that define an origin.

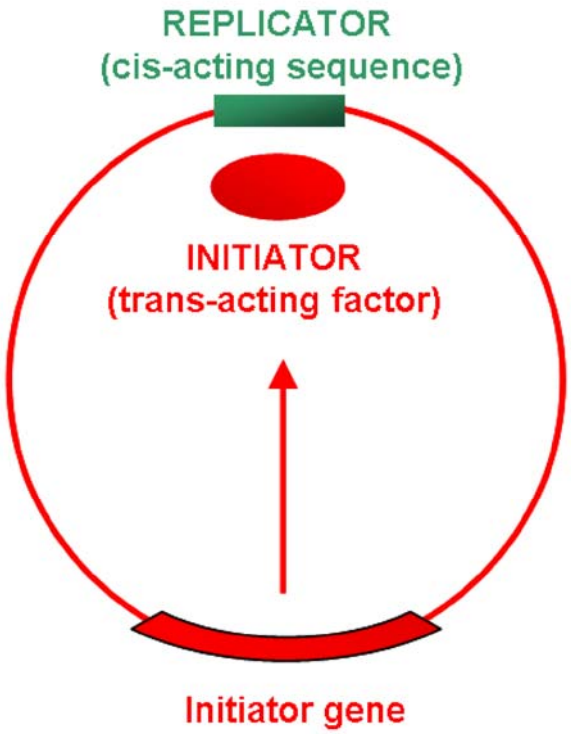

Figure 1. The replicon model. A trans-acting protein encoded by the initiator gene was proposed to recognize a cis-acting sequence (the replicator) that controls the initiation of DNA replication in the replicon.

1.1.3. Replication complex

The mechanism of initiation of eukaryotic replication is based on in the "origin licensing" model. In this model, origins are licensed once the pre Replication Complex (pre-RC)is entirely loaded onto them, and the ORC complex is the first to be involved in this assembly (Figure 2).

The ORC complex consists of six subunits (ORC1-6) discovered by specific isolation of ACS binding proteins in budding yeast. These proteins are conserved in evolution but their DNA specific binding dependence, present in budding yeast, is lost in the other eukaryotes. In *S. pombe*, ORC-origin binding to origins is mediated uniquely by ORC4, which is able to recognize and bind specifically AT-rich sequences through its AT-hook DNA binding motif[7]. This feature of ORC4 is limited to *S. pombe*. The specificity of mammalian ORC binding to DNA is very low, due to its limited ability to distinguish specific sequences[8]. Moreover, the difficulties in identifying well-defined ORC binding sites in species other than yeast raise the possibility that other DNA binding factors may contribute and facilitate ORC localization and origin selection. The human, frog and *S. pombe* ORCs preferentially bind DNA to AT rich tracts but there is not a clear consensus sequences among these regions. The AT regions are characterized by their helical instability that facilitates probably the DNA unwinding and their recognition as replication origins[9, 10]. Therefore, the chromatin structure, and not only the primary DNA sequence, might be the element recognized by ORC.

The ORC subunits are AAA^+ ATPases (ATPase Associated with various cellular Activities) but only the ATPase activity of ORC1 is required for DNA binding[11]. Contrary to yeast, in mammals some of the subunits of ORC complex are displaced from the origin site after initiation of DNA replication[12], suggesting a more dynamic interaction between mammalian ORC and origin DNA. In human cells, the ORC1 subunit in S-phase is selectively destabilized, ubiquitinated, partially degraded, and then stably bound to chromatin during the next M to G1 transition, to establish the pre-RC at specific genomic sites in G1 phase. Moreover, species-specific variations exist in the regulation of ORC activity, because although ORC proteins are highly conserved within a single taxonomic family, conservation among all species is modest[13].

The ORC1 ATPase is activated by the binding of Cdc6, an other AAA^+ ATPase, that induced conformational changes on DNA that increase the specificity of the binding of ORC-Cdc6 complex to origins. Moreover, Cdc6 ATPase activity determine its

dissociation from ORC-Cdc6-DNA complex, in non origin sequences, while inhibition of Cdc6 ATPase on origin DNA sequences results in a stable ORC-Cdc6-DNA complex, which can then promote MCM loading to origins. Therefore, Cdc6 ATPase activity regulates origin DNA sequence specificity for the assembly of the pre-RC, required for DNA replication initiation. In fact, mutations increasing Cdc6 ATPase activity result in a less stable complex on DNA[14].

After the origin binding by ORC-Cdc6 proteins, the next step is the loading of Cdt1. Duringthe M to G1-transition, in human cells Cdc6 is degraded and then resynthesized later during G1-phase[15].

Cdt1 has been shown to be a key element in the formation of the pre-RC, in particular, regulating the "once per cell cycle" replication feature. Cdt1 is periodically expressed under the control of the transcription factor Cdc10, which also controls the expression of Cdc6 in different species[16].

Moreover, the Cdc6 ATPase activity is seem to be required for Cdt1 binding to the origin, perhaps determining the following loading of MCM complex onto the ORC-Cdc6-DNA complex. Cdt1 activity is additionally regulated by differential stoichiometric binding of a DNA replication inhibitor, Geminin, and by proteolysis[17].

Geminin is a known as inhibitor of DNA replication that acts by preventing MCM loading onto origins throughout S and M phase and impeding unwanted additional firing events. This protein plays multiple roles in several fundamental cellular processes including proliferation, differentiation, development and transcriptional regulation. All these functions have been characterized by identifying Geminin binding partners[18]. Geminin was shown to interact with Cdt1 during the S phase, targeting it for degradation thereby preventing MCM loading until the following G1 phase and hence preventing re-replication[19]. The balances of the Geminin-Cdt1 association establishes the timing of DNA replication initiation and controls the cell cycle progression[20]. During S-G2/M phases, Geminin binds Cdt1 at increased stoichiometric ratios and suppress Cdt1 function. However, through pre-RC licensing, Geminin binds Cdt1 in a lower stoichoimetric ratio, allowing the presence of an active form of Cdt1 that can interact with Cdc6 and ORC[21]. The Cdt1-Geminin complex can exist in two distinct forms, as heterotrimer and as heterohexamer and the hexamer formation is critical for full inhibition of Cdt1 by Geminin. Probably the abbundance of heterotrimer and heteroexamer is regulated during the cell cycle[22]. The levels of Geminin rise during S phase and its degradation is mediated by the anaphase promoting complex (APC) in

mitosis or by a nonproteolytic inactivation of the fraction of Geminin that escaped from degradation.

Localization of Cdt1 to the origin is essential for MCM complex recruitment. Then, Cdt1 and Cdc6 dissociate from the origin and ATP hydrolysis by ORC completes the MCM loading[23].

The MCM genes were first identified in mutants defective on the maintenance of mini-chromosomes in budding yeast (Mcm phenotype)[24]. A subset of these MCM mutations were found in a family of six paralogous genes numbered MCM from 2 to 7, which are highly conserved in eukaryotes. The MCM complex consist of six AAA^+-ATPase subunits MCM 2-7, structured in a ring-shaped conformation around the chromatin[25]. After the initial recruitment of MCM to the origin, Cdc6 ATPase activity allows the disassociation of Cdt1 from the pre-RC, an important step that allows the MCM ring to close around the DNA[23]. This ATPase activity is followed by the ORC ATPase activity that completes the MCM loading reaction and may promote further rounds of MCM loading. Thus, the cooperation between replisome loading and helicase activation ensures coordinated replication of the two strands of DNA.

The assembly of pre-RC is not sufficient to initiate DNA replication, because the replication process requires the activation of two S-phase promoting kinases, CDKs (Cyclin-Dependent Kinases) and DDK (Dbf4-Dependent Kinase,Cdc7) during the G1/S transition. CDKs and DDK phosphorylate and activate pre-RCs promoting the loading of other factors like MCM10, Cdc45 and GINS, thereby forming the pre-IC (pre-Initiation Complex) and triggering origin firing[26, 27]. The large multiprotein complex formed after the licensing is called pre-Initiation Complex (pre-IC), from which start two diverging semi-conservative replicative forks that progress in opposite directions until converge with the replicative forks coming from the adjacent origins.

Phosphorylation of ORC, Cdc6, Cdt1 and MCM proteins also results in their inactivation and/or disassociation from the origin, nuclear export and proteasomal degradation; and finally establish post-RC (post-Replicative Complex), a state required to prevent origin re-licensing[21,28-30].

The CDK and DDK enzymes are independently regulated by similar mechanisms. Both kinases subunits are inactive as monomeric form and are activated by binding of Cyclins for CDK and by Dbf4 for DDK. Recent evidence indicates that the Mcm2-7 complex is a target of phosphorylation by DDK and this event is needed for loading of

the Cdc45 protein[31]. Hence, the cooperative action of these kinases trigger the recruitment of replication proteins necessary for origin unwinding and DNA synthesis.

In particular, Cdc6 is released from origins by the cyclin dependent protein kinase, CDK2/Cyclin A, and replaced by Cdc45 upon the concert action of Cdc7/Dbf4 and CDK2/cyclin E complex kinases. Then Cdc45 associates with DNA polymerase α and DNA primase and leads to initiate RNA-primed DNA synthesis.

While DDKs seem to act on MCMs, CDKs appear to play a direct role in preventing the assembly of new pre-RCs. Since CDK activity remains high from S phase to the end of the following mitosis, re-licensing cannot occur until the beginning of the next cell cycle[32]. At least three members of the pre-RC (ORC, Cdc6 and MCMs) are phosphorylated by CDKs to prevent pre-RC assembly and re-replication. Moreover, CDKs have also been implicated in controlling the time of replication initiation at specific origins[26]. In this way the activation and inactivation of the proteins restrict DNA replication to once and only once per cell cycle.

The fact that assembly of pre-RC is delayed until mitosis is complete and a nuclear membrane is assembled, suggest that nuclear structure plays a role in the initiation of DNA replication. Chromatin is looped into domains by attachment of the chromatin fiber to the nuclear matrix. DNA sequences that bind preferentially to nuclear matrices are named Matrix Attachment Region (MAR) or Scaffold Associated Region (SAR), which are supposed to mediate this loop formation in vivo. MARs are about 200 bp long, AT-rich and contain topoisomerase II consensus sequences. Moreover, they are often found near to cis-acting regulatory sequences, and their binding sites to the nuclear matrix are abundant (greater than 10,000 per mammalian nucleus). Any DNA sequence, including the origins of replication, could be attached to the matrix at a certain time during the cell cycle. Indeed, a large body of evidence indicates that DNA replication occurs on nuclear matrix but origins are not permanently attached with the nuclear matrix suggesting a dynamic association during the cell cycle[33].

Homologues of the indicated proteins were identified in all eukaryotic organisms, but the attempts to define an origin consensus sequence, or at least common specific features of the origin sequences, lead to the conclusion that no sequence specificity and no origin consensus is identifiable. However, DNA replication starts from specific sites and with a defined spatial programme. The preparation of this programme is crucial for the maintenance of genome integrity, because the disruption of pre-RC formation leads to genome instability[34].

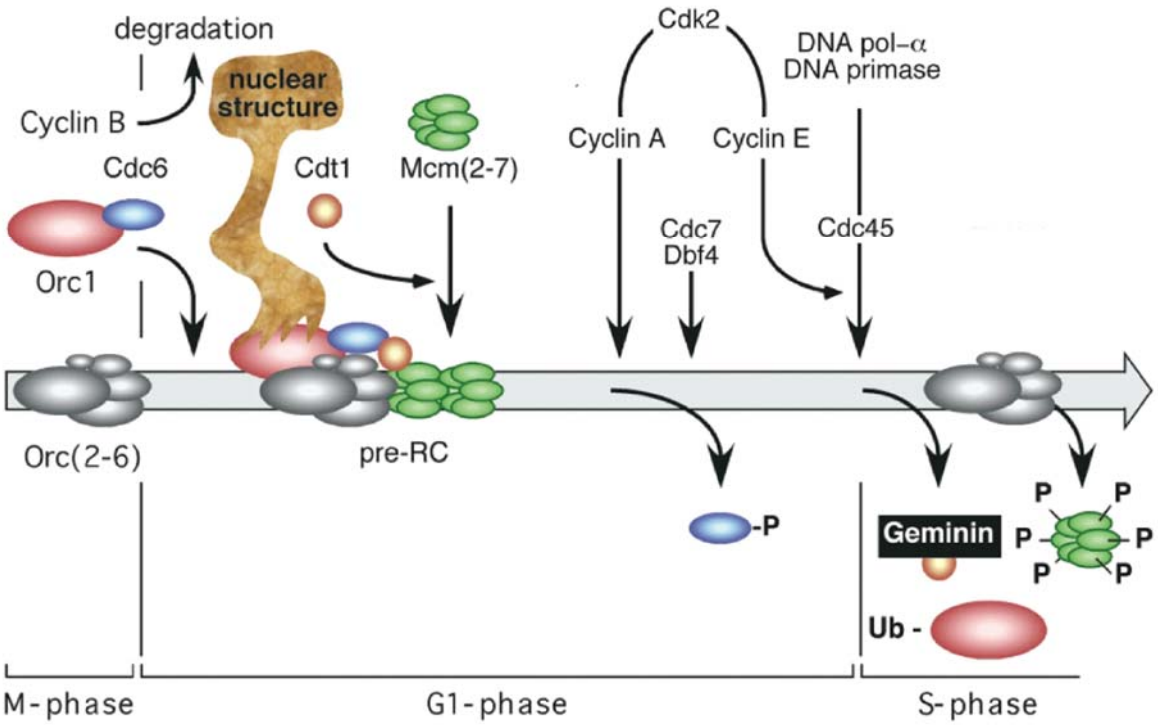

Figure 2. DNA replication in mammalian cells. ORC binds to DNA sequences recognised as replication origins. ORC exists in M phase as a stable complex of Orc(2–5) subunits, then ORC1 subunit is selectively bound to ORC complex during the M to G1 transition; the remaining ORC subunits remain stably bound to chromatin throughout the cell cycle. Cdc6 associates with ORC1, and this ORC1-Cdc6 complex is present in mitotic cells. In G1 phase, then Cdt1 associated to Cdc6 and loads at least one Mcm(2–7) hexamer per replication fork. In mammals, only ORC1 is associated to nuclease-resistant nuclear structure during G1-phase and then is released from origin previous to DNA replication initiation. Behind pre-RC assembly, DNA replication is initiated by the sequential loading of Mcm10, Cdk2/cyclin A that phosphorylate and release Cdc6, Cdc7/Dbf4 and Cdk2/cyclin E that subsequently modify other members of pre-RC and finally allow Cdc45 to bring polymeraseα and DNA primase to the replication origin DNA, the later enzyme initiates synthesis of the first RNA-primed nascent DNA strands. This event marks the beginning of S-phase. Concomitant with DNA synthesis is taking place the inactivation of Cdt1 by Geminin and the phosphorylation of MCM proteins. Cdc6 and Cdt1 are then released from chromatin and eventually degraded. MCM proteins remain in the nucleus where they are weakly associated with chromatin.

1.2 Eukaryotic origins of DNA replication

1.2.1. Replication origins

In multicellular organisms the origins of DNA replication are still not well defined. The main difference between prokaryotic and eukaryotic replication origins resides in the genome became larger in eukaryotes and the replication of chromosomal DNA depends on the activity of many different origins, that have to be activated with fine coordination in space and time[35, 36]. In fact, in higher eukaryotes, respect to the yeast, the genome size is about 100 times larger and this requires a bigger number of origins to ensure complete DNA replication. The increased number of origins also makes their mapping more difficult.

The discovery of the hexameric ORC complex in all eukaryotes from budding yeast to man now provides a strategy to look for the common features of origins of DNA replication. Although sequences bound by ORC complex could define the origins, differently from what observed for yeast, in *Drosophila* and human, the ORC complex have little or not sequence specificity binding activity[37]. The flexible sequence specificity observed in multicellular eukaryotes, respect to bacteria and *S. cerevisiae,* probably allows more dynamic changes in the organization of the genome and its replication.

If ORC complex does not preferentially bind specific sequences, probably other elements drive ORC to the origins. One possibility is that other proteins interacting with ORC carry it to a specific sequence. This might occur through the direct interaction of ORC with chromatin remodelling complexes or proteins that alter the local topological state of DNA, as occur for Epstein-Barr virus (EBV). The EBV origin of plasmid replication (OriP) provides an interesting model to study ORC recruitment in human cells. EBV is a human herpesvirus that establishes latent infections in multiple cell types and this latent form exists as an episomal minichromosome that replicate once per cell cycle and segregates similar to the cellular chromosome[38]. The EBNA1 protein of the virus interacts with ORC and with other cellular factors to recruit ORC and contribute to OriP replication activity[39]. By the same approach, it was found that not only the interaction with other factors, but also several features of the chromosomes, influence the firing of replication origins, as observed for methylation of DNA, that blocks the ORC binding and the initiation of DNA replication in *X. laevis* egg extract[40]; or histone

modifications that influence the activity of replication origins. The same was shown for *Drosophila* ORC binding, that is inhibited by the DNA methylation on GC nucleotides[40]. In addition, *Drosophila* ORC prefers negatively supercoiled DNA, proving that the DNA topology could be a more important determinant than DNA sequence for ORC binding[8].

Thus, even if the proteins involved in eukaryotic DNA replication are conserved, the sequences on which they bound are highly divergent because metazoan ORC complex exhibit no sequence binding specificity.

Despite the inability of investigators to isolate an autonomously replicating sequence in mammalian systems, approximately 20 mammalian origins have been identified. Mammalian origins could be classified in two groups, one containing regions referred to as zone of initiation, where replication begins from one or several potential sites within a large region of DNA, and the other group including origins where replication initiates from a localized site in each cell cycle.

Many examples of initiation zones exist, including the human rRNA locus, the Chine hamster rhodopsin and the DHFR (dihydrofolate reductase) loci. The best characterized is the Chine hamster ovary DHFR locus, with at least three primary initiation site (ori β, β' and γ) that account for most of the initiation events in this region[41]. Analysis of different DNA segments to drive initiation of replication when placed in different chromosomal context, revealed that in high eukaryotes replication requires specific DNA sequences that are both close to and distant from the site of initiation. Therefore, local sequence alterations can either enhance or repress origin activity, but no single consensus DNA-sequence motif is necessary or sufficient for replicator activity[42]. Moreover, by two-dimension gel mapping of a single copy of DHFR locus indicated that replication begins from multiple sites spanning 55 Kb intergenic region between DHFR and the adjacent locus. Preference for initiation was seen inside the central 35-40 Kb region, know to contain oriβ and oriγ[43]. This study suggested that mammalian initiation zone is composed of a primary initiation site coupled with multiple lower frequency sites.

An example of mammalian origin where firing is restricted to a circumscribed site is represented by the human β-globin origin. Originally, initiation of replication from this locus was thought to start from a single bidirectional origin of replication. Thn, more detailed studies indeed revealed that the locus is actually composed of two non

overlapping genetic elements that have been described to behave as redundant replicators when β-globin origin is assayed as an ectopic chromosomal site[44].

An other example of mammalian replicators is the lamin B2 origin, the best characterized among all isolated human DNA replication origins. It was mapped on chromosome 19, in the 3' untraslated region of the lamin B2 gene and consists of an AT rich region of 70-100 bp that contains 11 bp directed repeats, in one of which a bidirectional initiation site was precisely mapped[45]. Lamin B2 origin displays a cell cycle dependent footprint that is longer at G1 and shrinks as cells progress into S phase[46]. Inside this footprint area the transition between leading and lagging strands synthesis has been mapped to a single nucleotide level[45]. Moreover, it was demonstrated that cell cycle dependent footprints are in part due to the pre-RC binding. The chromatin surrounding the lamin B2 origin in G1 phase is bound by Cdc6, ORC1, ORC2 and MCM3, in S phase was found to bind by ORC2 and in M phase none of these pre-RC members were detected[47]. This data suggests that mammalian origins are bound by the same replication machinery as found in other eukaryotic systems, and this follow similar cell cycle regulation as in yeast. Additionally, this origin remains functional when transferred to other positions of the genome. Indeed, a DNA segment from lamin B2, comprising the larger footprint region, the start site of DNA replication and a CpG island, displays origin activity when moved to different chromosomal positions. This data confirm that this region contains a mammalian replicator and support the idea that sequence elements close to the replication start site play an important role in origin activation[48].

The ORC binding activity was also found in other replication regions like the TOP1 origin, which is localized in a GpC island usptream the human TOP1 gene on chromosome 20[49], and like the MCM4 origin that is located in a genomic region overlapping the control elements of the divergently transcribed genes MCM4 and PRKDC[49].

Thus, from all these examples of eukaryotic replication origins emerge that alternative factors could be responsible for ORC complex recruitment to the origins, and most probably different mechanisms of regulation are involved in their activation.

1.2.2. Origins spatio-temporal organization

The density of the origins is variable, with a minimum of no origin in 500 kb and a maximum of one origin every 11 kb, as the case of HoxA locus. Assuming that most origins have similar efficiency, the conclusion is that the genome is replicated by extremely variable replicons. In some loci, several origins located close together may be activated on the same DNA molecule at the same time[50].

The isolated origins posses a wide range of time for the activation, and it seems to act a sort of temporal programme for timing of origin firing, in which the density in origin is important but not sufficient to determine the replication time and probably the chromosomal environment plays a crucial role in this organization[50].

It was also observed that DNA sequences encompassing multiple replicons result to be less sensitive to replicative stress, and important to maintain genome stability. On the contrary, those regions with less origins are most sensitive probably due to the lower probability to use a backup licensed origin, located in between of two collapsed replication forks.

There is a considerable plasticity in the process of origin activation, because in cells whit many potential replication origins, most of them are not systematically activated in all cell cycles and only a small percentage are used, while the other remain "dormant". The dormant replication origins can be activated when DNA replication is slowed down. The usage of replication origins in fact, can change under different circumstances, like changes in chromatin organization that directly affect the way in which the origins are used in the subsequent cell cycle[51]. Also the presence of stalling replication forks, due to the damaged bases on DNA, can lead to activation of dormant replication origins. The dormant origins are activated as consequence of rapid and transitory response to several changes, and not affect the long-term behavior of cells. Their firing is like an adaptation mechanism of the genome to ensure the efficient DNA replication.

This adaptation correlates both with changes in chromatin organization and association of replication origins with the nuclear matrix[51]. Adjacent replicons are replicated together in "factories" with all the DNA replicated in a single factory co-localized within the nucleus[52]. This organization probably reflects the attachment of specific DNA sequences to an insoluble nucleoskeleton or nuclear matrix, creating in this manner chromatin loops that delineate functional defined units of transcription and

replication. Moreover, it is known that there is a relationship between the size of DNA loop that appear to be tethered to the nuclear matrix and the average spacing between replication origins.

From further observations of discrete sites within the nucleus detected by BrdUrd incorporation in mammalian cells at the beginning of S phase appear globular and constant in size but then grow and appear to form ring structures, it was concluded that the replication takes place at discrete sites in the nucleus and many replicons in a cluster might replicate in close proximity. These sites are called replication foci and indeed represent regions of DNA containing many replicons[53]. All studies to date demonstrate that in early S-phase replication sites are distributed throughout the nucleoplasm, with the exception of regions occupied by heterochromatin or nucleoli, and appear as discrete sites of variable size and number. Moreover, foci replication patterns and size of foci, change during S phase. In fact, in early S the number of the foci is bigger and their size is smaller than those found in mid and late S phase. This suggests that origin firing is programmed by a different temporal activation during S phase. In budding yeast, for example, the origins are activated principally in mid S phase, but the activation is visible throughout S phase[54].

The decision point for the activation of specific origins in metazoan cells is the early G1 phase. Therefore, transcriptionally active euchromatin replicates earlier respect the inactive heterochromatin, probably because the transcription open the chromatin and allows easy access to replication factors.

1.2.3. Origin activity and transcription

The relationship between transcription and replication has been documented in viral DNA genomes, in which transcription factors participate in the initiation of viral DNA synthesis by interacting with the viral initiator[55], but the involvement of transcription in replication and in regulating temporal activation of origins in eukaryotes is still unclear. Several lines of evidences indicate that transcription itself is not required for origin function in eukaryotes, but important is the competition between transcription and replication. In fact, during S phase in mammalian cells, there is an evident separation of the active replication sites and active transcription sites[56], indicating that transcription processes are stopped during replication. Moreover, it was observed that

the β-globin origin is activated both in transcriptionally active and inactive cells, and some kind of interferences between transcription and replication can occur when these two processes are not in the same orientation. The participation of transcriptional elements in site specific initiation of DNA replication might be due to the fact that they are part of the organization of chromatin domain competent for transcription[57].

However, replication origins tend to be associated with coding genes and the known origins are frequently distributed close the gene promoters. Studies of both coding and non-coding DNA regions, have shown that origin density and gene density are strictly correlated and reflect the coordinate organisation of replication and transcription. This finding suggests that these two processes may have regulatory factors in common, indeed it was also observed a strong association between origins and CpG islands. CpG islands are genomic regions bound by many transcriptional factors and the role suggested for them is the direct or indirect recruitment of pre-RC members for regulation of origin selection. In mouse the origins associated with CpG islands result to be more efficient because recruitment of the pre-RC proteins is favoured at this sites. Moreover, the presence of transcriptional factors at origin sites appear to stimulates replication also in others organisms[58, 59]. This stimulation may be a consequence of their direct interaction with components of the replication machinery or recruitment of chromatin remodelling complexes that facilitate the access of the replication complexes to DNA. Nevertheless, not all active promoters are efficient sites for DNA replication, so the active origins associated with transcription regulatory elements have to contain additional information. In fact, the sequences containing known origins remain strong sites of DNA replication also when inserted ectopically, indicating that specific replication starts point are recognized by replication machinery[48, 60]. Due most probably by recognition of specific combinations of transcription factors.

The CG rich regions seem to be important also for define a timing program. Several studies showed that CG rich regions tend to replicate earlier than CG poor regions, revealing a strong correlation between regional CG content and density in origins. In fact, late-replicating DNA, coincide with AT-rich regions on metaphase chromosomes with low transcriptional activity, while early-replicating DNA coincided with GC-rich regions with high transcriptional activity. Regions lacking origins, typically with a low content of CG, are replicated passively and hence relatively late, contrary to regions with high density of origins, that are rich in CG and replicate earier. However, in same studies, other regions with high density of origins were not early

replicated. The strong correlation observed between origin density and GC richness is a consequence of the strong association between origins and promoters or more distal transcriptional regulatory elements. Thus, as already said, the CG content and the density in origins are not predictive of replication timing and a major role for determining a precise spatio-temporal program is probably to ascribe to the chromosomal environment[50, 61].

Replication origins are found also in non open chromatin regions, probably bound by transcriptional factors that do not alter the chromatin structure. So, as for the transcription, probably different combinations of transcriptional factors are involved in the regulation of replication initiation sites. In this context, the lack of a consensus sequence in DNA replication origins could be explained[62].

So, replication in higher eukaryotes is clearly initiated by the interaction of an initiator with a replicator, as the replicon model proposed, but the replicator is defined by several features, between which the DNA sequence constitutes only one of the elements that influence where DNA replication starts, and the initiator is composed by all elements involving in the formation of the replication complex.

1.2.4. Chromatin structure influence on origin activity

Since is not possible identify a consensus sequence for origins, it was considered the possibility that the chromatin environment and modifications may play a central role in DNA origin selection. In fact, multiple structures contribute to eukaryotic replicator activity, suggesting that both DNA sequence and chromatin packaging influence replication initiation.

The basic units of chromatin are the nucleosome, octamers of histones, composed by two copies of each H2A, H2B, H3 and H4, around which DNA spools with about 146 bp. They form nucleosomal structures that coiled up to form a fibre of 30 nm diameter stabilized by linker histone H1. Iterative folding of these fibres generates the mitotic chromosomes. The packaging of eukaryotic genomes into nucleosomes and higher order chromatin structure seem to limit the access of replication factors to DNA. Now, origins of DNA replication might be delineated by poorly defined epigenetic factors, therefore additional information regarding the genetic marks contained within the chromatin associated proteins related to origins, known as "epigenome", is absolutely required[63].

A connection between the replicons and the chromatin organization came by the observation that proteins involved in DNA replication could also be involved in the assembly of specific chromatin domains. In fact, the amino terminal portion of Orc1 in higher eukaryotes is able to associate with heterochromatin-associated protein HP1, which contains chromo-domains and is known to be involved in heterochromatin formation. This binding indicates that ORC might also interact with other chromo-domain proteins to allow the assembly of replicons in euchromatin. Members of the chromo-domain family are the Polycomb group proteins that are involved in chromatin organization. The amino terminal portion of Orc1 contains also a bromo-adjacent-homology domain, conserved from yeast to human, also present in proteins involved in epigenetic regulation of transcription[57].

The chromatin environment influences both replication timing and frequency of origin activation[64]. Moreover, a close relationship has been observed between replication timing, chromatin structure and transcriptional activity. It was shown that early replicating genes could be either expressed or silent, while late replicating genes were almost always silent. This correlation between early replication timing and transcription was validated in higher eukaryotes also using various types of microarray

approaches. Additionally, this approaches showed a correlation between the transcriptional activity and the chromatin state. Chromatin indeed, exists in a decondensed or condensed state, called euchromatin and heterochromatin respectively. The euchromatin contain either actively transcribing genes or potentially active ones, and the heterochromatin is transcriptionally silent.

A regulatory role of chromatin structure in DNA replication was suggested also by the effect of chromosomal position on origin activity in *Drosophila*. It was observed that euchromatic domains generally replicates early in S phase whereas heterochromatic regions replicates later on[65].

1.2.5. Histone modification and Trichostatin A

Regulation of chromatin structure occurs through post-translational modifications of the histones tails, including acetylation, methylation, phosphorylation, ubiquitylation and sumoylation. Post-translational modifications of histone tails are largely investigated and also multiple modifications in the structured globular domains of histones are recently analysed. These modifications can generate different interaction affinities for chromatin-associated proteins and enable a dynamic chromatin state in which diverse nuclear processes can occur systematically[66].

Two of the most studied modifications are the acetylation and deacetylation of lysines on the core histones, which are controlled by histone acetyltransferases HATs and histone deacetylases HDACs respectively. The acetylation state of a chromatin locus results from the activities of HATs and HDACs on the nucleosomes. The reversible acetylation of the N-terminal tails of histones is a prominent chromatin modification that is thought to alter the degree of chromatin compaction. Typically, histone acetylation correlates well with increased DNA access, while histone deacetylation, and also the histone metylation, correlate with the formation of transcriptional silent chromatin. Thus, the result of histone acetylation is a change in chromatin structure and a corresponding increase in the accessibility of the DNA by trans-acting factors. The enzymatic activities that are required to the access to chromosomal DNA then would take benefit from chromatin modification by histone acetylation. An example of this mechanism might be the case of the HAT that binds to pre-RC complex, HBO1 (histone acetyltransferase binding to ORC1),this enzyme is a

MYST domain protein, characterized by a highly conserved zinc finger and a putative histone acetyltransferase domain. Data showing association with pre-RC components such as ORC1 and MCM2, suggest a role for this histone acetyltransferase protein in DNA replication. The presence of this HAT around the origins opens the possibility that this factor is recruited to the pre-RC by multiple proteins interactions and propose an active process in which chromatin is remodelled by replication initiators[67, 68].

Alterations of HDACs activities were identified in tumor cells and contribute to the massive perturbations of gene expression in numerous tumours. HDAC inhibitors leads to differentiation, cell cycle arrest and apoptosis in tumour cells and in some cases, prevents tumour growth. The most know potent inhibitors of the HDACs is Tricostatin A (TSA), which belong to the group of hydroxamic acids and is a natural product isolated from *Streptomyces platensis*. Crystallographic analysis indicate that TSA interacts reversibly with the HDAC catalytic site preventing binding of the substrate[69].

Histone deacetylase inhibitors in general, represent a new class of targeted anti-cancer agents. Several of these compounds are in clinical trials with significant activity against a spectrum of both hematologic and solid tumors at doses that are well tolerated by the patients[70]. One of the most effective and well studied HDAC inhibitor is the TSA itself.

The importance of chromatin structure in spatially and temporally regulation of DNA replication initiation was analysed using TSA treatment itself. The level of histone H4 acetylation correlates with the frequency of replication initiation, as measured by the abundance of short nascent DNA strands, mostly in the human c-myc and lamin B2 origins, and quite less with the frequency of initiation across the β-globin locus. Cells treated with TSA result in a reversible increase of the acetylation level of histone H4, both globally and locally to initiation sites at origins. In all three origins, TSA treatment transiently promoted a more dispersive pattern of initiations, decreasing the abundance of nascent DNA at previously preferred initiation sites. When cells arrested in late G1 were released into TSA, they completed S phase more rapidly than untreated cells, possibly due to the earlier initiation from late-firing origins. Thus, histone deacetylation might modulate replication origin activity through its effects on chromatin structure, by changing the selection of initiation sites, and promoting DNA synthesis at dormant initiation sites[71].

Other published data regarding TSA cell treatment lead to the same observation that pattern of initiation site selection in a replication loci was altered and becoming more dispersive. Thus, preferred initiation sites become less active while the less frequently used initiation sites become more active after treatment with TSA. It was also shown that the β-globin origin was induced to initiate DNA synthesis earlier in S phase after treatment with TSA. Since that in c-myc origin, the redistribution of MCM proteins was altered after TSA treatment, it was suggested that histone acetylation is a temporally upstream event leading to pre-RC formation or that pre-RC formation responds to other effects of TSA[72].

Moreover, the observation that origin identity can change during development, strongly supports that epigenetic regulation is central in origin selection. Chorion genes of *Drosophila* have been employed to analyse the relationship between chromatin modification and origin activity in metazoan. In this organism the somatic follicle cells undergo a developmental transition from genomic replication to continuous re-replication at 4 different chorion locus, two of them were identified on the X and 3rd chromosomes. Nucleosomes at these chorion origins were found to be hyperacetylated. These epigenetic modifications of chromatin contribute to different origin usage during the development. In this particular case, widespread acetylation might allows DNA access to additional origin binding proteins thereby impaired origin activity, that then results in a redistribution of the origins[73]. This observation is consistent with some observation in yeast, indeed mutation in a histone deacetylase resulted in advanced activation of late firing origin, and interaction of HBO1 HAT with replication proteins.

Interestingly, another component of the pre-Rc complex, MCM3, is endogenously acetylated and the acetylated MCM3 form is strictly chromatin-bound in late G1 phase. Moreover, MCM3 associated protein (MCM3AP), a protein isolated by two-hybrid screening using MCM3 as bait, is a specific MCM3 acetyltransferase of the GNAT superfamily[74]. The acetylase activity on MCM3 by MCM3AP is required to inhibit initiation of DNA replication and the association of MCM3AP to chromatin alone is not sufficient for this inhibition. The interaction between MCM3 and MCM3AP is essential for nuclear localization and chromatin binding of MCM3AP. Hence, MCM3AP is a potent natural inhibitor of the initiation of DNA replication whose action is mediated by interaction with MCM3[74].

Thus, chromatin acetylation seems to have an important effect on origin identity and activity, probably by relaxing chromatin and allowing proteins to gain access to the

origin binding sites. However, the precise step where this modification influences the origin activity remains still unclear. Nevertheless, in the chorion locus acetylation coincides with ORC binding to newly synthesied fibres and chromatin hyperacetylation determines redistribution of ORC2, suggesting that acetylation regulates DNA recognition by ORC.

1.2.6. Chromatin remodelling factors

In addition to histone modifications, nucleosome repositioning is involved in general chromatin remodelling events. Cell cycle changes in histone and in chromatin at eukaryotic origins are important regulatory feature controlling replication and access of licensing factors to DNA. Consistent with this, it was shown that a depletion of the ATP dependent chromatin remodelling complex ACF-ISWI, delayed progression of replication in late stage of S phase. ACF (ATP-utilizing chromatin assembly and remodeling factor) catalyzes the ATP-dependent assembly of periodic nucleosome arrays in vitro, and consists of ACF1 and the ISWI ATPase. ACF1 and ISWI are also subunits of CHRAC (chromatin accessibility complex), whose biochemical activities are similar to those of ACF. ACF1 forms a complex with the SNF2H isoform of ISWI and this complex is localized to pericentromeric heterochromatin during DNA replication. The depletion of this complex impairs the replication of pericentromeric region indicating the requirement of this complex to enables DNA replication through highly condensed regions of heterochromatin in mammalian cells[75]. The replication defect of ACF1 depleted cells was rescue by 5-aza-2-deoxycytidine treatments, which causes decondensation of heterochromatin by inhibition of DNA methylation. Although it is not known whether SNF2h alone plays a role in chromatin remodeling at replication origins or forks, it is worth noting that SNF2h is also recruited to remodel chromatin at the Epstein-Barr virus origin where host cell initiation machinery is utilized[76]. Also in *Drosophila* it was shown that chromatin-remodelling protein is implicated in promoting DNA synthesis and chromatin formation. Biochemical experiments with ACF1-deficient embryo extracts further indicate that ACF/CHRAC is a major chromatin assembly factor in *Drosophila*. The phenotypes of flies lacking ACF1 suggest that ACF/CHRAC promotes the formation of repressive chromatin[77].

Nucleosome remodelling factors are known to regulate chromatin structure at nucleosomal level by ATP-dependent alteration of the interaction between histones and DNA[78]. All these factors have ATPase subunits of the SWI-SNF superfamily. The ATPase subunits are related to helicases and are the key to understand the mechanism of remodelling. Probably, the interaction of the ATPase with the nucleosome, combined with directional translocation of the DNA, leads to a shift of DNA segments relative to the histone surface. To induce the sliding of nucleosomes the ATPase requires the presence of the basic amino acids at histone 4 (H4) tail. The same residues are necessary for the folding of the nucleosomal fibre indicating that nucleosome-remodelling factors directly interact with the histone determinants of fibres folding[79].

1.2.7. Role of DNA topology

Also the topological state of the origin DNA play a fundamental role in regulation of DNA replication. It was precisely demonstrated that the topoisomerase I and II interact with sequences bound by the replication proteins and close to the start site in the lamin B2 origin of replication[80]. Topoisomerases are enzymes that can modify the tertiary structure of DNA without changing its primary structure. The need for these enzymes comes from the structure of DNA itself, because during the elongation the two fork complexes unwind the double helical accumulating positive supercoils and tortional stress. When two replication forks, from adjacent replicons, converge and finish replicating the DNA segment, termination occurs. At this stage the two daughter double helices need to de separated to ensure a proper sister chromatids segregation. Topoisomerases are therefore required to rapidly relax these supercoils and allow progression of the replication fork. Topoisomerases must also function to ensure that DNA strands are completely unlinked and replicated chromosomes can be segregated to daughter cells.

In the Lamin B2 origin, topoisomerases I and II selectively bind to ORC2 binding sites in a cell cycle dependent manner. Additionally, ORC2 interacts with topoisomerase II at G1 phase and with topoisomerase I at the G1/S transition. Moreover, these particular protein interactions occur when they are bound to DNA replication origin, indicating a close association with replication complex. Indeed, topoisomerase I was also shown to be essential for replication initiation. These two

topoisomerases are never acting in the pre-RC binding area at the same time, so that they seem to have specialized functions in the context of topology modulation for regulation of origin activity. Topoisomerase II seems to be especially involved in pre-RC assembly and Topoisomerase I in origin firing. The dynamic interplay between ORC topology-modifying enzymes and DNA replication origin throughout the cell cycle, demonstrate the importance of DNA topology for the origin regulation.

1.2.8. Cruciforms structures

Numerous studies have shown that also the cruciforms structures in DNA, that naturally occur as secondary structures, serves as recognition signals at or near origins in yeast and mammas[81]. Stem-loops and cruciforms structures can be formed from DNA inverted repeats. These structures are widely distributed in eukaryotes and may affect the supercoiling degree of DNA, nucleosome positioning and the formation of other DNA secondary structures, or directly interact with proteins. A dynamic distribution of DNA cruciforms in mammalian nuclei exists and their number becomes maximal at the G1/S boundary. A lot of evidences support the hypotesis that particular inverted repeats located in potential DNA replication origins give rise to cruciform structures that serves as attachment site for initiator proteins. Probably some cruciform-specific binding proteins are involved in regulation of replication[82]. Thus, the mechanism proposed involved cruciform stabilization and recognition by replication initiator proteins. CBP and 14-3-3 proteins are indicated as proteins involved in DNA replication through binding to cruciform structures. CBP (Cruciform-Binding Protein) was identify for his specificity for cruciform DNA regardless of its sequence and was then revealed as a member of 14-3-3 protein family. This family is highly conserved and widely distributed in eukaryotes, consisting of several isoforms in each species and with many functions ascribed. Among the 14-3-3 associated proteins were also identified some DNA replication initiator proteins as Topoisomerase II and some MCM[83].

1.3. Hox proteins

1.3.1. HOX genes

During the development, animals specify many different types of cells. Each cell must be developed in the correct context and at in the opportune time. For this purpose cells own the genetic information that then are translated into reproducible spatial and temporal signals.

All bilateral animals possess a common genetic mechanism for the regulation of the development along the body axis and HOX proteins are among the key regulators in the specifying regional identities[84, 85]. These proteins are transcription factors that act during normal embryo development.

The Hox genes are expressed in defined domains in the anterior-posterior axis and the importance of their function in assigning positional identities along the embryonic body is evident when an Hox gene function is disrupted, which usually results in "homeotic transformations"[84,86]. The term "homeotic transformations" was introduced by Bateson in 1894 and derived from the word "homeosis". It is used to describe the transformation of a structure or body segment, in form and shape, into another homologous structure present in the body[87].

The genes encoded for the HOX proteins were first identify in the fruit fly *Drosophila melanogaster*. In this organism, mutations in such genes often resulted in transformation of the body plan. Initially they were discovered by the observation of two dramatic alterations in the fruit fly body, in which after mutation of bithorax gene, the haltere balancing organ on the third thoracic segment is transformed into part of a wing whereas mutation in the antennapedia gene resulted in transformation of the antennae into legs.

Further studies in *Drosophila* showed that these genes were found to cluster together in the genome, forming two groups with a total of eight genes. A group, named *Antennapedia* complex, regulates the development of the anterior body of the fly and is comprised of five genes, whereas the other group, named *Bithorax* complex, regulates the development along the anterior-posterior axis and comprise three genes (Figure 3).

An important feature of these genes is the collinearity, indeed the gene order in the cluster mimics the order of expression of genes and their function along the anterior–posterior body axis: genes at the 5′ end of the cluster are expressed in, and

pattern the posterior part of the body, whereas genes at the 3′ end pattern the anterior end of the body. In some species, homeotic genes also exhibit temporal collinearity in addition to the spatial one: anterior genes are expressed first during development and posterior genes later.

The loss of function of any Hox gene allows the expression of other overlapped Hox gene, resulting in the transformation of a segment identity of the body towards the identity of the neighbouring segment. However, mutations in Hox genes do not always determine such strong phenotypes, but they can also cause very subtle defects, as frequently observed in organisms with multiple Hox clusters. This is due to the overlapping expression and functional redundancy of paralogous Hox genes of different clusters. Indeed, individual Hox genes, at least in higher animals, show a degree of functional redundancy, possibly caused by the duplication events of the Hox cluster described below. Thus, morphologic specification by HOX proteins may also be partially determined by the concentration of the single Hox member in a dynamic interaction with other HOX proteins[88]. However, even in organisms with a single cluster, as in *Drosophila*, homeotic transformation are primarily observed after mutations in those Hox genes that either have overlapping expression domains or are engaged in a negative cross-regulation with other Hox genes[85, 89].

Although first described in *Drosophila melanogaster*, homeotic transformations are found in many other organisms, which led to the conclusion that Hox proteins act as principle regulators of morphogenesis.

The characteristic of these proteins is the homeobox, a highly conserved domain in HOX proteins, constituted by a 60 aminoacids motif. The tertiary structure of this domain consist of three alpha helices that form an helix-turn-helix motif and an additional domain know as the aminoterminal arm, located just adjacent to the first helix. DNA contacts are formed primarily by residues 47, 50, 51, and 54 in third alpha-elices, the so called "recognition helix", and by an arginine in position 5 of the amino-terminal arm[90]. This domain enables the homeobox proteins to bind to DNA at specific binding sites and activate the transcription of their target genes. The rest of the protein may be very different, but this homeobox motif is crucial for its function. In some cases, however, the specificity also required non homeodomain residues, in particular, those that are adjacent to amino or carboxyterminal of the homeodomain[91].

1.3.2. Hox cluster in *Drosophila* and Human

The *Drosophila* species contain eight homeotic genes distributed in two groups: the Antennapedia cluster is comprise of the five genes *lab, pb, Dfd, Scr* and *Antp*, whereas the Bithorax cluster comprise the three genes *Ubx, Abd-A* and *Abd-B*. The name of the genes was given based on the phenotype. If the first mutant isolated was dominant the symbol of the gene begins with capital letter as for Dfd, Scr, Antp, Ubx, AbdB and if the first mutant isolated was recessive the small letter as for lab, pb adbA. Actually, several other genes containing the homeodomain have been isolated from the *Drosophila* genome at later stage and are more commonly reffered to as homeobox genes.

The mammalian Hox genes family consists of 39 genes organised in four clusters labelled A, B, C and D, located on four different chromosome in the genome, respectively the chromosome 7, 17, 12 and 2 and numbered from 1 to 13, although no cluster contains a full set. They are structural and functional homologues of the homeotic complex of *Drosophila* and are thus thought to have arisen by two separate duplication events. For their homology with *Drosophila* homeotic genes, group 1 to 8 are considered analogous to the *antennapedia* (*Antp*-like) and group 9 to 13 to the *abdominal-B* (*Abd-B* like) ones. It is believed that the evolutionary amplification of the Hox genes started with a cis-amplification of a primordial Hox gene, producing 13 members, which was followed by a trans-duplication of most of the Hox cluster. The trans-duplication is further believed to have occurred twice, leading to the four Hox cluster present in mammalian genome.

These duplication events have had a direct consequence on the striking homology shown by the 39 Hox genes. The genes occupying the same relative position along the 5' to 3' chromosomal coordinate share a higher degree of sequence similarities than the genes occupying adjacent positions on the same chromosome. Indeed, for example, the comparison of the homeodomain sequence of HOXA1 with HOXB1, that have the same coordinate on different chromosomes, and both HOXA2 and HOXA13, distant but on the same chromosome, shows that HOXA1 presents an ~88% of homology with HOXB1, the ~67% of homology with HOXA2 and only ~48% with HOXA13.

The homeobox genes were first identified as developmental regulators and much attention was given to their functional role in embryogenesis and early development. Then, it has became clear that these regulatory genes are also active in normal adult

cells, where the expression of homeobox genes is somehow related to controlling cellular identity and regulation of genes necessary for cell division, cell adhesion and migration, morphological differentiation and apoptosis in metazoan cells. Several downstream targets of the HOX proteins in fact, play multiple roles in several pathways, mainly acting as transcription factors and regulating their own subset of genes. HOX proteins also participate in the regulation of expression of some binding cofactors that influence HOX genes expression.

As during the development, adult cells also possess a tissue-specific Hox gene expression characteristic of a specific organ tissue. This tissue-specificity expression might be responsible for the organ specific phenotype and for the positioned in the biaxial enviroment. The maintenance of HOX expression in adult stages depends on the action of Polycomb Group proteins, which are believed to restrict HOX expression to lineages in which the genes were initially activated. The Polycomb Group proteins act through epigenetic effects on chromatin structure. Another regulator of the activity of HOX proteins is Geminin. The ability of Geminin to block HOX function is linked to the interactions with HOX proteins themselves as well with Polycomb group proteins, wherewith forms multi-protein complexes that regulate chromatin structure to repress gene expression.

There are also a number of divergent homeobox genes present in both invertebrates and vertebrates. Their common feature is the presence of the homeodomain and the involvement in development and cellular differentiation, but differently to the Hox genes, they are located randomly at various chromosomal loci.

1.3.3. HOX and its cofactors

All the homeodomain-containing proteins seemingly recognise and bind to DNA without a stringent sequence specificity. All the HOX protein family bind to a very similar set of AT-rich DNA-binding sites. This poorly specificity in sequence recognition exhibited most in vitro, is probably due to the high conservation of the homeodomain. As a consequence of an almost invariant three-dimensional structure of the homeodomain, the majority of HOX proteins preferentially recognize a conserved, but fairly unspecific, ATTA motif[92]. Clearly, the presence of this rather common sequence cannot be sufficient for HOX regulation. Moreover, the same motif is readily

bound also by many non-HOX proteins, thus raising the question of how specificity is achieved.

Structural analyses revealed the mechanism used by HOX proteins to select specific binding sequences in vivo showing that HOX proteins recognize generic HOX-binding sites through the helix-turn-helix that bind DNA major groove, with an additional contact in the minor groove through the aminoterminal arm[90]. The selection among sites is critically dependent on minor groove interactions determined by positively charged aminoacid residues located in the aminoterminal arm of HOX proteins[93].

This low DNA-binding specificity however, strongly contrasts with the highly specific effects showed for HOX transcription factors. One well established way of HOX proteins to achieve specificity in vivo is to bind DNA cooperatively with other DNA binding cofactors. These other factors influence and determine the specificity of the HOX proteins. Many data, in fact, suggested that multiple domains within the HOX proteins are essential for the in vivo specific DNA-binding[94, 95]. Thus, the emerging idea is that HOX proteins would heterodimerize with many other factors and from these interactions result their sequence selectivity and specificity in DNA-binding. Thus, the more stringent selectivity in DNA binding observed in vivo, respect the relaxed one showed in vitro, is probably due to the cooperation with other proteins.

The factors involved in the specificity of the HOX proteins are called cofactors. These cofactors may reveal intrinsic latent specificities to the individual HOX proteins[96]. This would also partially explain the increase in binding affinity by HOX proteins bound to DNA in conjunction with cofactors compared to binding affinities of HOX monomer. The cooperative binding of HOX proteins with cofactors induces conformational changes revealing novel, specific binding properties of the complexes to their DNA targets.

The first cofactor identified in *Drosophila* was Extradenticle Exd. Mutations in the exd gene were originally identified as causing homeotic transformations of specific body segments in the fly, without altering the expression patterns of the HOX genes themselves[97]. Thus, Exd function together with HOX proteins to alters the morphological consequences of their activities. Exd encodes a divergent homeodomain protein related to a vertebrate protein Pbx1, that was independently identified by mutations that leads to human preB cell acute lymphoblastic leukemia[98]. Also some HOX proteins are involved in human leukemia. In both cases, genetic rearrangements

lead to the formation of chimeric proteins that retain the ability to interact each other and to bind DNA. The inappropriate expression of the chimeric Pbx or HOX proteins causes aberrant proliferation of hematopoietic progenitors and leukemic transformation[99-101]. This cooperative role of Hox and Pbx in leukemic cells may reflect a cooperative role for these two proteins in promoting cell proliferation during normal development.

Pbx was the first characterised Hox cofactor. In vitro, Pbx/Exd cooperates with a broad subset of Hox proteins to bind a paired recognition element on DNA with high specificity. The cofactors interactions are dependent on a hexapeptide sequence containing four specific core residues (YPWM) located in aminoterminal of the Hox paralogue groups 1-8. Hox paralogue groups 9-10, which lack the consensus sequence (YPWM) also cooperatively bind to DNA with Pbx1 with binding dependent on a conserved trp residue[102]. Exd and Pbx belong to the PBC subclass of TALE (three-amino-acid-loop-extension) homeodomain proteins.

The TALE proteins are divided in two groups, PBC family and Meis family. Meis is another cofactor belong to the Meis family, which acts as a DNA binding patner for the remaining paralogue groups 11-13[103]. TALE family homeodomain proteins also carry out many HOX-independent functions in vivo[104]. Since they have both HOX-dependent and HOX-independent functions, the genetic analysis of TALE family genes needs to consider that only a subset of the observe phenotypes is due to their role as HOX cofactors.

Pbx proteins interact with Meis/Prep family members in a DNA independent manner through highly conserved domains present in the aminoterminal region of these proteins[105]. In contrast, Pbx-HOX interaction appear to be more complicated involving the direct contact between TALE motif in Pbx and YPWM of HOX proteins[106]. For the HOX proteins that do not have the YPWM motif, there is a conserved W residue that plays an important role in this protein-protein interaction[103]. Moreover, some data, suggest that Pbx proteins may have other ways of interaction with different HOX proteins that are alternative to the classical YPWM-TALE interaction[107, 108].

Evidences suggest that Meis/Prep proteins are also required for normal function of Hox proteins. Hox-Pbx binding site is often associated with Meis/Prep sites, and Meis/Prep bind these sites in cooperation with Hox and Pbx in vitro[109].

The in vivo effects of cofactors binding on HOX function are still controversial. Depending on the target, HOX-Exd or HOX-Pbx complexes can act as transcriptional

activators or as transcriptional repressors. This has suggested that transcriptional effect is determined by the recruitment of other factors into the complex depending on the specific regulatory sequence involved or on extracellular signals[110].

Although cofactor binding clearly affects the specificity of Hox for DNA binding, the effect on DNA binding affinity differs for different Hox proteins. For some Hox proteins, interaction with Pbx confers vastly increased DNA binding affinity, whereas for others the interaction appears to have a minimal effect[106].

Given the high degree of specificity required for some HOX functions, and that there are many HOX genes in vertebrates, the HOX cofactors identify up to now are relatively few. This is also because, especially in vertebrates, multiple Pbx and Meis/Prep genes encode for proteins with different biochemical properties and thus expand the number of HOX cofactors.

HOX proteins used these cofactors to both activate and repress target genes, raising the question of how gene activation versus repression is determined. One possibility is that the presence of additional factors bind to HOX-targeted elements and contribute to their activities. These accessory factors are called " collaborators". Because HOX proteins work in many different developmental contexts, it is likely that HOX collaborators may include a very large number of different proteins, such as transcriptional effectors that are downstream of cell-cell signaling patways. Perhaps the ability of HOX proteins and Pbx-HOX dimer to interact with a large number of different collaborators makes these proteins ideal regulators of cell type and tissue identities[111].

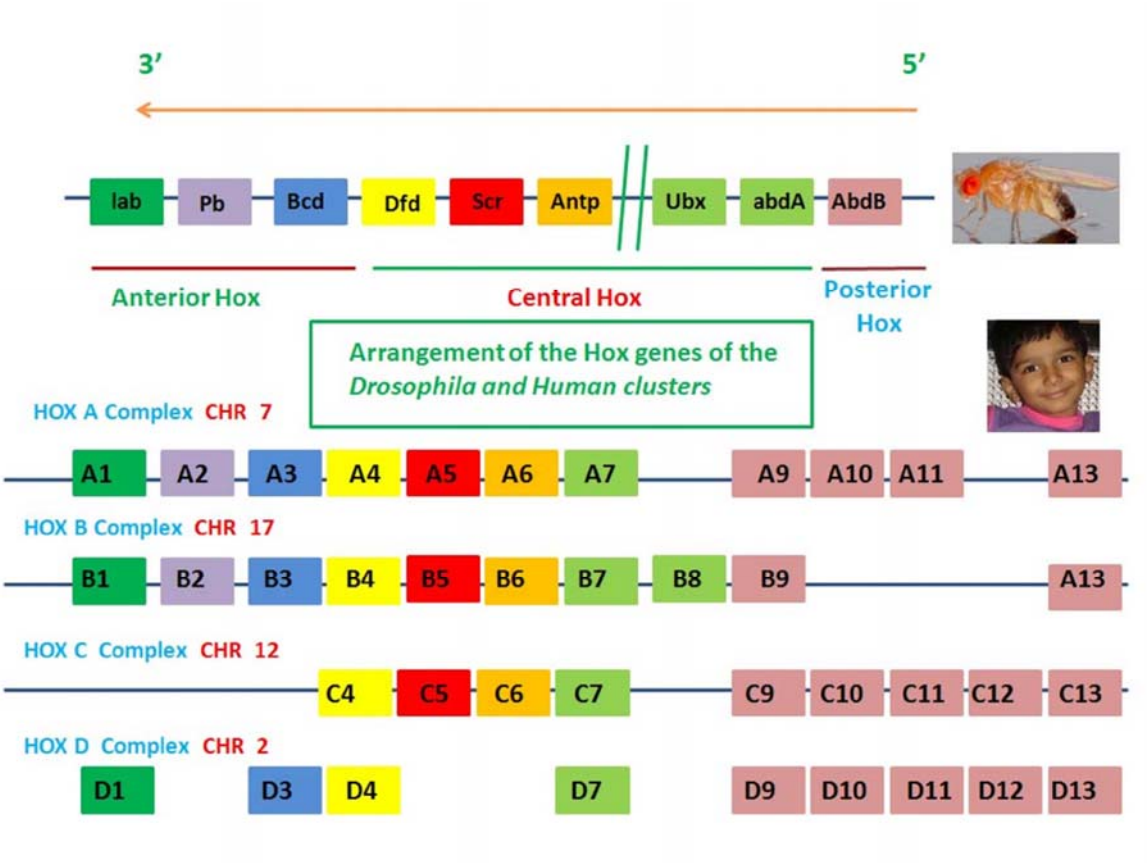

Figure 3. Alignment of Human with *Drosophila* HOX genes. The four clusters of Hox genes in vertebrates have arised by two duplication events of *Drophila* cluster, known as A, B, C, and D. Based on sequence similarity the genes can be sorted into 13 'paralog' groups. The order of paralogs along the chromosome is preserved in the four complexes. The genes within a complex are transcribed in the same direction and are numbered according to their paralog group from 1 at the 3′ end to 13 at the 5′ end. In several cases a member of a paralog group is absent from a complex, in this case the corresponding gene number is omitted. In the picture the same paralog group are colored identically.

1.4.Hox genes and cancer

1.4.1. Hox as oncogenes

The so called "oncogenes" are normal genes that cause cancer when their expression or function became aberrant and contribute to the development of a neoplasia. Generally, these mutations are dominant in that the oncogenic consequences can be manifest in the continue presence of the normal gene product. In some cases, the elevated or untimely expression of the normal gene product is sufficient to convert a normal gene in an oncogene.

Many oncogenes encode important proteins involved in cellular growth and differentiation, including transcriptional factors. Among these are the members of Fos and Jun family, transcriptional factors largely studied for the frequent overexpression of these genes in human tumours and in tumour cell lines. The overexpression of these two factors brings to cell transformation. In addition, Fos and Jun may play an important downstream role in transformation together with other activated oncogenes. Also MYC proteins, a family protein commonly associated with transcription factors, result to be involved in a large number of neoplasias. MYC has been studied in depth because it is involved in a large number of neoplastic conditions as well as for its contribution in cell growth regulation. In many cases, the chromosomal translocations cause a deregulation in MYC gene expression, as in the Burkitt's lymphoma and in AIDS associated lymphoma and T leukaemias[112].

Several lymphoblastic and myeloblastic leukaemias have cytogenetically visible rearrangement on chromosome 11q23. The gene on chromosome 11q23 has been called MLL and present homologies with the *Drosophila* homeotic trithorax gene. As trithorax, also MLL encodes for a protein containing a series of zinc-finger motifs and seems to be a mammalian homeobox gene. In all cases studied so far, its translocation generate fusion proteins that are involved in malignancies derived from hematopoietic lineages[113]. The fusion products often involved transcriptional regulators which can regulate the cell future. The chromosomal rearrangement involving MLL with homeobox protein CDX2 regulator, or the translocation of the Pbx1with a cofactor of HOX proteins, lead to an aberrant expression of HOX genes found in malignant cells, supporting in this way the involvement of the HOX genes in human leukaemia[114-116]. In leukaemia, in fact elevated levels of HOX genes have been frequently observed.

Homeobox transcription factors are known to be involved in both cells proliferation during early development and in cell differentiation. To control these two processes, homeobox proteins act regulating the cell cycle progression and, in many cases, interacting with factors of the cell cycle machinery. Since it is widely accepted that the process of normal development and oncogenic transformation have a great deal in common, it is quite predictable that homeobox genes have been often found deregulated in cancer cells, where they can contribute to the shift between differentiation and cell proliferation.

The first indication of a link between HOX proteins and activation of replication origins, and hence with proliferation and cancer, is represented by the protein Geminin. Geminin is involved in embryonic development by interacting with the Polycomb complex that represses HOX genes, or by direct binding with HOX proteins. HOX, and the others homebox proteins, compete with Cdt1 for the binding to Geminin. The binding of HOX proteins inhibits the expression of genes important for morphogenesis preventing the binding of HOX with DNA, while the binding of Cdt1 repress DNA replication and cell proliferation. When HOX is bound by Geminin, Cdt1 is displaced from its complex with Geminin itself[117]. Thus Geminin, playing multiple roles in several fundamental cellular processes including proliferation, differentiation, development and transcriptional regulation, appear to be critical in the strict coordination that occurs between embryonic patterning and the cell cycle. Geminin represent in this way a connection of HOX proteins function and DNA replication origin regulation, apparently also without a direct contact of HOX proteins with the origins. Indeed Geminin has no affinity for DNA and his connection with the origin is due to the binding of Cdt1. On the contrary, many data show that HOX proteins can bind directly the replication origins and the proteins involved in DNA replication[118].

Important indications of a connection between HOX proteins and cell proliferation, from which derives also the connection with cancer, were obtained using two different systems namely *Xenopus* embryo and retinoic acid-induced differentiation of pluripotent mouse P19 cells. It was shown that the first cell cycles following the mid-blastula transition in *Xenopus* are necessary and sufficient for HOXB activation, whereas later cell cycles are necessary for the correct expression of those genes. Similarly, in P19 cells, HOXB expression requires proliferation, and the entire locus is activated within one cell cycle. The activation of HOXB genes is colinear within a single cell cycle and concomitant with S phase. Thus, induction of HOXB genes occurs

in a DNA replication-dependent manner and requires only one cell cycle. Furthermore, in undifferentiated mouse P19 cells were detected several DNA replication origins in the 100 kb HOXB locus, indicating a relaxed origin use when the locus is transcriptionally silent. After the differentiation of the cells, when HOXB transcription is activated, a general silencing of DNA replication origins occurs in the locus except one located downstream of HOXB1, at the 3' boundary of the HOXB domain. Silencing of the replication origins is associated with histone hyperacetylation, whereas the active HOXB1 origin persists as a hypoacetylated island. These findings provide direct evidence for the differentiated use of origins in HOXB genes that might contribute to the regulated expression of HOXB genes during development[119].

Evidence for the role of HOX proteins in oncogenesis came from the retrovirally activated gene HOX 2.4 in NIH 3T3 fibroblast clones that produce fibrosarcomas in nude mice and from the observation that a translocation involving human chromosome 10, band q24, in a subset of T-cell acute leukemias, disrupts a region surrounding the putative oncogene HOX11[120, 121].

Many studies show that either loss of function or gain of function of homeobox genes may be involved in neoplastic transformation. In the majority of the cases, where the expression of a particular homeobox gene is associated with malignant transformation, this gene is also expressed during embryogenesis in development tissue later affected[122].

Overexpression of HOX genes in mouse bone marrow cells provided additional evidence of their role in specifically regulates different stages of hematopoietic development. Indeed, overexpression of HOXA10 gene showed that it affects myeloid differentiation, leading to myeloid proliferative disease or leukaemia[123].

1.4.2. HOX, NUP98 and leukaemia

HOX genes are involved in the pathogenesis of human leukaemia also through chromosomal translocations of HOX itself. HOX genes are in fact also the most common fusion partners of nucleoporin NUP98 in a growing list of identified NUP98 fusions in patients with acute myeloid leukaemia AML and chronic myeloid leukaemia[124]. AML is characterized by the uncontrolled proliferation of myeloid cells that accumulate at different stages where their further differentiation is blocked.

NUP98 gene encode for the NUP98 protein, a nuclear core complex that transports proteins and RNA/protein complexes between the nucleus and the cytoplasm and it is frequently involved in primary and therapy-related hematologic malignancies[125]. NUP98 has great propensity to form fusion genes, particularly with the member of HOX gene family and the rearrangements between these two genes are the cause of acute myeloid leukaemia. In this aberrant fusion protein, the resulted product have the amino-terminus of NUP98 containing a region of phenylalanine-glycine repeats attached with the carboxy-terminus of the HOX protein, containing the intact homeodomain. Not all Hox genes formed strong leukemogenic NUP98 fusion genes, but under certain circumstances, such as increased of the cofactor Meis1 expression, all induced myeloid leukemia.

Decisively, the Abd-B-like Hox genes formed NUP98 chimeric proteins with the greatest impact both in proliferation and differentiation in vitro and in leukemogenic potency in vivo, which may account for the fact that up to now, only HOX genes belonging to this group have been found rearranged with NUP98 in human leukemia. At the beginning of these type of studies, HOX partners for NUP98 appeared to be all Abd-B like genes, including HOXA9, HOXA11, HOXA13, HOXC11, HOXC13, HOXD13[126-129]. Nevertheless, the data reported about the leukemogenic potential ability of the Antennepedia-like genes, HOXB3 as fusion partner of NUP98 indicate that this ability is not restricted to Abd-B like group, but rather is determined by the intrinsic leukemogenic potential of the HOX protein itself. In fact, this property is not observed using another Antennepedia-like genes like HOXB4. The potent leukemogenic activity of Abd-B-like HOX genes is correlated with their strong ability to block hematopoietic differentiation while coexpression of the HOX cofactor Meis1 alleviated the requirement of a strong intrinsic HOX-transforming potential to induce leukemia. These observations support a model in which many if not all HOX genes may have the capacity to contribute to leukemic transformation, presumably reflecting a fine regulation in common pathways[130].

HOXD13 is one of the HOX genes involved in AML and his involvement in the induction of tumor is not limited to the hematopoietic system, but its deregulation of expression has been detected in breast cancer, melanoma, cervical cancer and astrocytomas[131].

Thus, the fact that the ability to form leukemogenic NUP98-HOX fusion genes is not restricted to the already identified HOX genes but it could be a common feature of

almost all the HOX genes, probably is due to the derivation from a common ancestor. Their homeodomains are closely related, which may indicate that their redundant function in leukemogenesis is mediated through regulation of a similar or overlapping set of genes and pathways that control hematopoietic cell growth and differentiation.

In attempt to reach a molecular classification of leukaemia, using DNA microarray, HOXA9 was identified as a predictive marker for the diagnosis of AML. HOXA9 expression was the only one among the many tested gene that showed a strong correlation with clinical outcome and its expression was also correlated with treatment failure in patients with AML[132]. HOXA9 and Meis are commonly co-expressed in myeloid cells lines and in samples from patiens with AML, of all type except promyelocytic leukaemia[133].

The long latency observed in HOX-induced AML in mouse model strongly indicates that additional genetic events are required for full leukemic progression[134]. Indeed, many data suggest that HOX fusions, as well as most transcriptional factors fused with oncogenes, alter the growth and the differentiation of early hematopoietic precursors leading to the establishment of a pre-leukemic population of cells that are then susceptible to the acquisition of cooperative mutations[135]. This idea is reinforced by the finding that the overexpression of Meis cooperates with multiple native and NUP98-HOX fusion genes to accelerate the onset of AML[130]. Meis is not leukomogenic by itself, but its important role in normal hematopoiesis is supported by its preferential expression in the hematopoietic stem cells compartment[134] and by the multiple hematopoietic defects observed in Meis deficient mice[136].

1.4.3. Oncogenic potential of homeobox genes

While the role of Hox genes in leukaemia is well defined, the oncogenic potential of some specific Hox genes in neoplasias is currently being studied. In many forms of cancers, chromosomal translocations lead to the creation of fusion genes that acts as proto-oncogene. Numerous reports have identified differences in Hox genes expression between normal and neoplastic tissues, but there are relatively few studies illustrating their function in cancer and the Hox genes functional relationship with the malignant phenotype is still elusive. Useful information has been achieved by monitoring biological responses in models in which the expression of a specific Hox gene has been

altered. These experimental approaches have widely recognized limitations and are complicated by the functional redundancy present in Hox gene system. For the same reason, in vertebrates, mutations in single Hox gene do not cause dramatic alterations in morphogenesis, differently from what observed in *Drosophila*. Indeed, gene duplication results in multiple members of each paralog group and therefore a significant degree of redundancy of the HOX organization. Moreover, vertebral morphogenesis depends on quantitative effects of HOX proteins on many effectors genes, while in *Drosphila* the proteins encoded by Hox genes regulates only a few downstream effectors which are sufficient to trigger alternative developmental pathways within each segment.

Abnormalities of homeobox gene expression were identified in many solid tumors and in cell lines derived from them. Three different categories of homeobox gene expression have been identified in malignancy. In the first, the homeobox genes expressed only during the development are re-expressed in neoplastic cells. This result in a gain of expression that is seen in many primary tumours and cell lines such as brain, mammary gland and kidney in which the homeobox are expressed during the embryogenesis. In the second category, homeobox genes are expressed in malignant cells derived from tissue in which a particular gene is not expressed during embryogenesis. In this case a new activity occurs in the cells, as in the case of PAX5 expression in medulloblastoma but not in cerebellum, that is the tissue from which PAX5 derived. The third category is the one where the homeobox genes are down-regulated in malignant cells derived from tissue in which a particular gene in normally expressed in the differentiated state. This is the case of the loss of expression of CDX2 and NKX3.1 in colon cancer[137].

Malignancy is associated with the failure of the cells to differentiate and several homeobox genes have been found to be involved with terminal differentiation. Over-expression of some HOX genes is associated with loss of differentiation, as in human prostate cancer where the over-expression of HOXC8 is associated with loss of tumour differentiation in human prostate cancer suggests that HOXC8 may play a role in the acquisition of the invasive and metastatic phenotype of this malignancy[138].

Some data also indicate a relationship between metastatic potential of melanoma cells and the expression of integrins, ICAM-1 and HOXC genes. In fact, in clones with high levels of integrins and ICAM-1 there is no detectable expression of HOXC locus. On the contrary, in clones with low levels of integrins and ICAM-1, the HOXC genes are actively express indicating an inversely related expression of HOX genes respect the

genes involved in the interaction between cells and between cells and matrix. Cells with low levels of proteins involved in these interactions are expected to have an higher metastatic potential. HOX gene expression reflects the intra-tumor heterogeneity of cancer cells and suggest that the expression of surface molecules involved in cell-cell and cell-matrix interactions may be related to the patterns of HOX gene expression[139].

1.4.4. Role of HOX-Pbx interaction in cancer

Despite the growing body of data implicating HOX genes in the development of various cancers, little is known about the role of HOX-Pbx interactions in the regulation of proliferation and induction of transformation of mammalian cells.

The first indication about a relationship between HOX proteins and cell proliferation come from a study in which was analysed the Pbx homeobox protein, also described in translocations in human leukemias, and its cooperation with HOX proteins as cofactor to the DNA binding. It was observed that, in rat cells, cellular transformation and proliferation induced by HOXB3 are greatly modulated by the levels of available Pbx1 present in these cells and that the transforming capacity of this HOX protein depends on his conserved tetrapeptide and homeodomain regions which mediate binding to Pbx and DNA, respectively. Taken together, results of this study demonstrate that cooperation between HOX and Pbx proteins modulates cellular proliferation and strongly suggest that cooperative DNA binding by these two groups of proteins represent the basis for HOX-induced cellular transformation[140].

Furthermore, normal Pbx1 homeodomain protein, as well as its oncogenic derivative, E2A-Pbx1, binds the DNA sequence ATCAATCAA cooperatively with the murine HOXA5, HOXB7, HOXB8, and HOXC8 homeodomain proteins, which are themselves known oncoproteins, as well as with the HOXD4 homeodomain protein. Cooperative binding to ATCAATCAA required the homeodomain-dependent DNA-binding activities of both Pbx1 and the HOX partner. It was found that HOXB8 suppressed transactivation by E2A-Pbx1 suggesting that oncogenic mechanisms of certain HOX proteins may require their physical interaction with Pbx1 as a cooperating DNA-binding partner[141]. The importance of Pbx-HOX interaction, which inhibition could be a strategy to control the abnormal proliferation of the cancer cells, for determining the cancerous phenotype is indicated by the effects of the disruption of this

interaction using specific small molecules containing a second homeodomain, that modifies HOX activity and act as antagonist of the interaction between HOX proteins and Pbx[142]. Another small designed peptide amphiphile (PA) which self-assembles into micelles, shows inhibition of pancreatic cancer cells, leukemia cells and melanoma cells, while non-cancerous fibroblast NIH 3T3 cells are less affected. This molecule contains critical regions designed to disrupt HOX-Pbx-DNA complex formation and significantly enhances the effectiveness of the molecule to slow cell proliferation[143].

In the case of genes corresponding to the Abd-B like, the same cooperation is observed with Meis cofactor. The Meis homeobox gene was identified as a common site of viral integration in myeloid leukemias arising in BXH-2 mice. These integrations result in constitutive activation of Meis1. The Abd-B like HOXA9 is frequently activated by viral integration in the same BXH-2 leukemias. HOXA9 protein physically interacts with Meis1 proteins by forming heterodimeric complexes on a DNA target. Also Hox proteins from the other AbdB-like paralogs, HOXA10, HOXA11, HOXD12 and HOXB13, form DNA binding complexes with Meis, while HOX proteins from other paralogs do not appear to interact with Meis proteins[103].

Given all these observations, the direct involvement of these transcription factors in the start of DNA replication, both in regulation of expression of target genes coding for replication protein and in oncogenesis have began to emerge.

1.5. HOX and DNA replication

Regulation of eukaryotic DNA replication involves a large number of factors. Most of these have been identified, as seen before, and appear to be conserved among many organisms[144]. The studies regarding these main factors however, did not lead to a complete elucidation about the mechanisms regulating the DNA replication. Thus, this suggests the presence of other elements operating in origin activation and regulation. In order to identify new proteins participating in DNA replication, a yeast one-hybrid screen for human proteins was used[145]. The probe used for this screening was a portion of the human origin of DNA replication lamin B2.

Lamin B2 origin is located near the 3' end of the lamin B2 gene, close to the promoter of the housekeeping gene TIMM13 on the short arm of chromosome 19 and it is currently the best characterized human origin[146]. It was shown that 1.2 kb of the lamin

B2 replicon promotes initiation of DNA replication when integrated at ectopic positions of the human genome and that the activity of the origin depends on a 290 bp[48]. It is an early firing origin which has been shown to be active in a variety of normal and cancer cell-lines and to be covered by a proliferation-dependent footprint which varies dynamically with the cell-cycle[147]. Indeed, *in vivo* footprinting studies of the lamin B2 origin identified a protected area within the 290 bp sequences, due to the binding of proteins involved in DNA replication, that changes during the cell cycle. As expected for a complex involved in origin activation, the protected area is absent in quiescent phase G0, because the incompetence for proliferation in the resting cells. This covered area increases in G1 phase becoming particularly evident in G1/S border, where the protection is extended for 110 bp, and shrinks in S phase, after origin activation, measuring approximately 74 bp. Then, the protection remains unchanged until the M phase when disappears, and reappears in the ensuing G1 phase[46]. Within the protected area, were identify in both the G1 and S phases, the transition points between continuous and discontinuous DNA synthesis for upper and lower strands, that represent the start sites. The start site of the leading strand in lamin B2 origin, correspond to a single nucleotide overlapping by four nucleotides on the complementary lagging strand[45]. Furthermore, *in-vivo* studies have shown that the human ORC1, ORC2, Cdc6 and MCM3 proteins, are in direct contact with this origin sequence at different moments of the cell cycle[47]. Nucleotide-level investigation, shows the position bound by the ORC1, ORC2 and Cdc6 proteins at the origin, throughout the cell cycle. In M phase, accordingly with the lack of covered area at the origin, none of these proteins were found on the DNA[46]. In the middle of G1 phase, ORC1, ORC2 and Cdc6 are located at the lamin B2 origin as part of the pre-replicative complex. In the middle of S, after the origin firing, only the ORC2 protein is still bound on the DNA, moreover its position has been shifted closer to the start-site[47].

In the one-hybrid screen, a protected sequence of 74 bp located within the origin area in S phase, was used as a probe to search for proteins able to specifically interact with the lamin B2 origin. The one-hybrid system is a highly stringent method of selection. To be selected a protein must be able to interact with the target sequence despite the large excess of unspecific binding sites in the used system. Moreover, binding has to occur in the presence of histones and other DNA-binding proteins that could hamper the recognition of the sequence by the protein. Thus, the selected proteins have doubtless a high affinity binding for the 74 bp sequence used as probe.

The identified proteins with this approach were three homeotic proteins, all orthologues of the product of the *Abdominal B* gene of *Drosophila*, namely HOXC13, HOXC10 and HOXA13. To validate the affinity between these homeotic proteins with the probe from lamin B2 origins, recombinant HOXC13 and HOXC10 were challenged in electrophoretic mobility shift assay EMSA, and both the proteins resulted able to bind specifically the 74 bp sequence. In addition, the ability of HOXC10 and HOXC13 to increase the activity of a promoter containing the 74 bp sequence, as assayed by CAT-assay experiments, demonstrates a direct interaction of these homeoproteins with the origin sequence in mammalian cell milieu[145].

The level of HOXC10 were analysed in different phases of the cell cycle and an oscillation was observed. It was shown that HOXC10 levels are reduced in early G1 phase, abundant from mid-G1 to G2 phases and became undetectable in mitosis. This protein in fact, is targeted for degradation in early in mitosis, when it is ubiquitilated, by the proteasome pathway. Mitotic proteolysis of this protein appears to be regulated by the APC. It was also show that failure of HOXC10 destruction causes mitotic accumulation by delaying the metaphase to anaphase transition suggesting novel function of a HOX family member in cell cycle progression.

About HOXC13 another confirmation of this interaction came from an *in vitro* analysis using nuclear extract in which, by immunological detection, was observed a specific binding between HOXC13 and the origin DNA on agarose beads[148].

The interaction of HOX proteins with the origins has been already described also for HOXA13, HOXC10, HOXD11 and HOXD13[145, 149]and in all these cases, depletion of these proteins is compatible with cell cycle progression, probably because the redundant expression of homeotic proteins.

HOXC13 is expressed in vibrissae, in the filiform papillae of the tongue, and in hair follicles throughout the body. Mice carrying mutant alleles of HOXC13 have been generated by gene targeting and it was shown that they have defects in hair, vibrissae, nail, and filiform papilla development as well as in caudal vertebrae, but the mouse are still viable.

Similar to both HOXA13 and HOXD13 mutants, HOXC13 mutants also show an alteration in the limb. Whereas both HOXA13 and HOXD13 mutants have alterations in the phalanges as well as other limb bones, however, the limb defects of HOXC13 mutants are restricted to the nails, which are derived from ectoderm while the expression in the limb appears to be ectodermal. The HOXC13 gene is expressed in

each of these areas, even those well outside the area expected for a paralogous group 13 gene, suggesting that the defects seen in each of these regions are defects directly attributable to a developmental program controlled by HOXC13[105] .

The function suggested for these HOX proteins in DNA replication is concerns the promotion of the pre-RC assembly and the stimulation, in a DNA-binding dependent manner, of replication. Indeed, a direct role in promoting DNA replication origin licensing and replication initiation was observed for HOXD13 and the sharing of the same role for other HOX proteins was hypothesized. In the case of HOXD13, this activity seems to be regulated by Geminin, that blocks the stimulation of origin licensing mediated by HOX protein. Furthermore, it was observed that the binding of Geminin interferes with the binding of HOXD13 and a pre-RC protein Cdc6, controlling also in this way the activity at replication origins[149]. A model proposed for the involvement of HOX in DNA replication is that increasing levels of Geminin would be sequester both HOX proteins and Cdt1, thereby blocking reinitiation of DNA replication, induced by HOX proteins and Cdt1, until Geminin becomes inactivated during mitosis.

HOX transcription factors appear in fact to possess adequate properties for the regulation of replication initiation at multiple origins given that they present a low intrinsic DNA-binding specificity and a preference for binding AT rich sequences.

2.Results

2.1. Characterization of polyclonal anti-HOXC13 antibody

The human recombinant HOXC13 protein fused to GST was obtained by cloning of aminoterminal portion of HOXC13 encompassing amino acids 1 to 259 in pGEX 2T (Amersham Biosciences) vector. This fusion protein exclude the highly conserved homeodomain of HOXC13 (amino acids 260 to 319) in order to obtained more specific antibodies against this HOX protein. The recombinant GST tagged HOXC13 protein was produced in *E. coli* BL21 strain and purified by reduced glutathione agarose (Sigma) beads. The purified protein was then used for polyclonal rabbit antibodies production.

Two rabbits (A and B) were injected with the purified GST-HOXC13 protein following the standard ICGEB immunization protocol. A series of 5 injections and bleedings were obtained from each rabbit and kept separately. Sera were then purified by IgG purification kit (Viva Science) and tested for specific recognition of recombinant HOXC13 used as antigen, and for recognition of endogenous protein in nuclear extracts from human cells. Nuclear extracts from HeLa cells were run in linear gradient gels. These are polyacrylamide gels having a gradient of increasing acrylamide concentrations and linearly decreasing pore size. These gels have the advantages over SDS-PAGE linear gels, with fixed concentration because a much greater range of protein molecular weight values can be separated on a gradient gel than on a fixed-concentration gel. Moreover, using this kind of gels, there is a greater likelihood of resolving proteins with very similar molecular weight values than on fixed-concentration gels.

We set the conditions for the Western blot analyses and we choose the sera derived from the immunization of the rabbit B, because it displays a greater efficiency in recognize the protein and the observed band appear at the expected molecular weight for HOXC13, about 35 KDa, as shown in figure 1.1.

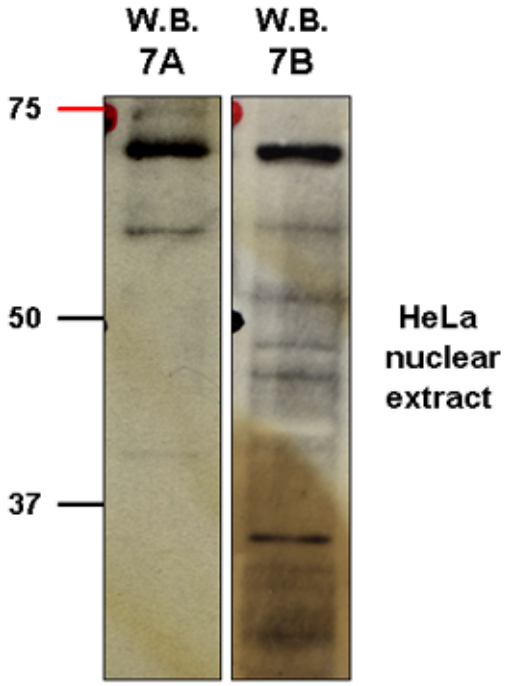

Figure 1.1. Identification of HOXC13 protein in human cells by immunoblotting. Western blot assays were performed with nuclear extracts from asynchronous HeLa cells. Extracts were run in a gradient gel and after blotting, the detection of HOXC13 was performed with sera from different rabbits. 7A and 7B indicate the sera from the last bleeding (7) of the rabbit A and rabbit B respectively. Only the sera from rabbit B recognized a band at the expected molecular weight for HOXC13, about 35 KDa.

2.1.1. Western blot detection of endogenous and double tagged HOXC13 in cells extracts

We also produced HeLa and T98G stable clones expressing a double-tagged version of full length HOXC13. The cDNA of HOXC13 was cloned into pFHIRES Neo vector and transfected in HeLa and T98G cells lines. This vector contains the Neo resistance gene for positive selection with G418 drug of transfected cells and expresses a N-terminal Flag-HA double-tagged fusion protein. The transfected cells were selected on G418 containing medium for more that two weeks to obtained a stably transfected

pool of cells. The expression of the double-tagged Flag-HA HOXC13 in stable cell lines was detected using the anti-HA and/or anti-Flag antibodies. Those antibodies recognise a band at about 45 KDa corresponding to the recombinant protein. We used total cell extracts to further confirm the specificity of polyclonal antibodies against HOXC13. As shown in figure 1.2, in extracts from clones expressing the double-tagged recombinant protein, HOXC13 was similarly detected by anti-HA and anti-Flag antibodies as a clear band migrating of 42 KDa above the endogenous one, and corresponding to the additional band of same molecular weight detected by the specific antibody against HOXC13 itself.

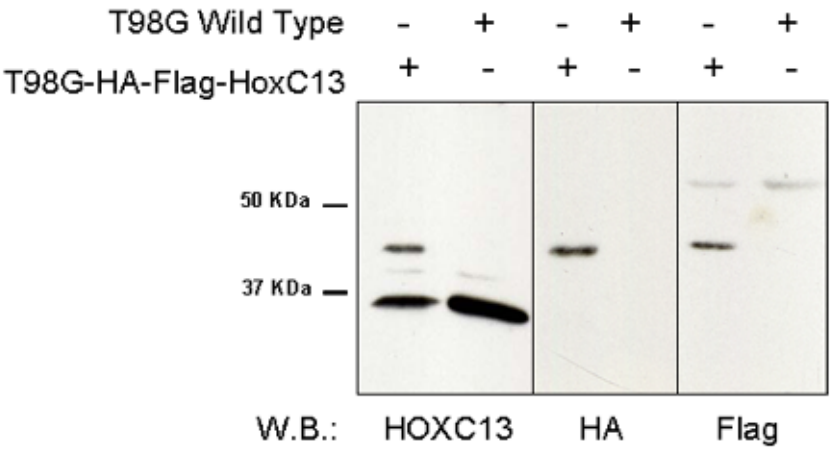

Figure 1.2. Characterization of purified polyclonal antibody against HOXC13. T98G cell line and T98G clone expressing HA-Flag-HOXC13 were used to check the ability of our antibody to recognize endogenous and recombinant HOXC13.

2.1.2. Partial proteolytic peptide maps

To further confirm the specificity of our antibody against the HOXC13 protein, we utilized a peptide mapping approach. In this assay, the treatment of two identical protein samples with a specific protease but at different concentration, will generate a series of products ranging in size from fully digested to undigested protein. When the

products of these individual reactions are run side by side on lanes of one-dimensional SDS-polyacrilamide gel, a ladder of digestion products is displayed that gives a diagnostic fingerprint of the polypeptide backbone. When two proteins are identical, these partial proteolytic maps can be used to confirm their identity. Even a minor changes in the polypeptide is enough to produce a vastly different pattern.

We compared the fingerprint obtained after trypsin digestion on purified HOXC13 protein and total cell extract. To prepare the samples for enzyme digestion, purified HOXC13 and total extract were run in a SDS-polyacrilamide gel that was then stained with Coomassie and dried on paper to excise the different bands. The HOXC13 band was the substrate for digestion with a fixed concentration of trypsin. After digestion were observed a clear fingerprint proteolytic profile for HOXC13. So, we treated the dried gel slices corresponding to molecular weight corresponding to HOXC13 in T98G total cell extract and, as negative control, we used a gel slices of the same lane of total cell extract that do not corresponds to HOXC13 molecular weight. The digestion occurs directly inside the wells of a modified polyacrilamide gel and then the gel is run to separate the different proteolytic products.

By a western blot analysis using the specific antibody against HOXC13, we compared the fingerprint obtained from the recombinant purified HOXC13 and total cellular extract digestions. As shown in figure 1.3, the peptide fingerprint detected by HOXC13 antibody in total cell extract is very similar to the digested recombinant HOXC13 protein, and this antibody did not recognize any proteins in the lane of the negative control, as expected. This result indicates in this way that the HOXC13 band protein recognized in cell extracts by our antibody specifically correspond to the endogenous HOXC13 protein.

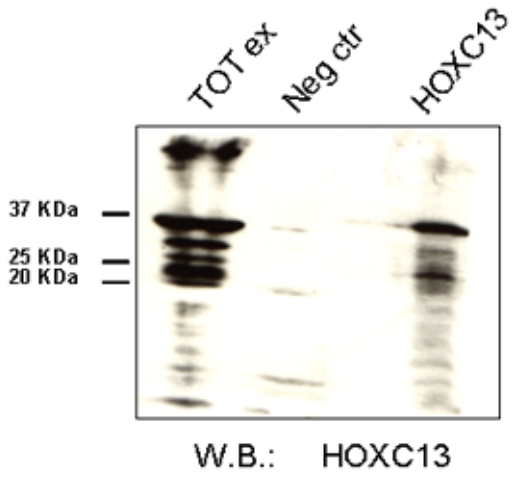

Figure 1.3. Western blot analyses of the proteolytic products produced after enzymatic digestion of T98G cells (TOT ex) and negative control (Neg ctr) cell extracts, and recombinant HOXC13 (HOXC13) protein by HOXC13 antibodies. The comparison of the detected bands in both samples demonstrates that HOXC13 antibody specifically recognizes the full-length protein and its proteolytic fragments in both cell extracts and purified recombinant protein digestions.

2.1.3. Immunoprecipitation of tagged HOXC13

To test the ability of our antibody to specifically immunoprecipitate HOXC13, we performed an immunoprecipitation (IP) using total cell extracts from HeLa stable clone expressing double-tagged Flag-HA HOXC13. The expression of the double tagged Flag-HA protein was monitored using the anti-HA and anti-Flag antibodies. In this way, we could detect the immunoprecipitated HOXC13 also by other antibodies specifically recognizing the tags of the fusion protein. The immunoprecipitation assays were performed using the antibody against HOXC13, while immunoblots were revealed using an antibody against HA tag. As shown in figure 1.4, a clear signal at the expected molecular weight for the recombinant protein, about 45 KDa -10 KDa more that the size of the protein due to the tags- was found in the HOXC13 IP lane, demonstrating that our antibody specifically immonuprecipitates HOXC13. As controls, we used the specific antibody in only IP buffer (without cell extract) to exclude those bands corresponding to the light and heavy chains of the antibody after IP; the IgG in the cellular extract as negative IP control and a protein tagged with HA (HA-GCN5) as positive immunoblot signal for the antibody against HA.

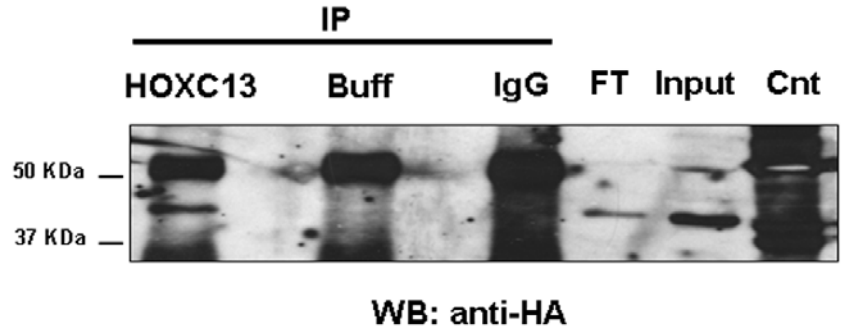

Figure 1.4. Immunoprecipitation of tagged HOXC13. Total cell extracts from HeLa stable clone expressing double-tagged Flag-HA HOXC13 were used for immunoprecipitation experiments. Anti-HA antibody detects a specific tagged HOXC13 after immunoprecipitation with the specific antibody against HOXC13 (HOXC13) corresponding to the same band detected in the input lane. This band was absent in the IP using the antibody alone (Buff) and in the IP with IgG used as negative control. The tagged protein was also detected in the flow through (FT) recovered after the immunoprecipitation. A strong signal for purified HA-GCN5 was observed in the lane used for control of efficiency of anti-HA antibody (Cnt).

2.2. Expression of HOXC13 in different human cell lines

Homeotic genes are principally involved in the specification and regulation of the patterns of development of the body. To carry out this function, their expression is modulated in a tissue-specific manner. We analysed the expression of HOXC13 in cell lines by Western blot analysis of total extracts from different cell types. As shown in figure 2, the cell lines analysed were: the U2OS cell line, with epithelial morphology, derived from the bone tissue of a fifteen-year-old human female suffering from osteosarcoma; SAOS cells, a non-transformed cell line from sarcoma osteogenic, derived from the primary osteosarcoma; the T98G glioblastoma cells, derived from a glioblastoma multiforme; 293T derived from the transformation of normal human embryonic kidney cells with adenovirus; the spontaneously transformed human epithelial HaCat cell line derived from adult skin, which maintains full epidermal differentiation capacity and are immortal but non-tumorigenic; the HeLa cell line, derived from cervical cancer cells and the U937 monocytic cells, isolated from the histiocytic lymphoma.

Immunoblot analysis in these different cell types showed that in some cases, like for 293T and U937, the expression of HOXC13 protein were not detectable, at least by immunoblot, as showed in figure 2. In these cell extracts no appreciable bands were detectable at the expected molecular weight, indicating absence or low expression levels of this protein. The bigger expression of HOXC13 was observed in HaCat cells and mostly in T98G cells, where the levels of expression of this protein results to be very high.

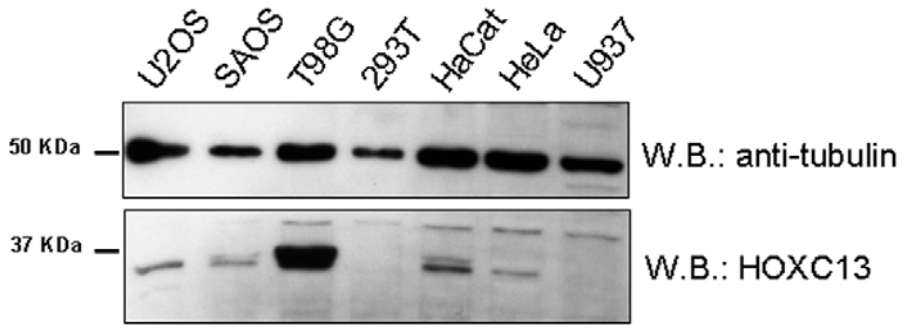

Figure 2. Expression levels of HOXC13 protein in different human cell lines. Western blot analyses of the total extract from the indicated cell lines were performed using antibody against HOXC13 to detect the levels of protein expression. To control the amount of total extract loaded in the gel, the levels of alpha tubulin were detected by antibody against tubulin (α tubulin).

2.3. Protein expression of HOXC13 throughout the cell cycle

The expression levels of HOXC13 protein were analysed during the cell cycle by Western blot using our antibody against HOXC13, in total extracts of synchronized HeLa and T98G cells.

First we synchronized HeLa cells by sequential treatment with thymidine and nocodazole, where the last a drug inhibits microtubule polymerization and arrests cells in G2/M phases. These cells were then harvest by shake-off, released from the block by

several washes, replated into new cell plates and harvest at different times. At 5, 10 and 15 h from the release, harvested cells correspond to early G1, G1/S transition and S phase cell population, respectively.

A second cell line T98G was chosen to analysed the expression levels of HOXC13 during the cell cycle. T98G cells are characterized by their high HOXC13 expression and their capability to be synchronized in G0 phase (quiescent state) by serum deprivation. T98G cells were synchronized by serum starvation method, first growing cells in completed medium for 24h and then for 72h in medium without growth factors (serum). This longer incubation blocks all cells in quiescent phase. The cells are then released from the G0 state by serum addition to the medium and then cells re-enter in the cell cycle all together. Upon re-addition of serum, cells synchronously progressed throughout G1 and S phases. Thus, the synchronized cells can be collected at different times, coinciding with different phases of the cell cycle of our interest.

The advantage of serum starvation method for cell synchronization is the possibility to avoid the use of drugs, but one limit of this method is that over the 28h after the release cells loose synchronicity after they rised G2 phase and only a portion of the entire population result to be in the expected next phase. So we used nocodazole drug to block the cells in G2/M phases, corresponding to the collection 32h after the initial release.

Cell synchronizations showed that contrary to what was observed for another HOX protein involved in DNA replication, HOXC10, which levels are reduced in early G1 phase, increased during the cell cycle until the G2 and became undetectable in M phase, the levels of HOXC13 are constant throughout the cell cycle in both cell lines. As shown in figure 3.1 and 3.2, the HOXC13 protein abundance detected in HeLa and in T98G cells, compared with tubulin abundance used as internal loading control, clearly showed that HOXC13 does not changes through the cell cycle.

The goodness of cell synchronization profiles of propidium iodide stained cells was observed by the FACS analysis, using a similar fraction of the cells collected and used for immunoblotting analysis (figures 3.1 and 3.2). We also confirmed cell synchronization profiling by immunoblotting using specific antibodies against cyclin A, which expression varies in different phases of the cell cycle as shown in figure 3.3.

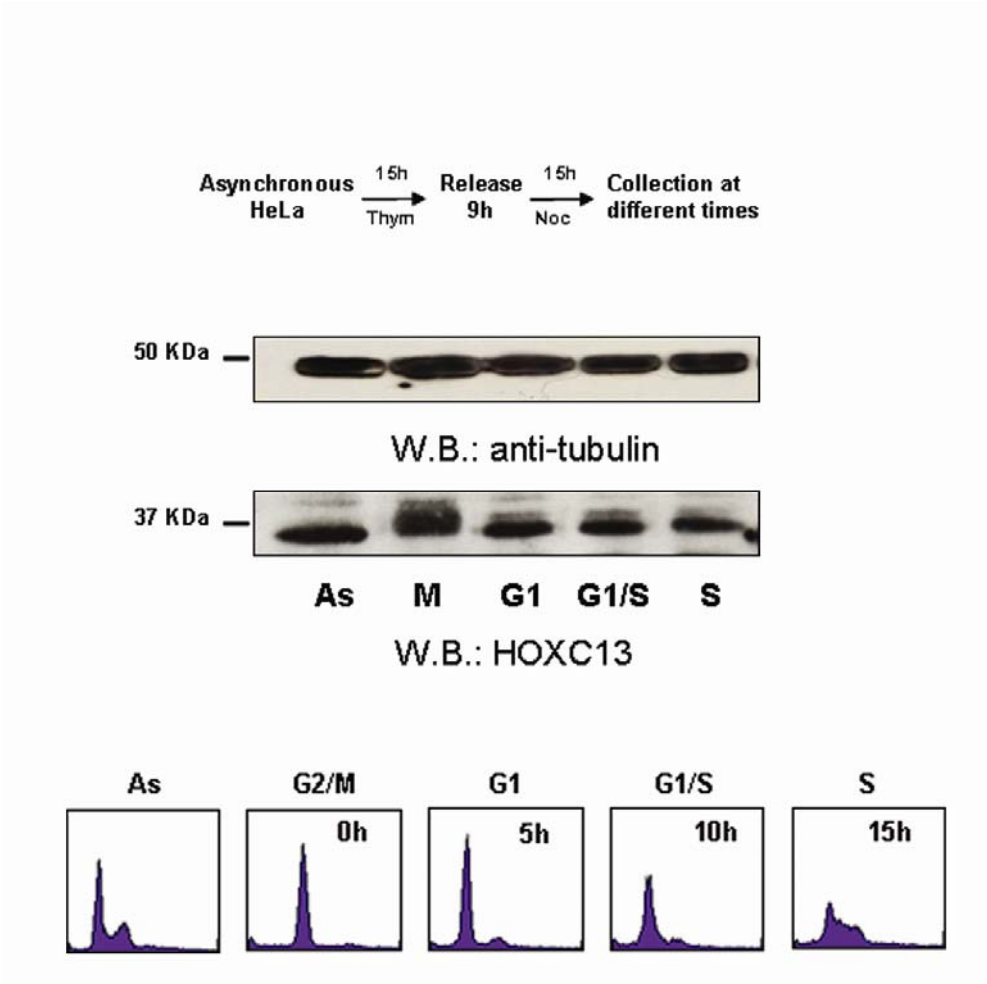

Figure 3.1. HOXC13 expression levels through the cell cycle in HeLa cells. The upper panel shows the experimental scheme used for cells synchronization. HeLa cells were synchronized in G2/M by a double thymidine/nocodazole block, and then release from the block to obtain cells in G1, G1/S and S phase. The total lysates corresponding to each cell population were analysed by immunoblotting with the tubulin and HOXC13 antibodies. In the lower panel, flow cytometry profiles of asynchronous cells (As), cells blocked in G2/M (0 hr) or cells at different times after release from G2/M block, are shown.

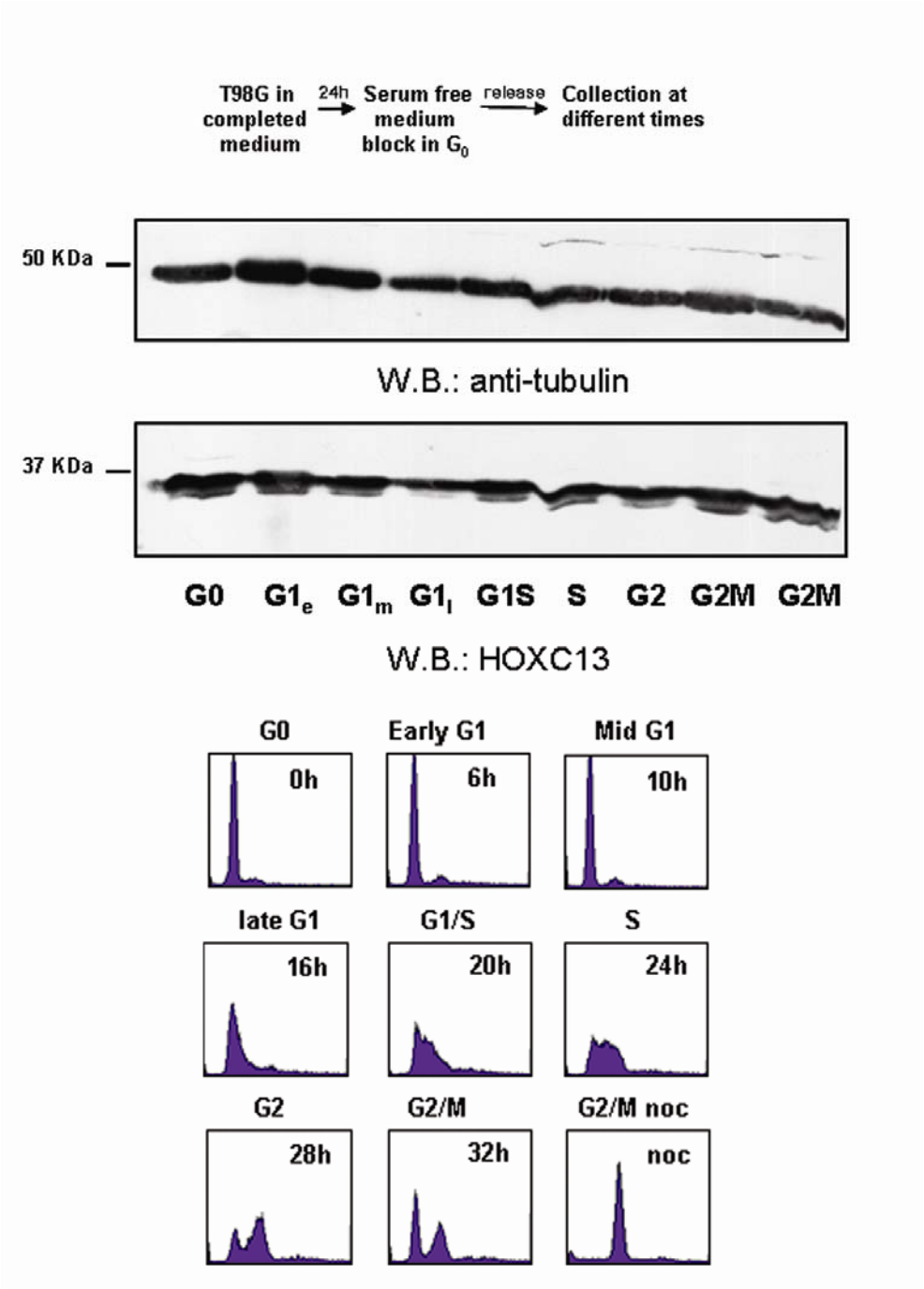

Figure 3.2. HOXC13 expression levels through the cell cycle in T98G cells. The upper panel shows the experimental scheme for T98G cells synchronization. T98G cells were synchronized by serum starvation in G0, and release at different time point after serum starvation, and by nocodazole treatment to obtain the G2/M phase. The total lysates corresponding to each cell population were analysed by immunoblotting with the indicated antibodies. In the lower panel are shown the flow cytometry profiles of propidium iodide stained cells blocked in G0 by serum starvation (0 hr), cells at different times after serum stimulation and cells blocked in G2/M phase by nocodazole treatment.

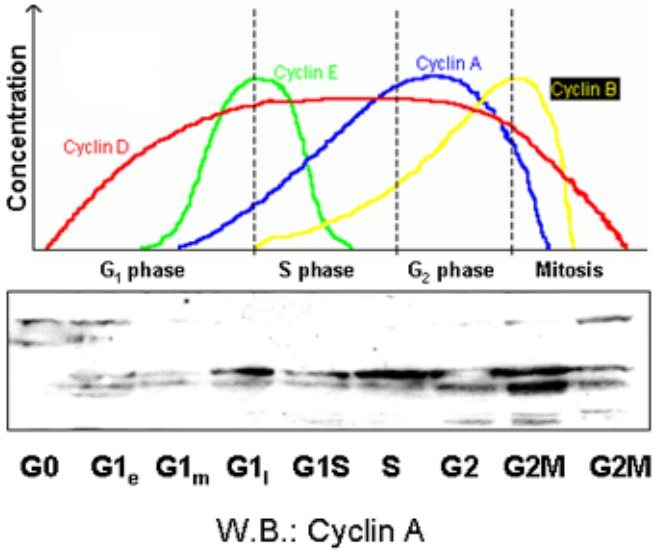

Figure 3.3. Cyclin A expression levels in synchronized T98G cells. The upper panel shows a draw indicating the classical profile of expression of cyclins throughout the cell cycle. The lower panel shows the immunoblotting of the same cell extracts used for analysis of HOXC13 expression in T98G cells. Levels of cyclin A expression were analysed in different stages of the cell cycle by western blot.

2.4. Interactions of HOXC13 with other replication protein complexes

To understand the direct involvement of HOXC13 with DNA replication, we first investigate the affinity of this protein for different components of the pre-RC by GST pulldown experiments. *In vitro* translated [^{35}S]-labeled Cdc6, ORC1, MCM2 and MCM3 were challenged for their ability to bind a GST-HOXC13 fusion protein. Figure 4.1 shows that Cdc6 and ORC1 were specifically retained on GST-HOXC13 agarose beads but not on GST beads used as control, indicating a direct interaction between HOXC13 with Cdc6 and ORC1, two proteins that specifically bind to DNA replication origins; differently, MCM2 and MCM3, two members of the MCM helicases complex, were not retained upon *in vitro* binding, suggesting an absence of direct interaction of this factors with HOXC13. To control the amount and integrity of the GST fusion proteins on agarose beads used for each assay upon *in vitro* binding, gels were stained with Coomassie reagent before autoradiography.

A more detailed analysis of the interacting moieties of HOXC13 and Cdc6 was carried out, as reported in figure 4.2. To determine which portion of Cdc6 interacts with HOXC13, two Cdc6 fragments, showed in the figure 4.3, one from the aminoterminal to aminoacid 363 and the other from aminoacid 364 to the carboxyterminal, were immobilized on the beads and challenged with full-length [^{35}S]-HOXC13. Neither of Cdc6 fragments tested was able to interact with HOXC13, indicating that the central portion integrity of Cdc6 is likely to be the interacting moiety. Conversely, using the *in vitro* translated [^{35}S]-labeled Cdc6 for *in vitro* binding with N-terminal and C-terminal HOXC13 fragments as GST fusion proteins, we only observed a clear interaction only with the C-terminal portion of HOXC13, that encompasses the homeodomain, indicating this moiety as important for Cdc6 interaction.

The same type of *in vitro* assays were used to test HOXC10 as interacting proteins with pre-RC protein components. This HOXC10 homeotic protein was identified together with HOXC13 for its ability to bind the lamin B2 origin in the one hybrid screening assay. Surprisingly HOXC10 was unable to bind Cdc6 in GST pulldown assay, thus suggesting that interaction observed between HOXC13 and Cdc6 was specific (figure 4.4).

This specific interaction was also confirmed by co-immunoprecipitation experiment performed in asynchronous T98G. The cells were first crosslinked by formaldehyde and the total cell extracts were then used for immunoprecipitation with specific antibodies against Cdc6. The immunoprecipitated material was separately divided, one part was used for immunoblotting with antibodies against Cdc6 and the other for HOXC13 immunoblotting to check the presence of both the proteins in the samples.

Hence, as shown in figure 4.5, HOXC13 was co-immunoprecipitated after Cdc6 immunoprecipitation

We can conclude that HOXC13 displayed a significant *in vitro* affinity for two members of the pre-RC, ORC1 and Cdc6. Additionally we showed that this interaction is mediated by the HOXC13 homeodomain, at least for Cdc6.

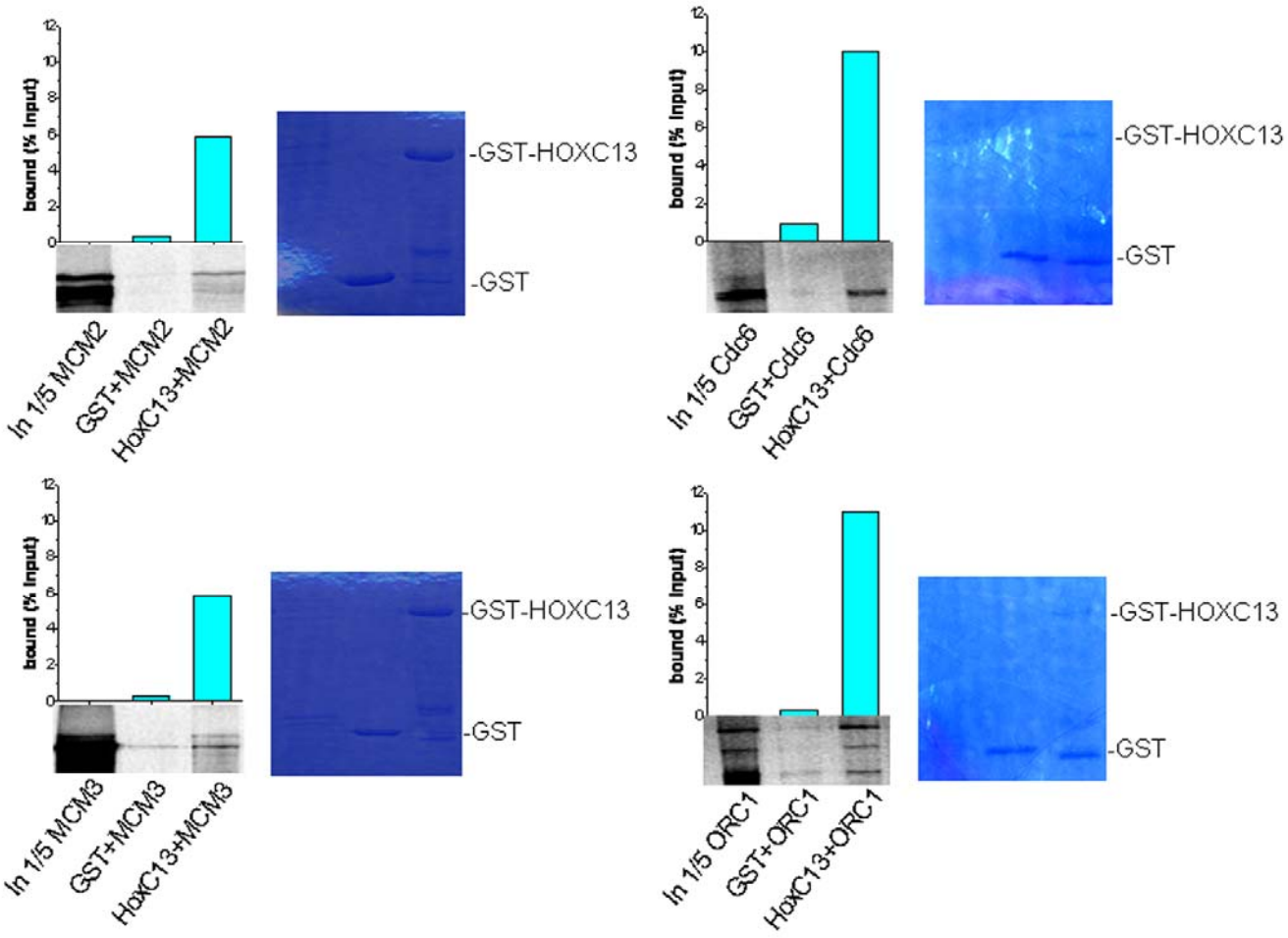

Figure 4.1. HOXC13 binds Cdc6 and ORC1 *in vitro*. GST or GST-HOXC13, immobilized on agarose beads were incubated with *in vitro* translated [^{35}S]-labeled MCM2, Cdc6, MCM3, and ORC1. Proteins retained by extensively washed beads were loaded onto a 10% acrylamide-SDS gel and the gel exposed to Cyclone screen. The input lanes contain a fraction of the radiolabeled proteins prior to binding (IN). Graphs show the amount of bound radiolabeled proteins as % of the input. On the right of each autoradiography, Coomassie-stained gels show the amount of proteins used in the assays.

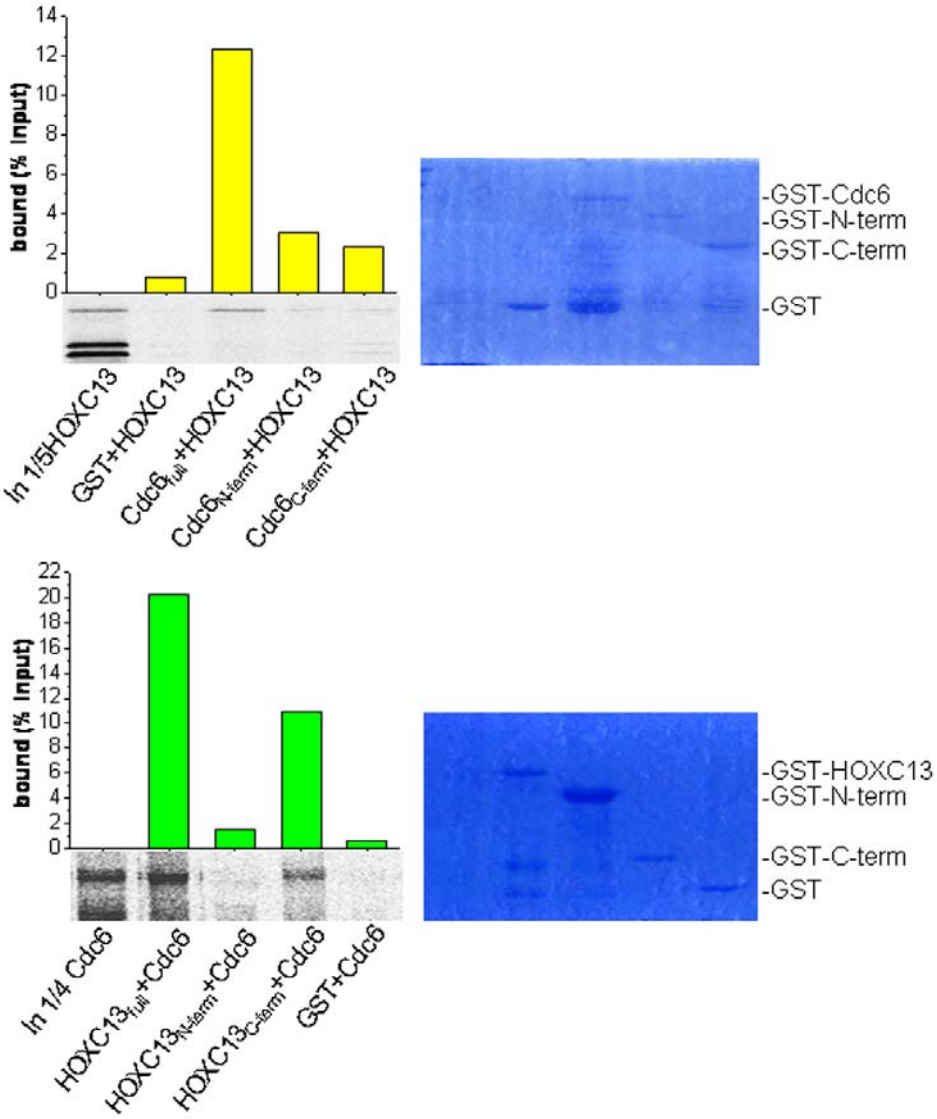

Figure 4.2. HOXC13 interacts by its homeodomein with Cdc6. In the upper panel, [^{35}S]-labeled HOXC13 was incubated with GST, GST-Cdc6 full length, GST-Cdc6 aminoterminal or GST-Cdc6 carboxyterminal and processed as described above. In the lower panel, [^{35}S]-labeled Cdc6 was incubated with GST-HOXC13 full length, GST-HOXC13 aminoterminal, GST-HOXC13 carboxyterminal or GST, and processed as already described. The input lanes contain a fraction of the radiolabeled proteins prior to binding (IN). Graphs show the amount of bound radiolabeled proteins as % of the input. On the right of each autoradiography, Coomassie-stained gels show the amount of proteins used in the assays.

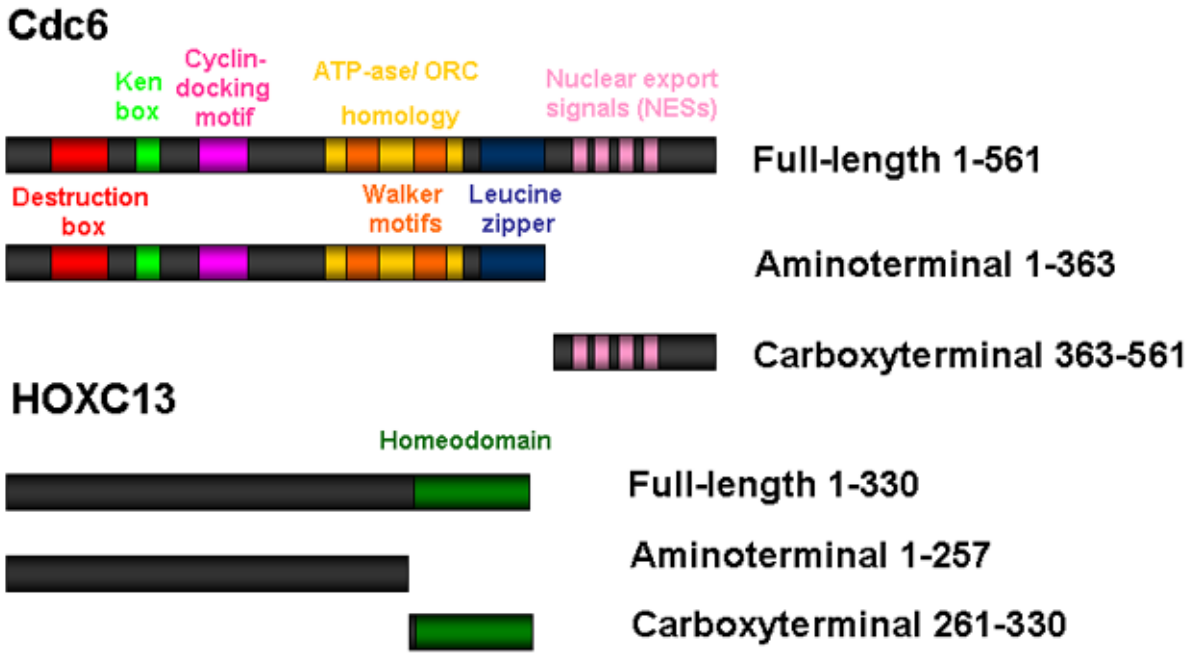

Figure 4.3. Representative scheme of Cdc6 and HOXC13 aminoacidic domains used in the GST pulldown assay shown in figure 4.5.

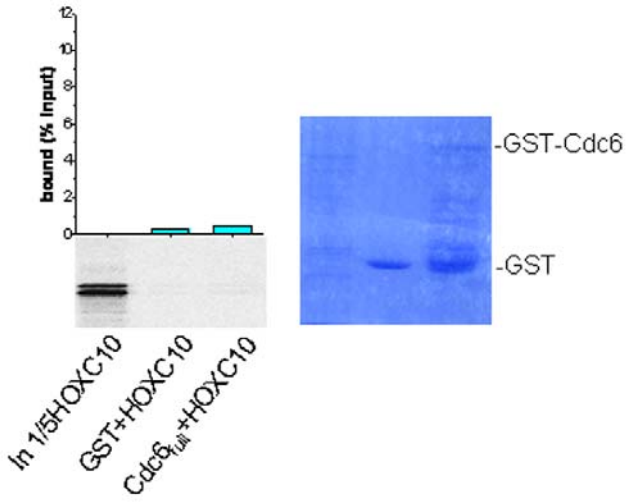

Figure 4.4. HOXC10 does not interact with Cdc6 *in vitro*. GST or GST-HOXC10, immobilized on agarose beads were incubated with *in vitro* translated [^{35}S]-labeled Cdc6 and processed as described above. The input lanes contain a fraction of the radiolabeled proteins prior to binding (IN). Graphs show the amount of bound radiolabeled proteins as % of the input. On the right of each autoradiography, Coomassie-stained gels show the amount of proteins used in the assays.

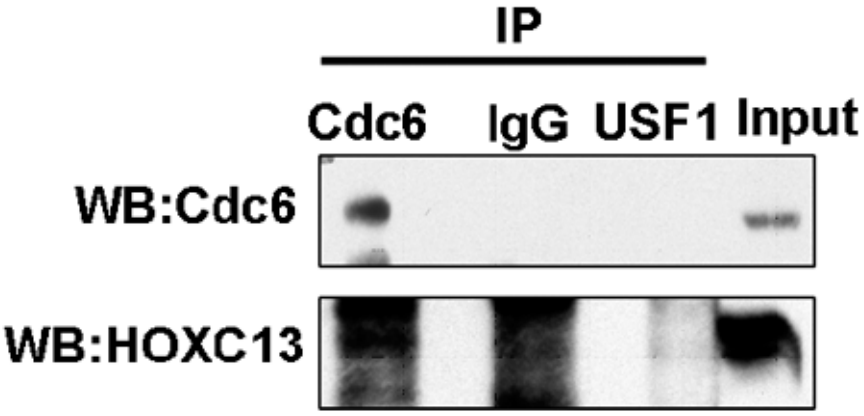

Figure 4.5. HOXC13 binds Cdc6 *in vivo*. Extracts of T98G (Input) cells were immunoprecipitated with antibodies against Cdc6 (IP Cdc6), IgG (IP IgG) and antibodies against USF1(IP USF1) as controls. The immunoprecipitated material was divided for separately western blotting analysis and revealed with antibodies against Cdc6 and HOXC13.

2.5. Binding of HOXC13 to lamin B2 origin

HOXC13 was identified in a one hybrid screening assay for its ability to specifically interact with the human lamin B2 origin. This interaction was confirmed *in vitro* by electrophoretic mobility shift assay (EMSA) and by immunological detection on origin DNA on agarose beads, and in cell milieu by CAT assay[145, 148]. We want to answer the question whether this protein is actually bound *in vivo* to this origin. In order to investigate this possibility we used the technique of chromatin immunoprecipitation (ChIP) in asynchronous cultures of the human glioblastoma T98G cell line, in which HOXC13 is abundantly expressed and this origin is fully functional, as previously shown by our group.

The cells were treated with 1% formaldehyde and the extracted DNA cross-linked protein material was chopped with subsequent micrococcal nuclease digestion and sonicated to a DNA fragment size ranging from a maximum of 1000 bp to a minimum

of about 160 bp, approximately corresponding to a nucleosome. The cross-linked and shared chromatin was then immunoprecipitated with our specific antibodies against HOXC13. These antibodies only recognize the N-terminal portion of the protein. This characteristic was essential to increase its specificity for the target protein and avoid interference with the DNA-recognition sequences that are mediated by the homeodomain of HOXC13. We analysed the immunoprecipitated HOXC13 by western blot analyses as shown in figure 5.1, and then we decross-linked DNA from the immunoprecipitated material. The resulted DNA was then quantitatively assayed by competitive PCR to determine the presence of lamin B2 origin and origin-unrelated sequences by B48 and B13 probes respectively[150]. As shown in figure 5.2, the immunoprecipitated DNA is significantly enriched in origin sequences relative to origin-unrelated DNA, demonstrating that HOXC13 protein is bound to the lamin B2 origin in dividing cells. The USF1 transcription factor, known to be constantly present on the promoter located upstream of the TIMM13 and downstream the lamin B2 gene, resulted also enriched and bound to the origin area, as expected The relatively lower degree of enrichment, besides the possible lesser efficiency of the antibody, may be due to the relatively small size of the DNA fragments, considering that the USF1 binds approximately 260 bp away from the start site and approximately 50 bp to the right of the border of the pre-replicative complex.

Since we found HOXC13 bound to lamin B2 origin in dividing cells, we wonder if its presence is connected or not with the DNA replication, thus we analysed whether HOXC13 is also bound to the origins in non-cycling T98G cells arrested in G0 phase by serum starvation. In this context, the USF1 factor was found bound to the downstream promoter close to the lamin B2 origin, whereas the HOXC13 protein results to be completely absent on the same origin in quiescent cells (figure 5.2), in agreement with the known absence of replicative complexes on the origin in these conditions. This is not due to an absence of the protein in non-cycling cells, because these contain the same level of the protein as the cycling ones, as previously shown.

Hence, not only the binding of HOXC13 with the lamin B2 replication origin, but also its absence on the origin in quiescent cells, suggest a role of HOXC13 in DNA replication.

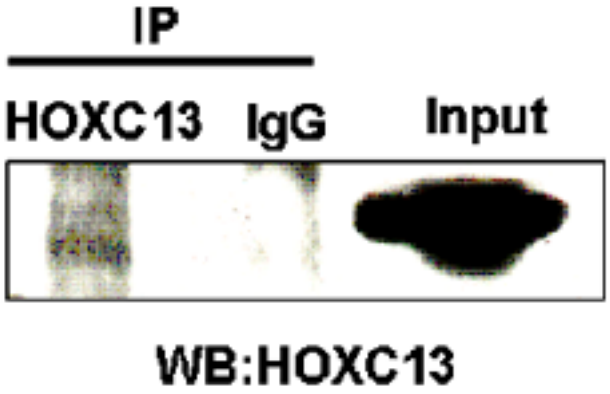

Figure 5.1. Immunoprecipitation of endogenous HOXC13. Immunoprecipitation experiments were performed with lysates from asynchronous crosslinked T98G cells (Input) using the specific antibody for HOXC13 and immunoglobulin (IgG) as control. The immunoblot analyses were performed with anti-HOXC13 antibody.

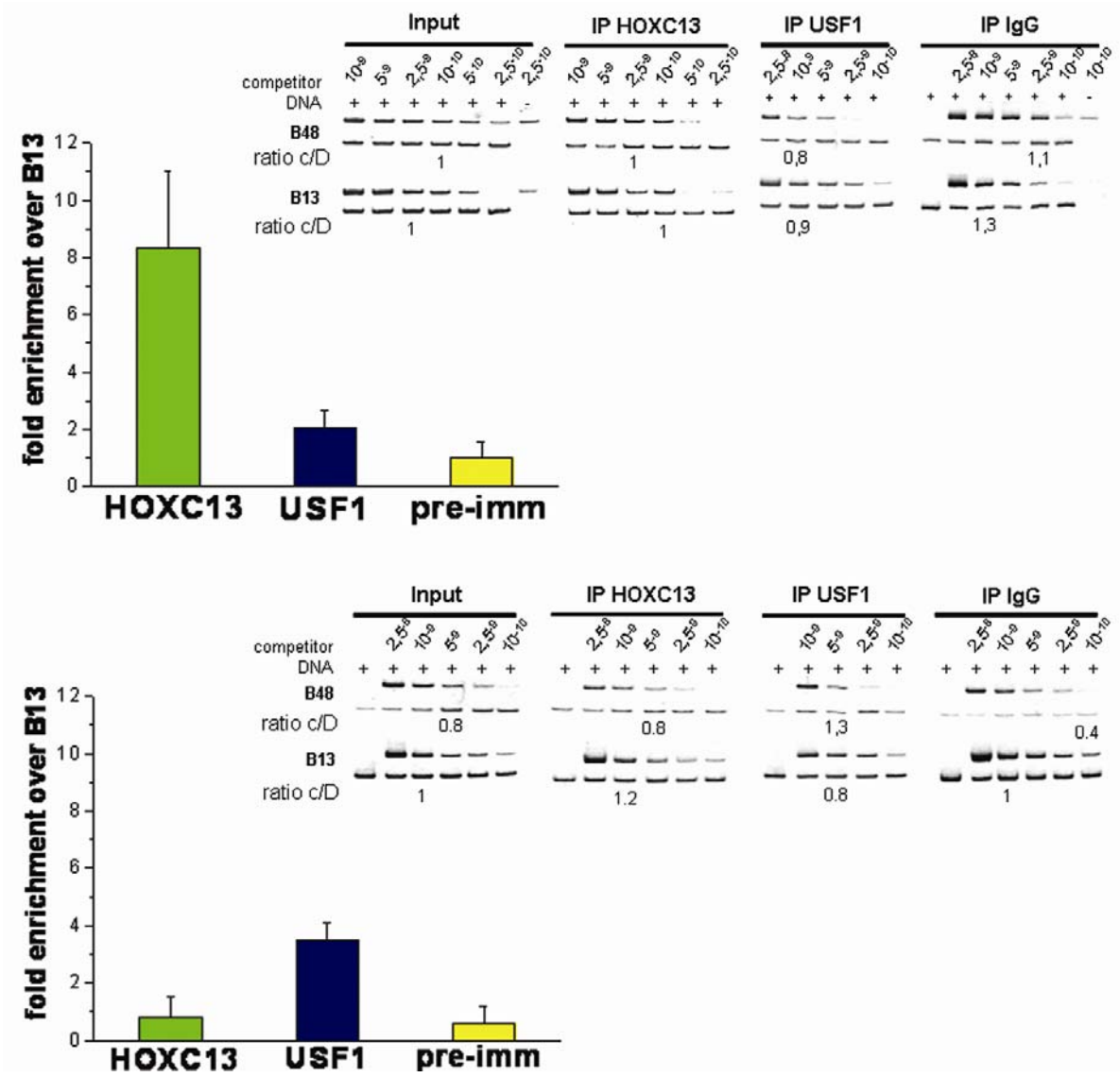

Figure 5.2. Binding of HOXC13 and USF1 proteins to the lamin B2 origin. ChIP analysis was performed on cross-linked chromatin from T98G cells and immunoprecipitated with the indicated antibodies. The immunoprecipitated DNA was then quantitatively assayed by competitive PCR to determine the presence of lamin B2 origin and origin-unrelated sequences by B48 and B13 probes, respectively. The histograms report the relative enrichment of B48 sequences over the B13 ones in the different immunoprecipitates, as derived from the analysis of the PCR reactions shown in parallel. The enrichment values obtained by a pre-immune serum control are also shown. Upper panel: results in asynchronously growing cells; bottom panel: cells starved in G0.

2.5.1. Cell cycle dependent binding of HOXC13 to lamin B2

Next, we asked whether HOXC13 binding to lamin B2 origin occur in a cell cycle dependent fashion, as expect for an interaction related with the origin function. For this purpose, a more precise and efficient cell synchronization in all phase of the cell cycle was used. T98G cells were synchronized by serum starvation and drugs. Then cells were tested for the binding of HOXC13 to the origin by ChIP assays, as already described. Competitive PCR analysis was used to measure the enrichment in the lamin B2 origin sequences in the precipitated DNA.

The G1 cells were achieved by the release of G0 cells from serum starvation. Late G1 cells were obtained by mimosine treatment, G1/S border and the S phases arrested cells were obtained by release from mimosine block. While G2/M arrested cells were obtained by sequential treatment with thymidine and nocodazole drugs.

As shown in the graph in the figure 5.3, HOXC13 appears bound to the origin region at the beginning of G1 and reaches a peak value at the G1/S transition, indicating that HOXC13 binding to the lamin B2 origin mostly occurs in G1 phase. As the cells enter into S phase, HOXC13 leaves the origin and then returns to it only at the next G1 phase.

Our data clearly demonstrate that HOXC13 binds to lamin B2 origin in a cell cycle dependent manner and its maximal binding occurs at the G1/S border, consistent with the time of origin activation expecting for an early firing replication origin like lamin B2. Therefore, this observation clearly suggests a direct involvement of HOXC13 in DNA replication activity.

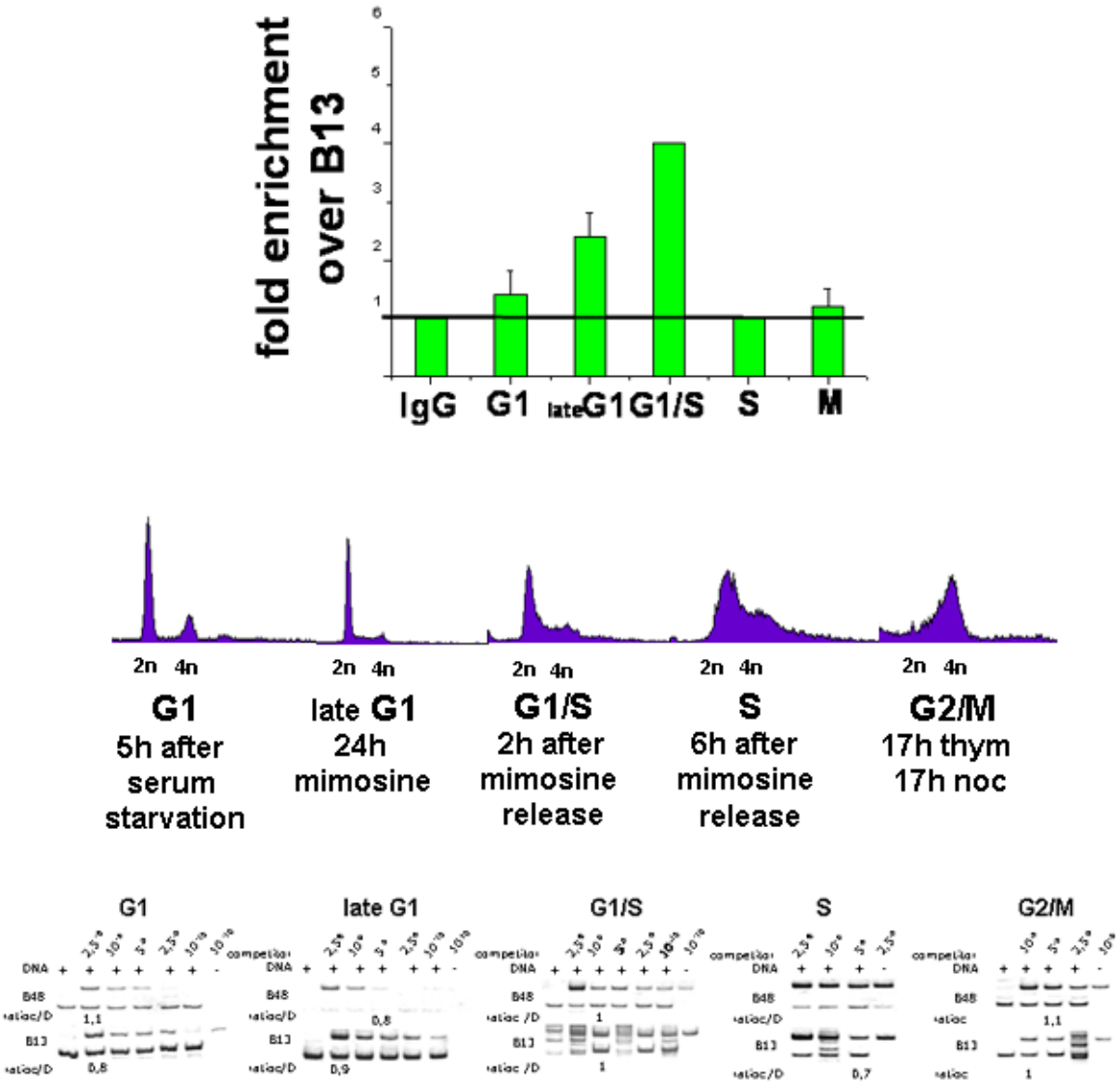

Figure 5.3. Temporal binding of HOXC13 to the lamin B2 origin by ChiP analysis. The histograms reported in the upper panel show the relative enrichment of B48 sequences over the B13 ones in the different cell populations. The relative enrichment reported in the histograms derived from the analyses of competitive PCR reactions showed in the bottom panel. The line in the graph indicates the threshold enrichment level obtained by using rabbit IgG antibody as negative control. Cell synchronizations were confirmed by FACS analyses using a portion of the cells from the synchronized samples, stained with propidium iodide (middle panel).

2.5.2. Analysis of the lamin B2 origin structure by footprinting assay

In order to confirm a direct interaction of HOXC13 with lamin B2 origin DNA, we utilized a combination of ChIP and *in vitro* DMS footprinting assay. The position of HOXC13 was explored in more detail by this fine technique, modifying the cross-linked, immuno-purified DNA to dimethyl sulfate (DMS) and analysing the produced fragments after degradation by ligation-mediated PCR (LM-PCR) technique. A clear footprint was observed (figure 5.4.) on the lower strand within nucleotides 3910 and 3990, corresponding to the area protected by the pre-RC in G1[151]. This area includes also the sequence recognized *in vitro* by HOXC13[145]. The large protection observed in Figure 4B is certainly not due exclusively to this molecule, in fact the formaldehyde treatment produces also protein-protein cross-links and the footprint is most probably due to the covalent linking of several RC members, including HOXC13. Also, the quantitative variations in the course of G1 may in part reflect a greater exposure of HOXC13 epitopes as the RC evolves rather than an actual increase in amount of bound protein.

We can conclude that HOXC13 assembles in G1 together with the other pre-RC molecules on the origin and leaves it after firing. This suggests that HOXC13 must closely interact with other members of the pre-RC.

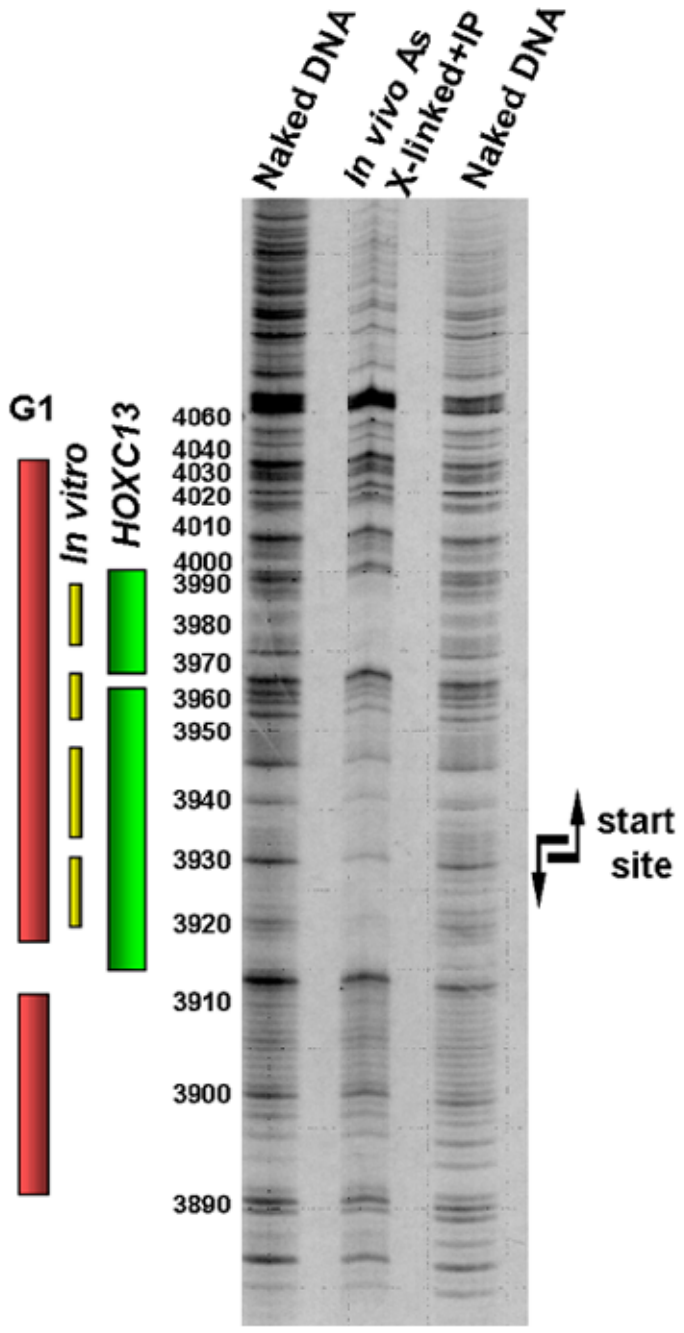

Figure 5.4. Analysis of *in vivo* HOXC13 interaction with the lamin B2 origin by a combination of ChIP and DMS footprinting. Footprint was obtained using *in vivo* cross-linked and sonicated chromatin. Chromatin was immunoprecipitated with anti-HOXC13 antibody and the resulting DNA-proteins complexes were subjected to footprint analysis by dimethyl sulfate (DMS) treatment and followed by ligation-mediate PCR (LM-PCR) of the produced fragments. DMS treated genomic DNA was used as control ladder for LM-PCR. In the figure, the bars indicate the covered area due to the binding of the prereplication complex in G1 phase *in vivo* (red) and *in vitro* (yellow) and the covered area due to the binding of HOXC13 in asynchronous growing HeLa cells (green).

2.6. Interaction of HOXC13 with other origins.

We showed that HOXC13 protein interacts with the lamin B2 origin and seems to be involved in origin function. Furthermore, HOXC13 interacts with two crucial proteins of the pre-RC complex, Cdc6 and ORC1, required for replication origin activity and involved in recognition and binding to replication origin sequences. Therefore, we asked at this point whether HOXC13 interaction is peculiar only for this origin or it is also present in other human replication origins, and whether those interactions are also linked to the proliferative state of the cells.

We probed two other replication origins identified in the recent years in the human genome, represented by the TOP1 origin, localized in the GpC island upstream the human TOP1 gene on chromosome 20[49], and MCM4 origin located in the genomic region close to or overlapping with the genetic control elements of the divergently transcribed genes MCM4 and PRKDC[152], that are both active in T98G cells.

Chromatin immunoprecipitation assays were performed on T98G cell line as previously described. As shown in figure 6, the resulted DNA was analysed by multiplex PCR using primers able to equally amplify the two origin sequences and the non-origin B13 sequence, contemporaneously. The PCR amplifications were performed on genomic DNA isolated from quiescent G0 cells (lanes of set A), on DNA derived from cross-linked chromatin isolated from G0 cells (set B), both synchronized by serum starvation, on DNA derived from cross-linked chromatin isolated from the same cells as set B but after immuno-precipitation with anti-HOXC13 antibody (set C), on DNA derived from cross-liked chromatin derived from asynchronously growing cells (set D), and on DNA derived from the same cells as Set D but after immuno-precipitation with anti-HOXC13 antibody (set E). The results show that each DNA sample of sets A, B, C and D contains, using previously established conditions, comparable amounts of the three probed sequences, also after the DNA treatment required for ChIP assay, whereas the sequences corresponding to the two origins are significantly enriched with respect to the B13 control following immuno-precipitation of the asynchronously growing cells (set E). Thus, we demonstrate that the HOXC13 protein interacts also with the human replication origins located TOP1 and MCM4, but only when the cells are cycling and not resting in quiescent G0 phase, according with the presence of the replicative complex on the origins in replicating cells.

This observation tends to downplay the possible role of this molecule as a sequence-specific transcription factor in the context of origin function. In fact, HOX proteins display a rather loose specificity for the DNA sequence bound, and moreover, the MCM4 and TOP1 origins do not display any similarity to the lamin B2 origin sequence, and, in particular, do not contain the tract of this sequence recognized *in vitro* by the HOXC13 molecule[145].

To further check the loading of HOXC13 preferentially on early firing origins to which all the three tested origins belong rather that late one, we analysed the binding of this protein with the late firing origin β-globin.

ChIP analyse was performed on asynchronous growing T98G cells and the resulted DNA was analysed by PCR performed using, in the same PCR reaction, primers able to equally amplify the β-globin origin sequences and the non-origin B13 sequence. By the comparison between the amplification products from cross-liked chromatin derived from asynchronously growing cells and the DNA derived from the same cells but after immuno-precipitation with anti-HOXC13 antibody, we found that after immuno-precipitation the sequence corresponding to the origin is not significantly enriched with respect to the B13 control, indicating that HOXC13 does not bind the late firing origin β-globin.

We conclude that the presence of HOXC13 on replication origins appear to be a common feature exclusively of the human early replicating origins, that also represent the majority of the origins.

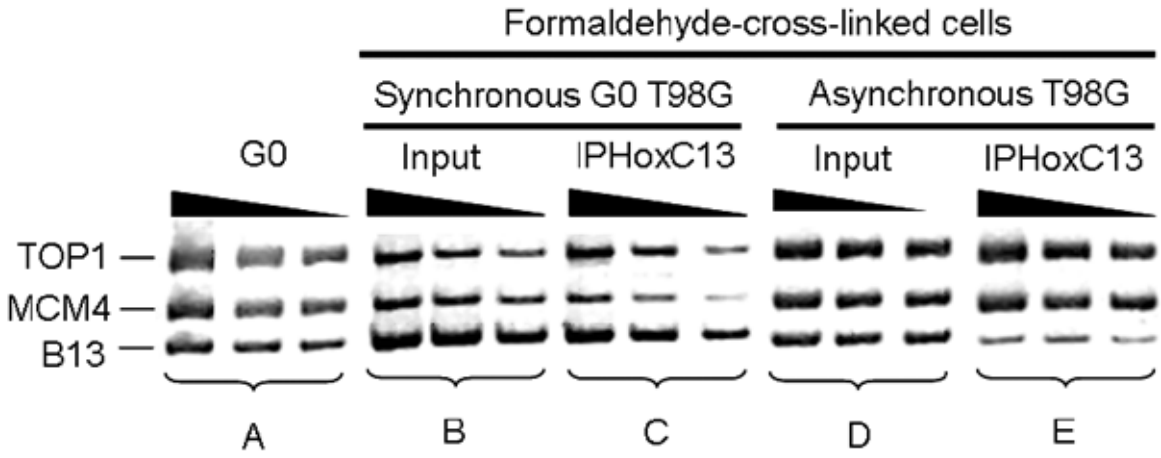

Figure 6.1. Binding of HOXC13 to TOP1 and MCM4 origins. The relative abundance of sequences corresponding to the TOP1 origin, to the MCM4 origin and to the B13 non-origin control was evaluated by probing the extracted DNA with sets of three serial two-fold dilutions of the samples, amplified through 35 PCR cycles using specific primers for the three sequences[153]. Set A: control DNA extracted from T98G cells in G0. Set B: DNA extracted from formaldehyde cross-linked T98G cells in G0. Set C: DNA extracted from formaldehyde cross-linked T98G cells in G0, after immuno-precipitation with anti-HOXC13 antibody. Set D: DNA extracted from asynchronously growing, formaldehyde cross-linked T98G cells. Set E: DNA extracted from asynchronously growing, formaldehyde cross-linked cells (as in set D), after immuno-precipitation with anti-HOXC13 antibody.

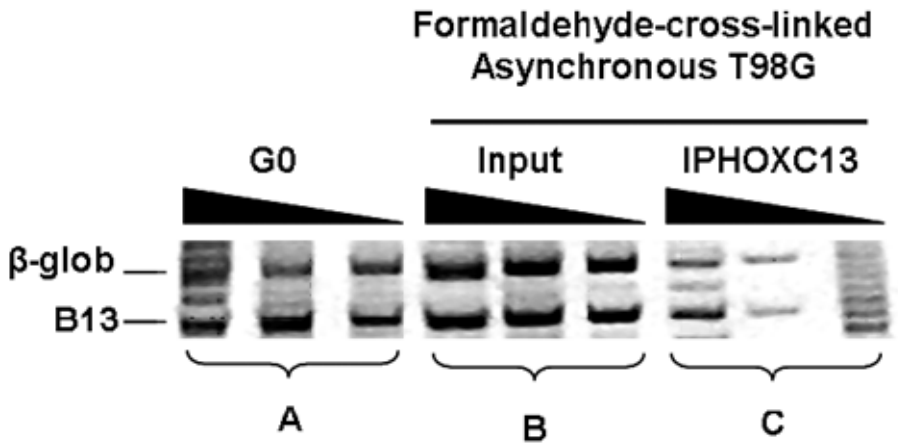

Figure 6.2. Binding of HOXC13 to the beta globin orgin. The relative abundance of sequences corresponding to the beta globin origin, and to the B13 non-origin control was evaluated by probing the extracted DNA with sets of three serial two-fold dilutions of the samples, amplified through 35 PCR cycles using specific primers for the two sequences[153]. Set A: control DNA extracted from T98G cells in G0. Set B: DNA extracted from asynchronously growing, formaldehyde cross-linked T98G cells. Set C: DNA extracted from asynchronously growing, formaldehyde cross-linked cells after immuno-precipitation with anti-HOXC13 antibody.

2.6.1. Cell cycle dependent binding of HOXC13 to TOP1 and MCM4

As for the lamin B2 origin, after the observation of the binding between HOXC13 and the two origins TOP1 and MCM4 occurred exclusively in replicating cells, we asked whether these binding change in relation with cell cycle progression. Chromatin immunoprecipitation assays were performed on T98G cell line synchronized by the same protocols previously described. As shown in figure 6.2, the resulted DNA was analysed by multiplex PCR as above.

The amplification was performed on genomic DNA isolated from quiescent G0 cells (lanes of set A), on DNA derived from cross-linked chromatin isolated from different phases of the cell cycle (lanes of set B), on DNA derived from cross-linked chromatin isolated from the same cells as set B but after immuno-precipitation with anti-HOXC13 antibody (set C). The results show that each DNA sample of sets A and B contains, using the same established conditions, comparable amounts of the three

probed sequences, and also the sequences corresponding to the two origins are not enriched with respect to the B13 control following immunoprecipitation in all the phases, except in G1/S border, where it was observed an enrichment of the DNA corresponding to TOP1 origin (set C). Thus, we demonstrate that the HOXC13 protein interacts with the human replication origins TOP1 in G1/S border, the same moment in which it bound the lamin B2.

About the MCM4 origin, it is not clear by this methods and using these time of observation, in what precise steps of the cell cycle it is bound by HOXC13.

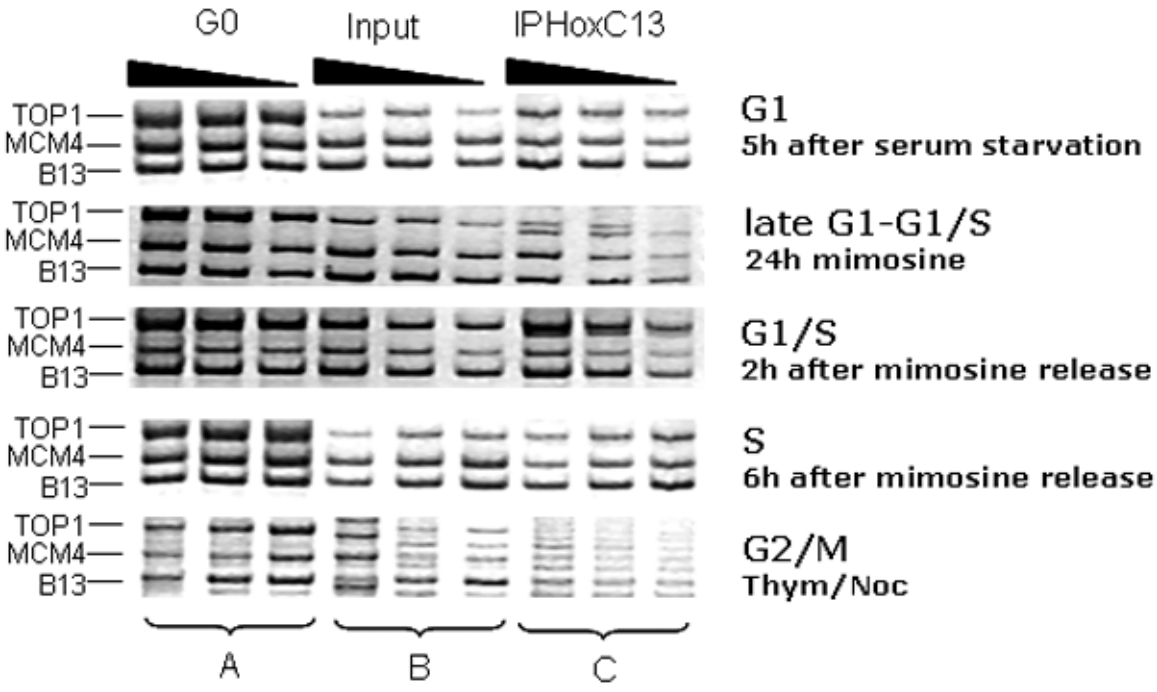

Figure 6.2. Binding of HOXC13 protein on the two origins TOP1 and MCM4 during the cell cycle. The relative abundance of sequences corresponding to the TOP1 origin, to the MCM4 origin and to the B13 non-origin control was evaluated by probing the extracted DNA with sets of three serial two-fold dilutions of the samples, amplified through 35 PCR cycles using specific primers for the three sequences[153]. Set A: control DNA extracted from T98G cells in G0. Set B: DNA extracted from formaldehyde cross-linked T98G cells in different phase of the cell cycle, indicated on the right of each panel. Set C: DNA extracted from formaldehyde cross-linked T98G cells in synchronized cells, after immuno-precipitation with anti-HOXC13 antibody.

2.7. UV-photofootprinting of origin structure

In eukaryotic cells the DNA replication origins are not defined by a single consensus sequence, as observed for budding yeast, but many features other than sequences seem to participate in the definition of a region as replication starting site. Also in the case of HOXC13 in fact, we observed a binding to replication origins that does not depend on DNA sequence, since the three analysed origins have not sequence similarity. Hence, other elements have to be involved in the binding of HOXC13 with the origins, like the chromatin structure, that in general is indicated as a factor influencing the chosen of the sequence as replication origin.

In order to investigate the role of DNA topology in origins specification, we analysed the impact of the disruption of chromatin organization at the origin sequence, inducing histone hyperacetylation by treating cells with trichostatin A (TSA), an inhibitor of histone deacetylase. The disturbance caused by TSA reduces origin activity of the lamin B2 origin of at least fifty per cent, as demonstrated by our group and also previously reported[71].

We applied the UV photofootprinting technique coupled with TD-PCR in synchronized cells. This technique is able to detect topological changes at or around the site of association of proteins with DNA. The UV irradiation causes pyrimidine dimer formation but the binding of proteins to DNA or changes in DNA flexibility block the induction of these dimer. In this way we can detect the topological DNA changes, induced by TSA treatment, at or around a site of association with proteins and furthermore this detection can be done directly in native cells.

HeLa cells were synchronized in M, G1 and S phases of the cell cycle and treated with TSA 100ng/ml for 1h and 4h. In this way we can analyze not only the topological changes happening in different phase of cell cycle, but also the influence of TSA in specific phase of cell cycle. For cell synchronization were used nocodazole in order to obtain M phase cells, a drug which inhibits microtubule polymerization, aphidicolin for G1/S border - early S cells, a DNA polymerase inhibitor, while middle of G1 cells were obtained replating M cells in complete medium for 5 hours. Hence, DNA was isolated after UV irradiation from synchronized HeLa cells without treatment or with 1h or 4h of TSA treatment. The DNA was subjected to TD-PCR with primers probing the upper and the lower strand. The region probed corresponds to the sequence covered by the replicative complexes of the lamin B2 origin. Figure 7.1 shows results obtained with specific primers, in different phase of cell cycle, with or without TSA. We measured,

using the image J program, the intensity of the bands visualized by autoradiography. The increasing in the intensity of the bands suggest that the TSA treatment induced modifications in DNA topology and revealed a dynamic variations of the protein-DNA interactions. We found that TSA treatment modifies significantly the pattern observed at the different moments of the cell cycle, particularly at the region around and at the start site. We can conclude that the TSA treatment induced modification in DNA topology in the analysed region of the origin replication.

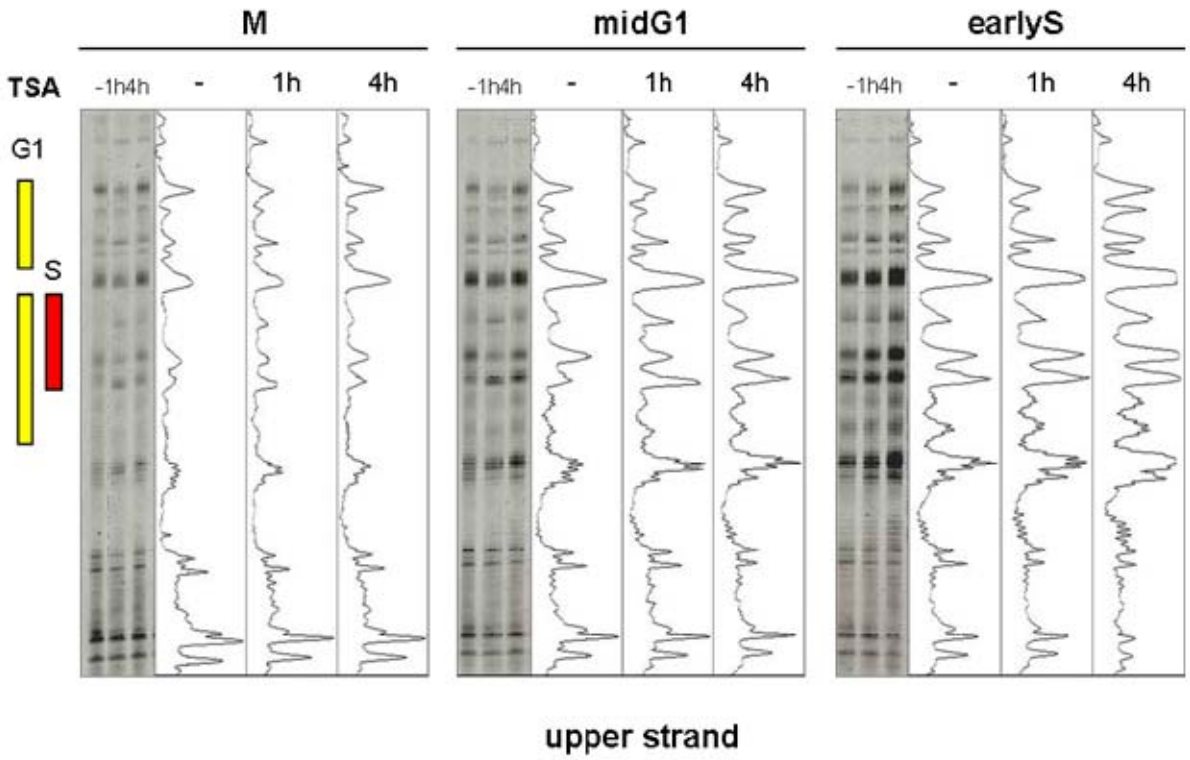

Figure 7.1. Photofootprinting analyses of a region encompassing the lamin B2 origin area in synchronized HeLa cells treated or not treated with TSA. The yellow and red bars indicate the covered regions in the G1 and S phases of the cell cycle. The TSA treatment was made for 1 and 4 hours. M, midG1 and early S are the phase of the cell cycle in which the cells were synchronized. After the UV irradiation, the DNA from these cells was subject to TD-PCR whit primers probing the upper and the lower strand (not shown). The intensity of the bands were analysed by Image J program and the resulted intensity were represented by the graph next to each PCR products.

2.7.1. Effects of disruption of origin chromatin

At this point we wonder if in addition to the showed reduction in the origin functionality and the observed modification in DNA topology, other effects could be detected in parallel on the replication origins after TSA treatment. Thus, we analysed the effects of the incubation with TSA on the ability of the origin area to bind some members of replicative complex and we found that this was reduced. In fact, by ChIP assay on T98G cells treated with TSA before the crosslink with formaldhyde, using antibody against Cdc6 and HOXC13, we observed a loss of binding between both, Cdc6 and HOXC13, with lamin B2 origin, a binding that indeed we observed without TSA treatment. Thus, as shown in figure 7.2, Cdc6 and its partner HOXC13 cannot bind the origin as a consequence of TSA treatment.

The TSA-induced disturbance in HOXC13 binding to the origin is observed also for the TOP1 and MCM4 origins that shown to be bound by HOXC13. Figure 7.3 reports the results of PCR performed using, in the same PCR reaction, primers able to equally amplify the two origin sequences and the non-origin B13 sequence, as already described for the identification of the binding between HOXC13 and the two origins TOP1 and MCM4. The comparison of the PCR products of set B (non-immuno-precipitated chromatin of cycling cells) and set C (the same, in presence of TSA) shows that the abundance of the three sequences, two corresponding to the origin sequences and the non-origin control B13, is not significantly changed; conversely, in the samples immunoprecipitated with anti-HOXC13 antibody, the ones obtained from TSA-treated cells (set E) lose the enrichment in origin sequences present in the non-treated ones (set D). Thus, an appropriate chromatin organization of the origin area appears to be required for the formation of an active replication complex. Notably, chromatin disruption does not affect other different protein-DNA interactions even if occurring in close proximity, but not involved in the prereplication complex formation. In fact, the binding of the USF1 transcription factor to the nearby promoter for the TIMM gene[147] is unaffected by chromatin disruption, as shown in figure 7.2.

we conclude that the structural features of chromatin at the origins appear to play an important role in origin function also through the modification of proteins binding, rather independently from the detailed sequence features.

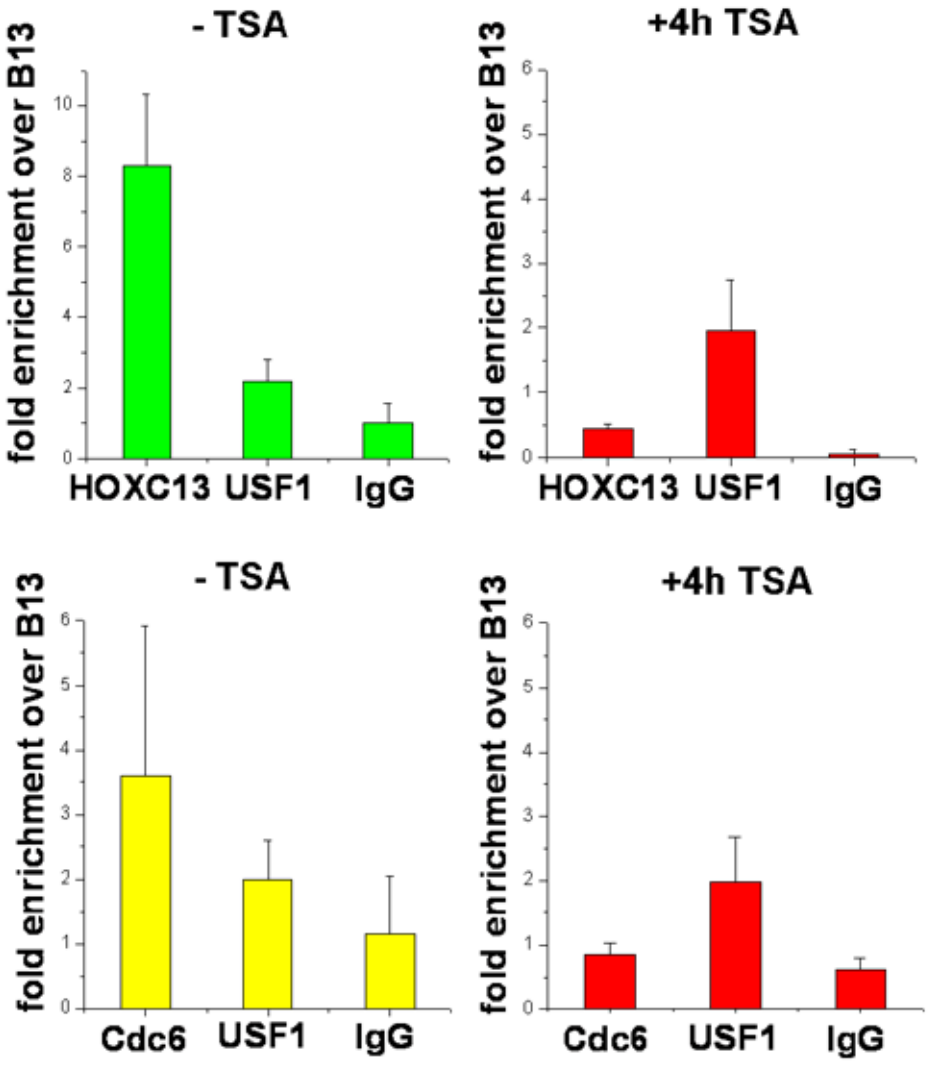

Figure 7.2. Comparison of ChIP performed without and with TSA treatment. The histograms report the relative enrichment of the B48 sequence over the B13 one in the different immunoprecipitates, as derived from the analysis of the competitive PCR reactions, using HOXC13 antibody (upper graphs) or Cdc6 antibody (bottom graphs). For each ChIP experiment USF1 and rabbit IgG were used as positive and negative control, respectively.

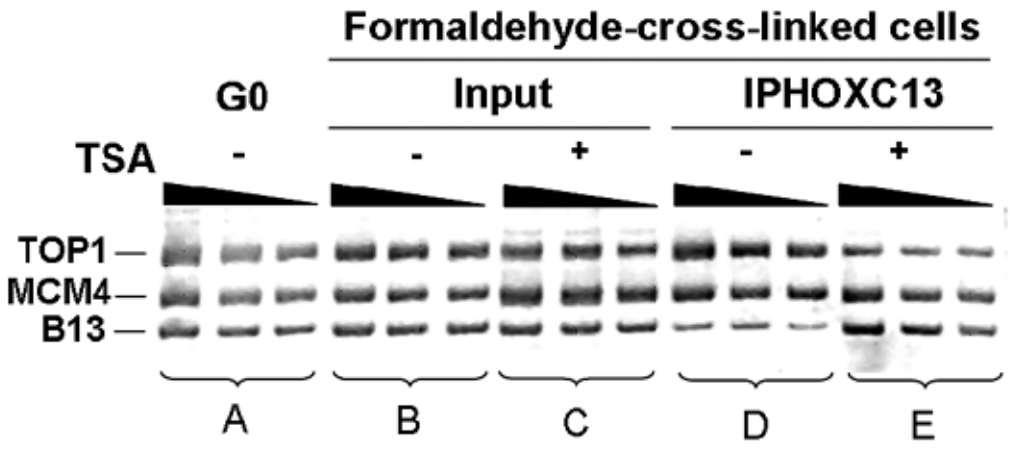

Figure 7.3. Association of HOXC13 with TOP1 and MCM4 origins *in vivo*. PCR was performed on genomic DNA isolated from quiescent cells (G0), on chromatin fragments from cross-linked asynchronously growing cells with and without TSA treatment (Input) and on chromatin fragments isolated after the immunoprecipitation with HOXC13 antibody with and without TSA treatment (IP HOXC13), in serial twofold dilution. The relative abundance of sequences corresponding to the TOP1 origin, to the MCM4 origin and to the B13 non-origin control was evaluated by probing the extracted DNA with sets of three serial two-fold dilutions of the samples, amplified through 35 PCR cycles using specific primers for the three sequences[153], as previously described.

2.8. Dispensability of HOXC13 on the lamin B2 origin

The general character of the interaction of HOXC13 with pre-RC proteins and human DNA replication origins, raises the question about the essential role played by HOXC13 for origin function. To answer to this question, we studied the effect of HOXC13 depletion in the T98G cells and in another cell line, the U2OS, derived from osteosarcoma, by either transient or stable shRNA transfection, using a lentiviral vector (Open Biosystems) and Polyfect tranfection reagent (Qiagen) following the manufacturer's instructions.

Cells were transiently transfected with lentiviral vector expressing shRNA against HOXC13 and harvest for immunoblotting and FACS analysis at 24h, 48h and 72h after the transfection. Control transfections with an empty lentiviral vector were

used as control. The efficency of the transfections or transductions were monitored by GFP expression from the lentiviral vector itself, detected by FACS analyses. We obtained at least 70% of the cells transfected in each experiment. Immunoblot analysis of RNAi experiments show marked depletion of HOXC13 after 48h from transfection with shRNA against HOXC13 in both T98G and U2OS cells.

As shown by FACS profiles, HOXC13 depletion did not significant modify the cell cycle aspects compared with the profiles from the control cells (figure 8.1). The same results were obtained using stable transfected cell lines using lentiviral vector expressing shRNA against HOXC13 and the empty lentiviral vector as control. Clones were obtained by drug selection for 12 days driving by the puromycin resistence gene included in the vector itself (figure 8.2). These results suggest a non essential role for HOXC13 in the cell cycle progression that is not surprising in light of the observation that in mice the mutation of HOXC13 alleles is compatible with life and only causes a defect in hair morphogenesis[105]; actually, HOX-gene knock-out often gives viable progenies[154]. This is suggestive of a redundant availability of other HOX proteins that can surrogate structurally and functionally the missing one.

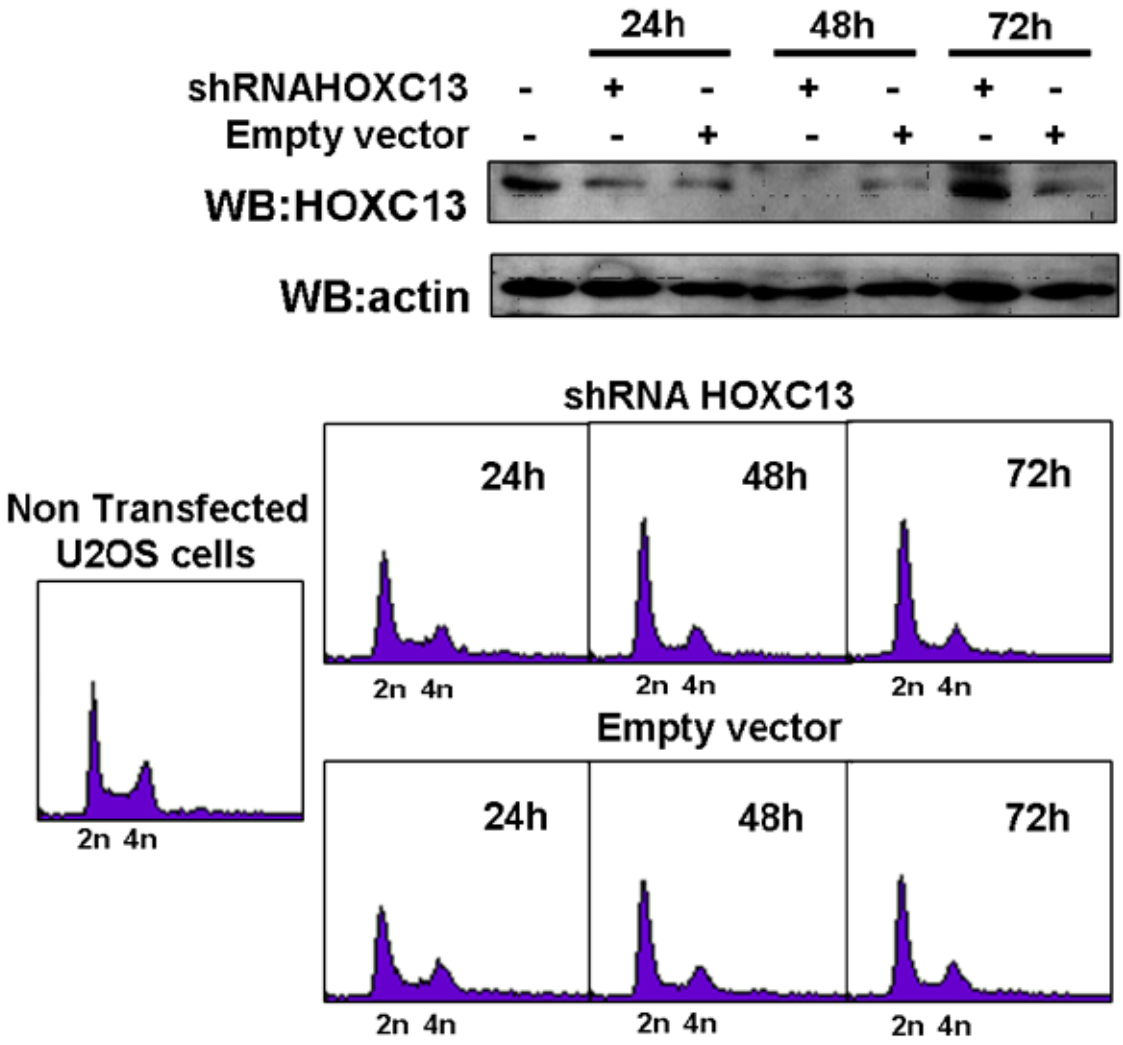

Figure 8.1. Transiently down regulation of HOXC13 in U2OS cells. Immunoblot analyses shows a nearly complete depletion of HOXC13 after 48h from transfection with lentiviral vector expressing anti-HOXC13 shRNA. Actin was used as a blot control (upper panel, lower blot). Flow cytometry profiles of U2OS cells before and after transient transfection with lentiviral vector expressing anti-HOXC13 shRNA or with empty lentiviral vector. The cells were harvested and stained with propidium iodide at 24h, 48h and 72h after the transfection.

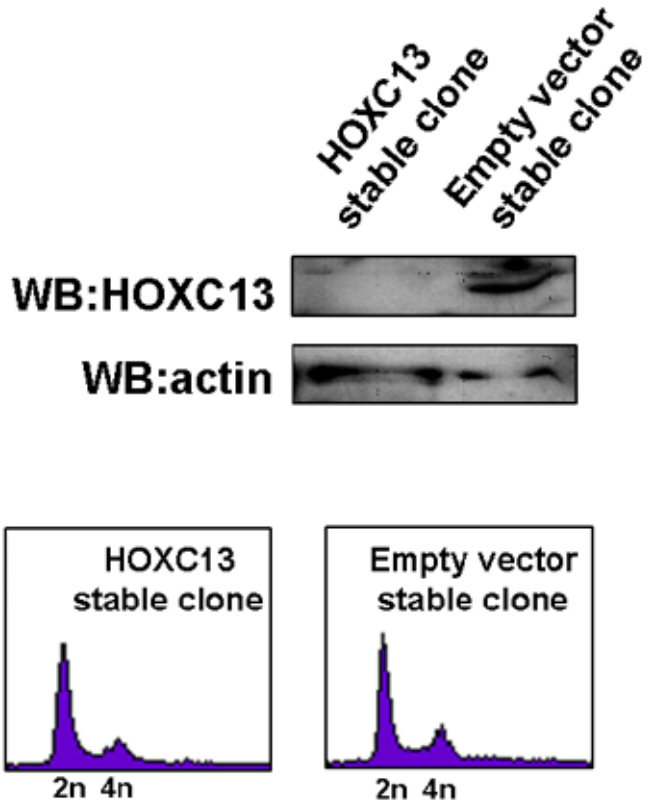

Figure 8.2. Stably downregulation of HOXC13 in U2OS cells. Immunoblot analyses shows a stably depletion of HOXC13. Actin was used as a loading protein control (upper panel). Flow cytometry profiles of stably transfected U2OS cells with anti-HOXC13 shRNA or empty vector (lower panel).

2.8.1. HOXC13 silencing do not alter the cell cycle progression

We next analysed more in detail the effects of the silencing of HOXC13 in cell cycle progression by the analyses of the DNA synthesis. For this assay we used the U2OS cells stable transfected with lentiviral vector expressing shRNA against HOXC13 and with empty vector, used as negative control, and we measured the DNA synthesis by bromodehoxyuridine incorporation. We analysed asynchronous growing cells and cells synchronized in different phases of the cell cycle by nocodazole, aphidicolin or mimosine.

In the case of asynchronous growing cells we pulsed U2OS stable clones with Brdu for 30 minutes or 1 hour. We didn't detect any differences in the Brdu

incorporation, and hence in the number of cells in active DNA synthesis, between stable clone expressing shRNA against HOXC13 and the stable clone with the empty vector.

So we went to analysed U2OS stable clones after synchronization to check the effects of HOXC13 depletion during the phases of the cell cycle in which HOXC13 result to be bound to the origins and in which the synthesis of DNA starts. In synchronized cells the Brdu was pulsed for 30 minutes before the collection. We were mostly interested in the analysis of the G1 and the G1/S transition phases. Hence, Brdu incorporation was measured 0, 2, 4 hours after the release from nocodazole, from what we obtained the G1 phase, and 6 and 8 hours after the release from what we obtained the G1/S transition and the entry in S phase. We also analysed the DNA synthesis in cells incubated with mimosine or aphidicolin, the blocks the cells in G1-G1/S and G1/S-S respectively and 4 hours after the release from the same drugs to obtain cells in S phase.

In all the cases, we did not observed a significant difference in the proportion of Brdu incorporation in all the analysed phase between the stable transfected with lentiviral vector expressing shRNA against HOXC13 and with empty vector. Especially, we did not observed any influence of the HOXC13 depletion in the entry in S phase, either after release from nocodazole (figure 8.19) neither in cells synchronized by mimosine or aphidicolin (figure 8.2 and 8.3), suggesting, as said in the previous paragraph, a non essential role of this protein in the DNA synthesis initiation and cell cycle progression. Likewise, the depletion of the others HOX proteins known to interact with replication origins is compatible with the continuation of the cell cycle. This is probably due to the high redundancy of the homeotic proteins and the strong conservation in the structure that allows the interchangeability between the HOX proteins.

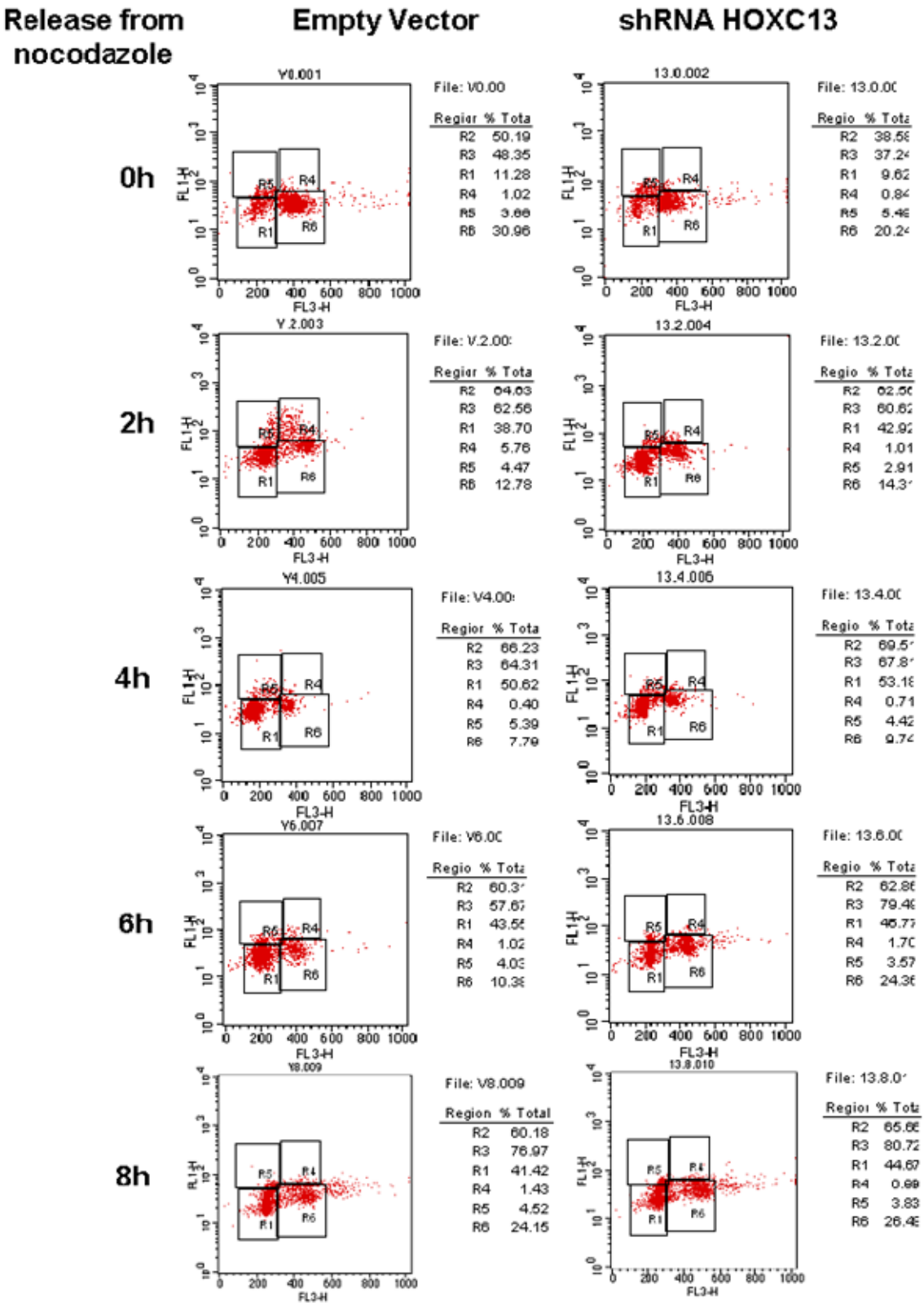

Figure 8.1. HOXC13 depletion does not affect cell cycle progression in cells synchronized by nocodazole. Flow cytometry profiles of synchronized U2OS stable clone with empty vector (V) and expressing shRNA against HOXC13 (13). The cells were collected directly after the release from nocodazole (0h) and 2h, 4h, 6h and 8h after the release. The cells were pulsed with Brdu and stained with

an anti-Brdu antibody and with Propidium Iodide and analysed by flow citometry. The percentage of cells in the different phases of the cell cycle are indicated: R2 indicates the total amount of the analysed cells; R1 indicates the cells in G1; R5 cells in early S; R4 cells in late S; R6 cells in G2/M.

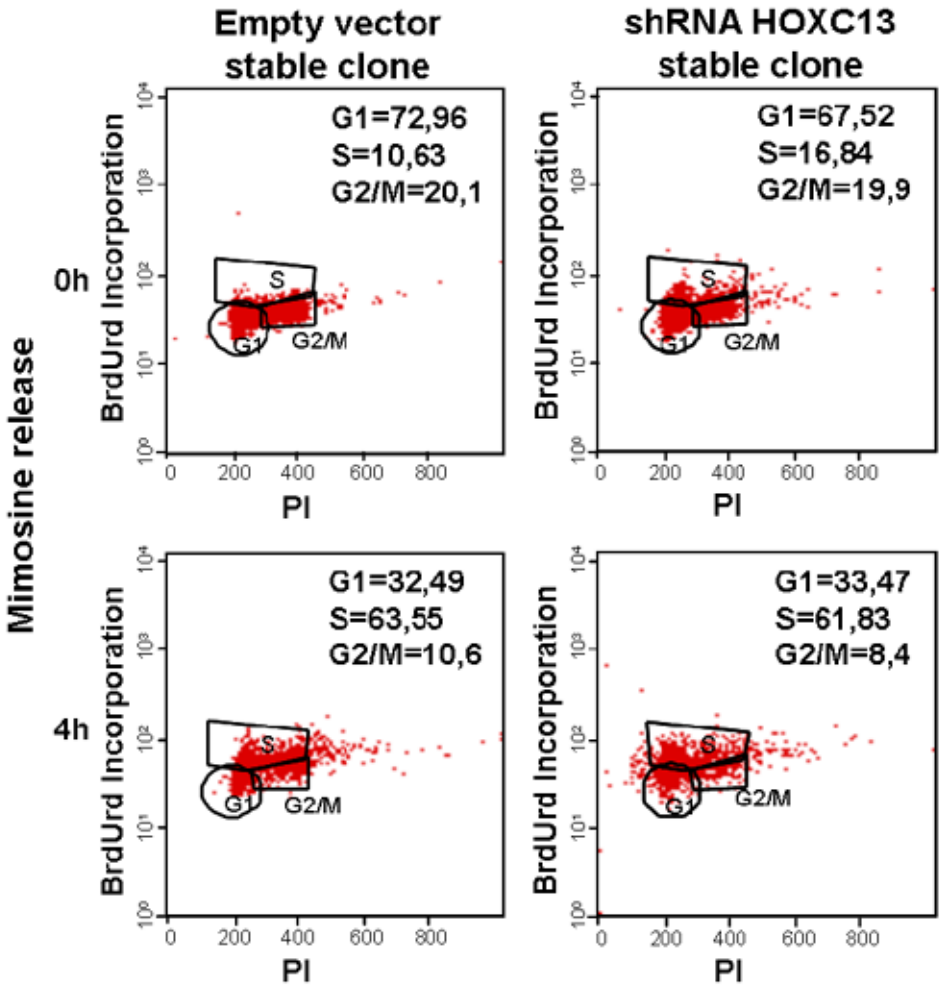

Figure 8.2. HOXC13 depletion does not affect cell cycle progression in cells synchronized by mimosine. Flow cytometry profiles of synchronized U2OS stable clone with Empty vector and expressing shRNA against HOXC13. The cells were collected directly after the release from mimosine and 4h after the release. The cells were pulsed with Brdu and stained with an anti-Brdu antibody and analysed by flow citometry after Propidium Iodide staining. The gated cell population in the different phases of the cell cycle and the number of cells in each gate are indicated.

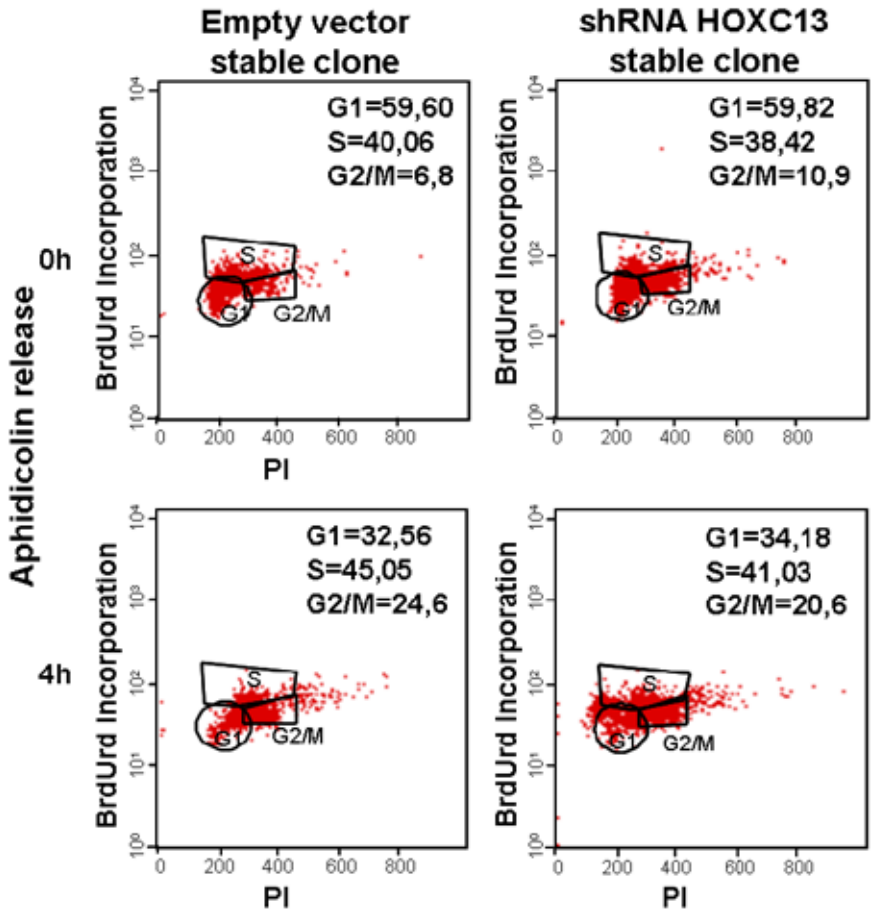

Figure 8.3. HOXC13 depletion does not affect cell cycle progression in cells synchronized by aphidicolin. Flow cytometry profiles of synchronized U2OS stable clone with Empty vector and expressing shRNA against HOXC13. The cells were collected directly after the release from aphidicolin and 4h after the release. The cells were pulsed with Brdu and stained with an anti-Brdu antibody and analysed by flow citometry after Propidium Iodide staining. The gated cell population in the different phases of the cell cycle and the number of cells in each gate are indicated.

3.Discussion

DNA replication is a complex, highly regulated process that leads to the duplication of the genetic material during cell division. Faithful transmission of genetic information to daughter cells requires a coordinate action of several processes and the recruiting of replication machinery on thousand sites that function as starting sites, called replication origins.

The knowledge of the mechanisms controlling the initiation of DNA replication is the key point to understand the mechanisms of regulation of DNA synthesis and cell proliferation. For these reason many efforts are being done to detect the factors involved in the initiation of DNA replication, how they are connected each other and how these connections are coordinate in space and time.

In bacteria cells and in budding yeast *Saccaromyces cerevisiae* it was found that the sites used as DNA replication origins were characterized by a conserved sequences recognized and bound by the proteins involved in DNA replication. In all the others eukaryotes however, the conservation of a consensus sequence used as DNA replication origin has never been found and the characterization of the starting sites, lacking a common consensus sequence, result to be more difficult. Also in human cells, currently, there are very few origins well characterized showing no sequence homologies between them.

In all eukaryotes, the origins are activated by the loading of the pre Replication Complex that lead to the formation of two replication forks moving in opposite direction to allow the synthesis of new DNA strands. The hexameric ORC complex is the factor that for first recognizes a sequence of DNA as replication origin and its binding allows the binding of all the other proteins of the pre replicative complex that lead to the origin licensing and then to origin firing. In the yeast *S. cerevisiae* the ORC complex recognizes and binds a specific AT rich regions, the ARS sequences, that function as starting site. AT richness is a feature generally observed for all eukaryotic origins and is assumed to be useful for facilitating the opening of DNA strands at the origins. The human ORC complex however, contrary from what observed in *S. cerevisiae,* has no sequence specificity[11]. Thus, its selective binding to DNA replication origin is probably modulated in a different way. The choice of regions to be used as origin in fact seems to be influenced by a combinations of elements depending not only on the presence of asymmetric AT-rich regions, but also on the presence of GpC island,

specific chromatin structure and its post-translational modification, the presence of binding sequences for specific transcription factors, the presence of cruciforms structures and specific DNA topology. The binding of ORC could be defined by these elements and also by the interaction with other unknown factors.

Moreover, among the mechanisms controlling the origins activity, a great relevance for cell cycle progression is assumed by those that limit the initiation of DNA replication to once per origin in each cell cycle. Not all origins are used during each cell cycle and those fired cannot be used a second time until the following cell division. Several mechanisms have been identified that limit initiation of eukaryotic DNA replication to once per origin for each cell division. Probably, these mechanisms evolved to avoid the genomic instability that would result if some sequences were duplicated more than others and replication forks remained during mitosis.

Many studies are ongoing with the scope of identification of other proteins involved in the specification and activation of DNA replication origins. To this purpose, well characterized and genetically defined origins were used to analysed the factors that influence the choice of a site as DNA replication origin. It was found that some of characterized origins are functional when moved to ectopic chromosomal locations, as for the lamin B2[150]. Hence, the origins seem to be site-specific and genetically determined in all organisms, but the nature and the role of the sequence determinants of origins might be given to others levels of organization of the origins[57] and even if the origins might be defined by specific combinations of sequence elements, it might not be a single consensus combination that defines an origin. Difficulties in studying whether specific sequence elements can initiate DNA replication might arise from a requirement for synergistic effects between separate, weak, but nevertheless specific, sequence elements that collaborate with other factors.

In particular, replication origins are shown to be frequently associated with coding genes and with CpG islands, regions that are largely bound by transcription factors. Thus, the transcription factors result to be between the factors involved in the regulation of origin selection and for them the suggested role in this process is the recruitment directly or indirectly of subunits of the pre replicative complex.

From the study of the mechanism regulating the specification and activation of DNA replication derives the indication of an involvement of the HOX proteins as factors able to play a role in replication origin activity. The HOX proteins are transcriptional factors involved in the development of the metazoan assigning positional

identities along the embryonic body. They are characterized by the presence of a very highly conserved homeodomain, that enables the HOX proteins to bind the DNA.

Beside the increasing evidence for a close association between transcription factors and DNA replication, a clear connection with HOX proteins and the machinery controlling the cell cycle progression have been identify also in the interaction with the regulator Geminin. Through the interaction with Cdt1, Geminin prevent the replication by the transiently association with members of the HOX-repressing polycomb complex, with the chromatin of HOX regulatory DNA elements and with the HOX proteins themselves[117].

One of the first indication of a direct connection between DNA replication and HOX proteins comes from our group, that in a previous work found a specific binding of three different HOX proteins to a portion of the lamin B2 origin. The portion of the lamin B2, previously showed to be bound by the pre Replication Complex in S phase, was used as bait in a one hybrid screening assay. This lead to the identification of three homeotic proteins: HOXC10, HOXC13 and HOXA13. For both HOXC10 and HOXC13 was further showed an high affinity *in vitro* for the origin, demonstrated by EMSA assay and CAT assay. About HOXC10, it was shown a timely degradation by anaphase promoting complex APC, that lead to the disappearance of the protein in the cells during the mitosis, and the affection of this degradation on the progression from metaphase to anaphase.

In the present study we wondered if the same affinity observed *in vitro* between HOXC13 and the lamin B2 origin is present also *in vivo* and how this protein could be involved in the DNA replication.

To analyse the behaviour of HOXC13 in human cells and its *in vivo* interaction with lamin B2 origin, our group produced polyclonal antibody against this protein. Rabbits were injected with a purified GST protein lacking the very highly conserved homeodomain, in order to generate a more specific antibody. First of all, it appeared essential to prove the goodness of this antibody in recognition clearly and with high specificity the right protein, considering the conservation and the redundancy of the these factors, especially for those homeoproteins derived from the same paralogue group. The many proves made demonstrate that the antibody is specific in antigen recognition also in a total cell extract and that it can be used for different application.

Hence, we obtained first general indications about the expression in different cell types, that is variable respect to the analysed cells, showing some cell lines with

irrelevant level or with big expression of HOXC13 protein. The cells with an appreciable level of HOXC13 also show a constant level in its expression. In the analysed cases, HeLa and T98G cells, in fact, during the entire cell cycle the amount of HOXC13 does not varies, as showed for HOXD13[149], another HOX protein interacting whit the human origins, which level is constant throughout the cell cycle, but differently from what already observed for the HOXC10[155].

In order to explore the direct interaction of HOXC13 with the human DNA replication origin lamin B2 we performed a chromatin immunoprecipitation assays and we found that HOXC13 binds *in vivo* to the lamin B2 origin. This interaction is also extensible to others origins, as we observed for the TOP1 and MCM4 origins, leading to infer that HOXC13 is stably bound to at least a good fraction of the origins, specifically the early replicating ones, to which all the three tested origins belong and that represent anyhow the majority. This hypothesis was further confirmed by our observation that HOXC13 does not bind the late firing replication origin β-globin.

HOXC13 is bound to origin chromatin, at least for the lamin B2 origin, at a precise site within the pre-RC and at a specific phase of the cell cycle. The interaction with the origin occur within the area covered by the pre-replicative complex in G1 phase and its binding is very close to the start site of leading strand synthesis and to the binding sites of ORC1, ORC2, Cdc6, topoisomerase I and topoisomerase II[47, 80]. Using synchronized cell cultures, we found that the protein is absent from the origins in mitosis and appears on them at the beginning of G1, reach a peak at G1/S, and in S phase, when the DNA synthesis starts, the interaction of HOXC13 with the origins fades, in parallel with the transition from the large pre-RC to a smaller and differently organized post-RC, that marks the end of the origin activity and prevent the re-licensing of the origins[30].

The mechanism by which HOXC13 interacts with the origins involved also its interaction with at least two important factors of the pre Replication Complex. We found indeed, by GST pull down and co-immunoprecipitation assays, that HOXC13 interacts with Cdc6 and ORC1, two crucial elements for origin recognition and licensing, that allow the pre-RC assembly. The observed interaction with Cdc6 is mediated by the central portion of Cdc6 and HOXC13 homeodomain. The same was found about HOXD13, where the interaction of the protein with Cdc6 is mediated again by its homeodomain[149]. The binding of HOXC13 with these two proteins is significant because these are the first factors involved in the origins licensing.

Since the connection between the morphogenetic function of homeotic proteins and the regulation of cell cycle progression, as well as the cancer role of HOX proteins in modulating cell proliferation was amply demonstrated[118, 122], we asked which specific role can we ascribe to the binding of HOXC13 to a large fraction of replication origins, in the context of the pre-RC assembly. It is tempting to hypothesize that homeotic complexes could contribute to the specification of origins in the genome, considering the lack of consensus sequence in metazoan origins, apart from a relative frequency of asymmetric AT-rich sequences[156]. Actually, also the HOX proteins recognize and bind the DNA without a stringent sequence specificity *in vitro* and preferentially bind AT rich regions. The HOX proteins achieve a high specificity *in vivo* by the cooperation with other DNA-binding cofactors. This observation suggests the hypothesis that the observed interaction of HOXC13 with the origins, in all probability, occurs in the context of a multi-protein homeotic effectors and other factors, rather than sequences, provide the appropriate conditions for the assembly of the pre-RC in defined starting sites. The finding that the interaction of the origins with HOXC13 occurs in the absence of sequence conservation, as demonstrate by the fact that the three origins lamin B2, TOP1 and MCM4 do not contain any motif similar between them, support this idea. This conclusion is further suggested also by the recent observations that other homeotic proteins as HOXC10, studied by our group, HOXA13, also previously found in the one hybrid assay, HOXD11 and HOXD13 bind the lamin B2 origin as well as other origins *in vivo*[6,7]; in this perspective, we should expect that, besides HOXC13 and the others already showed, other homeotic proteins may display an interaction with DNA replication origins. Depletion of one of these proteins however is compatible with the continuation of the cell cycle, probably thanks to the redundancy of homeotic protein structures that permit to a missing protein to be substitute in the complex by an analogous one, indicating a relatively generic function for the HOX protein rather than a precise requirement for a specific homeotic protein. In this view, it was not surprising our finding that depletion of HOXC13 does not alter cell cycle progression or S phase entry. The recent observation by Salsi *et al.* that depletion of HOXD13 instead delays the initiation of DNA synthesis represent the only case in which the depletion of an homeotic protein causes a delay in DNA synthesis initiation but this strong result could be due to a synergistic silencing effect when using siRNA targeted to highly conserved protein sequences like that of the homeodomain as in Salsi *et al.*. This further reinforce our model suggesting that the complete or partial depletion of several HOX proteins, at

least those paralogs to the HOXD13 one (the Abd-B paralogs, to which class all HOX proteins so far found to bind the replication origins belong to), and not the depletion of just one homeotic protein, lead to a delay of DNA synthesis.

Among the elements indicated as factors that influence the choice of a region as DNA replication origin, much relevance is given on the DNA topology and chromatin structure. It is widely accepted that chromatin structure around a promoter modulates the transcriptional activity of genes. Activators and suppressors of transcription often remodel chromatin by modifying or repositioning the histones. Histone acetylation is linked to opening chromatin for gene activation. Notably, histones around ORC binding sites are hyperacetylated during gene amplification in *Drosophila* follicle cells[14] and also nucleosome repositioning is involved in general chromatin remodelling events. These aspects are closely interrelated, affect each other and depend also on relatively long-range effects of DNA sequence, such as the tendency to form unusual DNA structures and to bend the duplex.

The lamin B2 origin sequence is characterized by an intrinsic tendency to form such unusual structures[157] and contains a bent DNA sequence between nt 3923 and 3928. Furthermore, topological status is clearly critical for origin recognition by ORC in yeast and *Drosophila* origins[8, 158], while the topoisomerases I and II, that are the modulators of DNA topology inside the cell, bind precise sites in the lamin B2 pre-RC area in precise moments of the cell cycle, topoisomerase I being essential for DNA synthesis initiation[80].

To analyse the influence of the topological state of the chromatin on the observed involvement of HOXC13, we study the effects of the Tricostatin A (TSA), the most potent inhibitor of histone deacetylase (HDAC). The acetylation of chromatin is an important post-translational modification that regulates the chromatin structure and it is the result of the activity of the histone acetyltransferases and deacetylases on the nucleosome.

We found that the disruption of chromatin structure caused by TSA treatment not only sharply reduces origin function, as already shown[71] but also disturbs the binding of replication complex members like the new HOXC13 and the well known Cdc6. After TSA treatment in fact, HOXC13 lose its affinity for all the three analysed origins. We observe the same for Cdc6, that we also found interacting with HOXC13, while the treatment does not affect the binding of USF1 to the *TIMM* promoter, that is located only 50 bp away from the right border of the replication complex, and that, contrary to

what happens with replication complex members, remains constant bound throughout the cell cycle.

This finding lead us to conclude that an appropriate chromatin organization of the origin area appears to be required for the formation of an active replication complex since chromatin disruption does not affect other different protein-DNA interactions even if these interactions occur in close proximity, but are not related with the pre-replication complex formation.

Previously, the positions of the human ORC1, ORC2 and Cdc6 at the lamin B2 origin were mapped with nucleotide resolution for the mitosis, middle of G1 and S phases of the cell cycle. Subsequently, the data obtained for topoisomerase I and topoisomerase II lead to new picture of the dynamic interaction of the origin binding proteins with the DNA[80].

Now our data indicate that HOXC13 represent a new member of the pre Replication Complex. Its function in the origin activity could be related to its specificity for DNA binding that might allows, in association with other cofactors, the specification of the sequence that have to be bound to start the DNA replication. Moreover, we further demonstrate that structural features of chromatin at the origins appear to play an important role in origin function also through the modification of proteins binding, rather independently from the detailed sequence features.

The important role of HOXC13 in DNA replication is also underlined by their involvement in oncogenesis. Principally, it was clearly demonstrate to determine an acute myeloid leukaemia when fused with NUP98 protein[128], such as observed for the other two components of the pre replicative complex, topoisomerase I and II.

Nevertheless, the emerging idea is that all the HOX proteins identified up to now, HOXC13 included, and the others that probably would be discovered on the origins, are not essential for the cell cycle progression, probably because the large redundancy of the HOX proteins in the metazoan cells. It appears indeed that the combination of the structural features in an appropriate cellular context, with certain essential proteins available and with accessibility to an HOX protein could be influence the origin selection.

3.1.Conclusion and future perspective

We show in this study that HOXC13 is a new member of the human pre Replication Complex. This protein in fact, directly interacts with two crucial factors of the pre replicative complex, ORC1 and Cdc6. About the interaction with Cdc6 we also found that the domains involved are the central portion of Cdc6 and the homeodomain of HOXC13.

Furthermore HOXC13 is bound to the early firing replication origins in a time dependent manner. Its binding reach a maximum in the G1/S border phase while disappear in S phase, when the cells start the DNA synthesis, and then remains unbound to the origin until the following G1 phase.

We also found that the binding of HOXC13 to the origins depend on DNA topology. After a TSA treatment in fact, we observed the loss of the interaction with all the tested origins. This behaviour is in common with the behaviour of one of the HOXC13 interacting protein, Cdc6, further demonstrating the involvement of this protein in the pre-RC. Conversely in fact, the binding of proteins bound near the pre replicative complex but not included in the complex itself, is not affected by the TSA treatment. We conclude that DNA topology strongly influence the binding between factors of the pre Replication Complex and DNA replication origins. This influence could be the key element in origin specification.

Since recently also others HOX proteins have been identify as proteins involved in regulation processes of DNA replication, we suggest that the interaction of HOXC13 with the origins might occur in a multi-homeotic proteins complex.

Hence, the sequel of this project could be the study of the interaction of HOXC13 with others factors of DNA replication, either the already known members of pre replicative complex and others HOX proteins.

It could be interest also investigate the presence of specific HOX cofactors in the pre-RC. The HOX proteins in fact, achieved a stringent sequence specificity cooperating with their cofactors. The presence of these cofactors directly in the pre replicative complex or associated with its members in specific moments preceding the loading of pre-RC could be explain also the presence of HOX proteins as driving factors for selection of specific sequences as DNA replication origins.

4.Materials and Methods

4.1.Cell culture, synchronization and TSA treatment

HeLa, T98G and U2OS cells were grown as monolayers in complete Dulbecco's modified Eagle's medium (DMEM), supplemented with 2 mM glutamine, 1 mM sodium pyruvate, 10 U/L penicillin, 10 µg/L streptomycin and 10% fetal bovine serum FBS (Gibco) and kept in a 5% CO_2 atmosphere at 37°C.

T98G cells were synchronized in G0 phase by incubating cells for 72h in DMEM serum free and collected at different times after the release in completed medium with 20% FBS to obtain synchronized cell populations at different stages of cell cycle[159].

In the analyses of the HOXC13 protein levels during cell cycle, T98G cells were synchronized and collected immediately after the 72h of incubation in serum free medium to obtain G0 phase population; 6h, 10h, 16h for the early, mid and late G1 respectively; 20h for G1/S transition; 24h for S phase; 28h for G2 and 32h for G2/M phase. To arrest the cells in G2/M phase, T98G were initially treated 17h with 2.5 mM thymidine (Sigma), release 9h in completed medium and treated 17h with 50 ng/ml nocodazole (Sigma) and subjected to mitotic-shake off for the collection.

For ChIP assays, the T98G cells were synchronized in G1 by collecting them 5h after serum deprivation; to arrest the cells in late G1, G1/S border and S phases, they were treated with 0.5 mM mimosine for 24h and collected directly, and immediately after 2h and 6h the mimosine release respectively. For cell synchronization in G2/M transition was also used the same thymidine/nocodazole treatment described before.

HeLa cells were synchronized in M phase by sequential treatment with 2.5 mM thymidine for 17h, washed and released in fresh medium for 9h and, finally blocked with 50 ng/ml nocodazole for 15h. For subsequent synchronization through the G1 to the S phase, mitotic HeLa cells were shaken-off, washed and released in fresh medium at different times: 5h for G1 phase, 10h for G1/S border and 15h for S phase.

For BrdU incorporation experiments, U2OS cells were synchronized by: 100 ng/ml nocodazole treatment, release in nocodazole-free medium and collected 0, 2, 4, 6, 8h after the release to obtain G2/M, early G1, mid G1, late G1 and G1/S transition. Cells in late G1 and early S, were obtained by 0.5 mM mimosine treatment, collected after treatment and release in mimosine free-medium and collected at 4h after treatment

respectively. Cells in early S and late S were obtained by 5 µg/ml aphidicolin treatment, collected after treatment and release in aphidicolin free-medium and collected after 4h of release, respectively. Cells were collected, stained with propidium iodide and analyzed by flow cytometry analysis on a FACSCalibur (Becton Dickinson) instrument, to determine the cell cycle profile based on DNA content, and follow S phase progression by incorporation of BrdU (SIGMA).

The HeLa and T98G stable clones expressing recombinant HOXC13 tagged with HA and Flag were produced by transfection of pFHIresNeo vector in which HOXC13 was previously cloned. The transfections were performed using Polyfect according to the manufacturer's instructions. The stably transfected clones were selected by adding G418 drug in the medium for at least 2 weeks.

For TSA treatment, asynchronously growing T98G cells were incubated for 4h with 100 ng/ml TSA (Sigma) in complete medium. Before the ChIP assay, the cells were washed twice with PBS supplemented with the same drug concentration and lysed in buffers containing the same amount of drug.

4.2.Antibodies

The following antibodies were used: rabbit polyclonal anti-HOXC13 produced and purified by immunization of rabbits with GST-tagged HOXC13 1-259, that lacks the conserved homeodomain; rat monoclonal anti-HA (Sigma), mouse monoclonal anti-tubulin (Sigma), mouse monoclonal anti-Cdc6 (clone D-1, Santa Cruz Biotechnology), mouse monoclonal anti-Cyclin A (Sigma), rabbit polyclonal anti-Cdc6 (clone H-304, Santa Cruz Biotechnology), rabbit polyclonal anti-USF1 (clone C-20, Santa Cruz Biotechnology), mouse monoclonal anti-β-Actin (clone C-4, Santa Cruz Biotechnology), mouse monoclonal anti-BrdU antibody (Abcam), anti-mouse secondary antibody Alexa Fluor 488 (Cat.A-11017).

4.3.Preparation of Nuclear Extracts

HeLa cells were washed in PBS, harvested and resuspended in 5 vol. of RSB buffer (10 mM Tris pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$). After 10 min of incubation on ice, samples were centrifuged at 1200 rpm for 10 min. The pellet was resuspended in 2.5 vol of RSB buffer with 0.2% NP40 and incubated on ice for 15 min. Cellular debris were removed by 3500 rpm centrifugation. The purity of the nuclei was checked by microscopy. Then the pellet was resuspended in 1 ml of RIPA (1% deoxycholic acid, 1% Triton X-100, 0.15 M NaCl, 0.05 M Tris PH 7.5) and the nuclei were incubated 10 minutes in ice. After 1 ml of RIPA 0.1% SDS was added to the lysate and incubated for 15 minutes in ice. The described buffers were supplemented with protease inhibitors 1 mM PMSF and protease inhibitors cocktail tablet (Roche).

The sample was then sonicated 3 times for 30 seconds and centrifuged to pellet debris.

4.4.Western blot analyses

The extracts and the immunoprecipitates were run on a 10% or 12% SDS-PAGE, the proteins transferred to a nitrocellulose membrane (Amersham) using a wet transfer system (Biorad). The membrane was block in 5% milk in TBST for 1h at room temperature. The primary and secondary antibodies were diluted in 5% milk in TBST. All the primary antibodies were incubated overnight a 4°C, with exception for antibody against HOXC13, incubated for 90 minutes at room temperature. The secondary antibodies were all incubated for 1h at room temperature.

For immunoprecipitations the membranes were blocked 1h with ReliaBLOT reagents (Bethyl Laboratories) to reduce the background from heavy chains. The membranes were then incubated for 1hour and 30 minutes with the primary antibody against HOXC13, diluted in the same kind of milk and TBST and followed by 3 washes in TBST (45 minutes in total). The secondary antibody, diluted in milk and TBST, was incubated for 1h and then 3 washes with TBST were repeated. All the steps were performed at room temperature. The SuperSignal® West Femto Maximum Sensitivity Substrate (Pierce) diluited 1:5 was used for signal detection.

4.5.Gradient gel

Linear gradient gels were made mixing two solutions with different acrylamide concentration. We choose a range of concentration of 9%-16%, based on the size of HOXC13. The mixture with lower acrylamide concentration was composed by 17.6 ml of H_2O, 10.1 ml Tris 1.5 M pH 8.8, 0.2 ml sodium thiosulfate 5%, 12 ml 30% acrylamide, 0.002 TEMED and 0.2 10% ammonium persulfate. The mixture with 16% of acrylamide is composed by 3.3 ml H_2O, 9.3 ml Tris 1.5 M pH 8.8, 4.3 ml of 85% glycerol, 0.18 ml sodium thiosulfate 5%, 19.7 ml 30% acrylamide, 0.002 TEMED and 0.2, 10% ammonium persulfate. Both solution were degassed under vacuum for 1 minute and poured in gradient maker. The mixed solution filled the gel apparatus using a pump speed. When the separating gel was prepared, a normal stacking gel was added.

4.6.Partial proteolytic peptide maps

To prepare samples for enzyme treatment, we run on a preliminary gel multiple lanes of the recombinat protein and 50 μg of total extract from T98G cells boiled in loading buffer. Recombinant protein bands from different lanes were the substrate for enzymatic digestions with different enzyme concentrations.

The preliminary gel is a normal SDS-polyacrylamide gel. This gel was stained with Coomassie and dry on a paper. The bands of interest were excised and the backing paper removed from each slice with a scalpel.

We prepared another polyacrylamide gel including 1 mM EDTA, 15% of acrylamide for separating gel and with a stacking gel of about twice the normal depth, 3 cm from the bottom of the well to the top of the separating gel. The wells of the stacking gel were filled with 0.1% SDS, 1 mM EDTA, 2.5 mM dithiothreitol and 0.125 M Tris pH 6.8. In this wells were pushed the dried gel slices previously prepared, and there were incubated for 10 minutes to allow the equilibration between slices and buffer.

Then the gel slices were overlay first with a mix of 20 μl of 20% glycerol, 0.1% SDS, 1 mM EDTA, 2.5 mM dithiothreitol and 0.125 M Tris pH 6.8, and then with another mix of 10 μl of 10% glycerol, 0.1% SDS, 1mM EDTA, 2.5 mM dithiothreitol, 0.001% bromphenol blue, and 0.125 Tris pH 6.8 containing a dilution of Trypsin.

The right concentration of Trypsin was previously checked to set the condition for the final experiment using different times, from 30 to 80 minutes of incubation and different dilutions, between 50 ng and 20 μg of this enzyme on recombinant HOXC13 protein. We obtained the best result using 20 μg of trypsin for 80 minutes. Thus we used 20 μg of trypsin in the mix..

The gel was run in normal running buffer at 125 V until the bromphenol blue formed a sharp line and had migrated about two-thirds of the distance into the stacking gel. At this point the run was turned off for 80 minutes to allow the protein digestion at room temperature.

Then the run was continued until the bromphenol blue reached the bottom of the separating gel and we proceeded with Western Blot analyses using antibody against HOXC13.

4.7.Chromatin Immunoprecipitation

Formaldehyde was diluted to 1% in PBS and added to the monolayer of cells at about 70% of confluence for 10 minutes. The reaction was stopped by glycine at a final concentration of 0.125 M. The cells were incubated on ice in hypotonic RSB buffer (10 mM Tris-HCl pH 7.5, 10 mM NaCl, 3 mM $MgCl_2$) for 10 minutes, then an equal volume of RSB with 0.2% NP40 was added for 10 minutes again. After washes with PBS, high salt NBS buffer (1 M NaCl, 10 mM Tris-HCl, 0.1 NP40,1 mM EDTA) and with low salt NBS buffer (0.1 M NaCl), the cells were collected by scraper in MNase buffer (20 mM Tris-HCl pH 7.5, 15 mM NaCl, 1 mM $CaCl_2$) and incubated on ice for about 1 hour. The described buffers were supplemented with protease inhibitors 1mM PMSF and protease inhibitors cocktail tablet (Roche).

At this point the cells were subject to micrococcal nuclease digestion with 125 U of MNase (Roche) for about 1-2x10^7 cells at 37°C for 15 minutes to limit the maximum lengh of each fragment to 160 bp to less than 1000 bp. Digestion was stopped adding SDS lysis buffer (SDS 10%, EDTA 0.5 M, Tris-HCl 1 M, pH 8). The cell lysetes were also sonicated 5 times for 30 seconds to shear further the DNA and centrifuged to pellet debris.

Protein concentration was estimated by Bradford assay (Pierce) and ~1.5 mg of protein was used for each immunoprecipitation. We performed a pre-clearing step

incubating the samples with 30 µl Protein A from Chromatin Immunoprecipitatio (ChIP) Assay Kit (Millipore) for 1h at 4°C and then the lysates were used for the immunoprecipitation with polyclonal rabbit anti-HOXC13 generated against GST–HOXC13: 1-260 that lacks the conserved homeodomain and affinity purified (Viva Science kit, Sartorius group), or with Cdc6 rabbit polyclonal anti-Cdc6 (clone H-304, Santa Cruz Biotechnology), with polyclonal antibody against USF1 as a positive control and with specific rabbit IgG's as negative control. After an overnight incubation with the antibodies, the immunocomplexes were recovered by incubation of 70 µl of Protein A for 5h at 4°C.

The immunocomplexes were washed with low salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl, pH 8, 150 mM NaCl), High salt buffer (500 mM NaCl), LiCl buffer (0.25 M 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 blotting and DNA extraction.

Before the DNA extraction, the immunopurified complexes were incubated at 37°C for 1h with Ribonuclease A enzyme to degrade RNA in the sample, and at 65°C over night with Proteinase K enzyme, to revert the crosslink and to digest and remove the proteins from the preparations. DNA was extracted with phenol:chloroform:isoamyl alcohol (Invitrogen), ethanol precipitated and resuspended in 30 µl of TE buffer. The quantification of the abundance of different sequences in immunoprecipitated DNA was performed by competitive PCR.

4.8.Plasmids

The human full length HOXC13 cDNA was constructed by PCR amplification of HOXC13 cDNA and cloned into pGEX 20T vector for GST fusion protein production and in pcDNA3 vector (Invitrogen) for in vitro transcription and translation experiments. The HOXC13 deletion mutants for the GST pulldown assays (1-257, 260-330) were obtained as *Bam*HI-*Eco*RI fragments. The HOXC13 full length, aminoterminal and carboxyterminal, the Cdc6 full length, aminoterminal and carboxyterminal (1-363, 364-561), ORC1, MCM2 and MCM3 cDNA were cloned into the pcDNA3 vector (Invitrogen) for in vitro transcription and translation.

For the stably expressing HOXC13 HeLa and T98G cell lines, full length HOXC13 cDNA was cloned into pFHIresNeo vector and transfected in the cells. This vector contains the NeoR resistance gene for mammalian selection in G418 drug and Flag and HA aminoterminal tags for tagging the target protein.

All constructs were verified by nucleotide sequencing before use.

4.9.GST pull-down assay

[^{35}S]-labelled proteins used for in vitro binding assays were produced by using the TNT Reticulocyte Lysate System (Promega) according to the manufacturer's instructions, by using the corresponding pcDNA3 vectors as templates. The recombinant GST fusion proteins were produced and purified from BL21 bacteria transformed with the respective plasmids. Bacterial cultures were grown in culture broth + ampicillin and protein production was induced with IPTG 1 mM for 4 hours at 30°C until the OD600 was between 0,6 and 0,8. Bacteria were then resupended in Cold lysis buffer (50 mM TrisHCl pH 8, 5 mM EDTA pH 8, 250 mM NaCl, 5% glycerol, proteases inhibitor) and sonicated. Bacterial lysates were mixed with a 50% slurry of glutathione cross-linked agarose beads and the GST-fusion proteins were allowed to bind the beads at 4°C on a rotating wheel for 1 hour. The suspension was then loaded on an empty plastic column, letting the unbound proteins pass through, and the beads were washed with lysis buffer. The purity and integrity of the proteins were routinely checked by SDS-PAGE and Coomassie blue staining. To remove contaminant bacterial nucleic acids, recombinant proteins were pretreated with nucleases (0.25 unit/µl DNase I and 0.2 µg/µl RNase) for 1 hr at 25°C in 50 mM Tris HCl, pH 8, 5 mM $MgCl_2$, 2.5 mM $CaCl_2$, 100 mM NaCl, 5% glycerol, 1 mM DTT. Then the GST fusion proteins immobilized on agarose beads were washed and resuspended in NETN buffer (20 mM Tris HCl, pH 7.5, 100 mM NaCl, 1 mM EDTA, 0.5% Nonidet P-40, 1 mM DTT/1 mM phenylmethylsulfonyl fluoride) supplemented with 0.1 mg/ml ethidium bromide to impede the possible formation of nonspecific interactions between residual DNA and proteins. 35S-labeled proteins was added and incubated at 4°C on a rotating wheel. After 1hr, bound proteins were washed five times with 1 ml of NETN buffer and separated by electrophoresis in an SDS-7% polyacrylamide gel. Dried gels were quantitated by phosphoimaging (Cyclone).

4.10.Competitive PCR analysis

Competitive PCR was performed using primer sets B48 and B13:

B48 Dx: 5'-GACTGGAAACTTTTTTGTAC-3'

B48 Sx: 5'-TAGCTACACTAGCCAGTGACCTTTTTCC-3'

B13 Dx: 5'-GCCAGCTGGGTGGTGATAGA-3'

B13 Sx: 5'-CCTCAGAACCCAGCTGTG-3'

The quantification of the abundance of the origin (B48) and non-origin (B13) DNA fragments was performed as described previously[150]. The B48 amplification reaction contained 10µl 5x Taq polymerase buffer Green (Promega), 3 mM $MgCl_2$, 1µl (10pmol) of each primer, 1µl dNTPs (10 mM each), 0,5µl GoTaq polymerase (2 U/µl) (Promega), 5µl of purified DNA, 5µl competitor DNA and H_2O up to 50µl.

The B13 amplification reaction contained 10 µl 5x Taq polymerase buffer Green (Promega), 1 mM $MgCl_2$, 1µl (10pmol) of each primer, 1µl dNTPs (10 mM each), 0,5µl Taq polymerase (2 U/µl) (Promega), 5µl of purified DNA, 5µl competitor DNA and H_2O up to 50µl. The amplification conditions for both B48 and B13 were:

In PCR cycle, an initial denaturation of 4 minutes at 94°C for 5 minutes was followed by 35 cycles with denaturation for 30 seconds at 94°C, annealing for 30 seconds at 56°C, polymerization for 30 seconds at 72°C, and a final extension for 5 minutes at 72°C.

A constant volume of immunopurified DNA was coamplified with decreasing amounts of competitor template. The competitor consists of a 130 bp stuffer DNA flanked by the target sequences for B13 and B48 primer sets and was obtained as already described[146]. Amplification products were resolved on 8% acrylamide and stained with ethidium bromide. The intensity of the amplification band was quantified with the NIH-Image program (version 1.62).

4.11.PCR with multiple pairs of primers

PCR primers were designed from the known sequences of the analysed origins TOP1[49], MCM4[152], Beta globin and B13 was used as non origin control:

TOP1 Dx: 5'-CACTGCCTAGCAGAGGGGCT-3'

TOP1 Sx: 5'-GCAGTTGTGTAACAGCCTAAGTTCG-3'

MCM4 Dx: 5'-GTCTGACCTGCGGAGGTAGTTTGG-3'

MCM4 Sx: 5'-TGGTCTCGAACTCCTGCGATCCCC-3'

Betag Dx: 5'-GTTGCCCATAACAGCATCAG-3'

Betag Sx: 5'-CTGCCGTTACTGCCCTGTGGG-3'

B13 Dx: 5'-GCCAGCTGGGTGGTGATAGA-3'

B13 Sx: 5'-CCTCAGAACCCAGCTGTG-3'

Three pairs of primers were used together in each PCR reaction. The conditions for the equal amplification of the three distinct fragments were set up using purified DNA from synchronized quiescent T98G cells. Three serial two-fold dilutions were used for the DNA of each sample.

The amplification reaction contained 10 µl 5x Taq polymerase buffer Green (Promega), 3 mM $MgCl_2$, 1µl (10pmol) of each of three pairs of primers, 1.25 µl dNTPs (10 mM each), 1 µl Taq polymerase (2 U/µl) (Promega), 10 µl of purified and diluted DNA and H_2O up to 50µl.

In PCR cycle, an initial denaturation of 4 minutes at 94°C for 5 minutes was followed by 35 cycles with denaturation for 30 seconds at 94°C, annealing for 30 seconds at 60°C, polymerization for 30 seconds at 72°C, and a final extension for 5 minutes at 72°C.

Amplification products were resolved on 8% acrylamide and stained with ethidium bromide. The intensity of the amplification bands was quantified with the NIH-Image program (version 1.62).

4.12.ChIP and dimethylsulfate (DMS) treatment

HeLa cells were crosslinked by adding 1% formaldehyde and incubation for 10 min. After stopping the crosslinking reaction by adding 0.125 M glycine and incubation for 5 min, the cells were pelleted at 2000 r.p.m. The cells were then washed twice in 25 ml ice-cold phosphate-buffered saline (PBS) including protease inhibitors. Nuclei were isolated by resuspending the cell pellet in 1 ml ice-cold swelling buffer (5 mM PIPES pH 8.0, 85 mM KCl, 0.5% NP-40 and protease inhibitors), split into two aliquots and incubated on ice for 10 min. Chromatin was fragmented by subjecting the nuclei to restriction enzyme digestion with 200 U *Pst*I for 4 h at 37°C and 100 U *Pst*I for an additional 16 h at 37°C. The nuclei were then incubated with 200 U RNase cocktail

(Ambion) and an additional 100 U aliquot of *Pst*I for 2 h at 37°C. Nuclei were pelleted at 4°C for 5 min at 5000 r.p.m. and lysed in 1 ml lysis buffer (1% SDS, 10 mM EDTA and 50 mM Tris–HCl pH 8.1) on ice for 20 min. The lysate was combined and transferred to a 15 ml conical tube and diluted with 9 ml dilution buffer (0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris–HCl pH 8.1, 167 mM NaCl and protease inhibitors). An aliquot of 500 µl protein A–Sepharose beads (Pharmacia) was added to the diluted nuclear lysate and incubated for 2 h at 4°C while rotating. The beads were pelleted for 10 min at 2000 r.p.m. and the supernatant was divided into three aliquots. An aliquot of 25 µl of the appropriate antibody (USF1 or NF-E2; Santa Cruz Biotechnology) or no antibody was added to the aliquoted supernatant and incubated at 4°C overnight while rotating. Protein A–Sepharose beads were washed twice in blocking buffer [3% BSA, 0.05% sodium azide, and protease inhibitor in 1× TE (10 mM Tris–HCl pH 8.1, 1 mM EDTA)]. The chromatin was then immunoprecipitated with 600 µl blocked protein A–Sepharose for 2 h at 4°C on a rotator. The immunoprecipitates were pelleted at 13 000 r.p.m. for 30 s and 1 ml of the no antibody supernatant was saved and labeled as 'input'. Half of the input chromatin was ethanol precipitated and resuspended in two aliquots of 20 µl ddH_2O and 800 µl DMS buffer (50 mM sodium cacodylate, 1 mM EDTA) and the other half was saved for the ChIP/PCR analysis. The supernatants of the samples precipitated with USF1 and NF-E2 antibodies were discarded and the pellets were washed by rotating at 4°C for 5 min with 1 ml each of low salt wash (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris–HCl pH 8.1 and 150 mM NaCl), high salt wash (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris–HCl pH 8.1 and 500 mM NaCl), LiCl wash (0.25 M LiCl, 1% NP-40, 1% sodium desoxycholate, 1 mM EDTA and 10 mM Tris–HCl pH 8.1) and twice with TE. Of the immunoprecipitates, 80% was resuspended in 800 µl DMS buffer and 20% was left in TE buffer for the ChIP/PCR analysis.

DMS treatment of the immunoprecipitated chromatin was performed using the Maxam and Gilbert guanine-specific sequencing reaction with 0.1% DMS for 15, 45 or 90 s at room temperature. The reaction was stopped by adding 50 µl DMS stop buffer (1.5 M sodium acetate pH 7.0, 1 M 2-mercaptoethanol), followed by two ethanol precipitations in a dry ice bath. The DMS-treated and non-DMS-treated chromatin was then eluted from the beads by incubating twice with 250 µl elution buffer (1% SDS, 0.1 M $NaHCO_3$), shaking at 1000 g for 15 min at 65°C, each time saving the supernatant. An aliquot of 200 mM NaCl was added to the eluates and crosslinking was reversed by

incubation at 65°C for 5 h. Proteins were digested with 40 µg/ml proteinase K in 10 mM EDTA and 40 mM Tris pH 6.5 for 1 h at 37°C. Immunoprecipitated DNA was purified using a Qiagen kit and eluted with 180 µl ddH_2O. To cleave the DMS-treated DNA, 20 µl piperidine were added and incubated at 95°C for 30 min. The DNA was washed twice by adding 1 ml ddH_2O, dried in a Speed Vac and resuspended in 50 µl TE. Of the DMS-treated immunoprecipitated DNA, 10% was used for ligation-mediated PCR (LMPCR)-assisted *in vivo* footprinting. An aliquot of the precipitated DNA was also analyzed by PCR using primers specific for the murine β-globin downstream promoter region (forward primer, 5′-GACAAACATTATTCAGAGGGAGTACCC; reverse primer, 5′-AGGTGCACCATGATGTCTGTTTCTGG).

4.13.RNA depletion and stable clones production

T98G and U2OS cells were transiently transfected for 24h, 48h and 72h with lentiviral pGIPZ shRNA vector encoding a short hairpin RNA against HOXC13 (NM_017410) (Open Biosystems) by Polyfect tranfection reagent (Qiagen) following the manufacturer's instructions. RNA interference control experiments were performed using an empty lentiviral pGIPZ vector (Open Biosystems). The HOXC13 downregulated clone and the control clone were selected using puromycin (3µg/ml) after 48h of transient transfection followed by 12 days of selection.

4.14.Brdu incorporation experiment

Brdu incorporation experiments were performed on U2OS cells stable transfected with shRNA against HOXC13 and with empty lentivar vector (Open Biosystem). Asynchronous growing cells were pulsed for 30 h or 1 h with Brdu; the synchronized cells were pulsed for 1 h at final concentration of 45 µM. After fixing the cells, they were incubated overnight at 4°C and then incubated for 45 minutes at room temperature with HCl 3N. After a wash whit 0.1 M $Na_2B_4O_7$, the BrdU-positive cells were detected by using an anti-BrdU antibody (Abcam) incubated for 1h and a secondary rabbit antibody Alexa Fluor 488.

Cells were stained with propidium iodide 0.02 μg/μl and incubated with RNAse A 0.5 mg/μl and analyzed by double-flow cytometry analysis on a FACSCalibur (Becton Dickinson) instrument, to simultaneously determine the cell cycle profile (DNA content) by incorporation of propidium iodide (SIGMA), and the S phase cell population by incorporation of Brdu.

Acknowledgements

I would like to thank the Scuola Normale Superiore, the International Centre for Genetic Engineering and Biotechnology (ICGEB) Trieste and the Istituto di Fisiologia Clinica (CNR) Pisa for gave me the possibility to perform this project. Mainly, I want to thank all the people of these institutes that help me during this PhD program: Roberta, Vania, Arianna, Francesca, Barbara, Antonella, Laura, Tiziana, Silvia, Anita, Awatef, Alvaro, Cristina, Paola, Emanuele, Luca, Caterina, Claudia, Lorenzo, Pino, Monica, Letizia, Francesca, Milena, Antonella, Gianluca, Raffaella and Gulnara for the suggestions about the experiments to do. Many thanks go to Ramiro for being always helpful and for having always believed in the project. A special thank go to my "dining companions" for making these yaers more enjoyable, for helping me and to have became friends as well as colleagues.

I can not thank Prof. Arturo Falaschi because he didn't want, but I would like to say that was an honor to work with him.

The biggest thank go to my parents for always support me and trust me.

References

1. Jacob F, Brenner S. [On the regulation of DNA synthesis in bacteria: the hypothesis of the replicon.]. C R Hebd Seances Acad Sci 1963; 256:298-300.

2. Kaguni JM, Fuller RS, Kornberg A. Enzymatic replication of E. coli chromosomal origin is bidirectional. Nature 1982; 296:623-7.

3. Arai K, Kornberg A. Unique primed start of phage phi X174 DNA replication and mobility of the primosome in a direction opposite chain synthesis. Proc Natl Acad Sci U S A 1981; 78:69-73.

4. LeBowitz JH, McMacken R. The Escherichia coli dnaB replication protein is a DNA helicase. J Biol Chem 1986; 261:4738-48.

5. Lohman TM, Bjornson KP. Mechanisms of helicase-catalyzed DNA unwinding. Annu Rev Biochem 1996; 65:169-214.

6. Norio P, Kosiyatrakul S, Yang Q, Guan Z, Brown NM, Thomas S, Riblet R, Schildkraut CL. Progressive activation of DNA replication initiation in large domains of the immunoglobulin heavy chain locus during B cell development. Mol Cell 2005; 20:575-87.

7. Kong D, DePamphilis ML. Site-specific ORC binding, pre-replication complex assembly and DNA synthesis at Schizosaccharomyces pombe replication origins. EMBO J 2002; 21:5567-76.

8. Remus D, Beall EL, Botchan MR. DNA topology, not DNA sequence, is a critical determinant for Drosophila ORC-DNA binding. EMBO J 2004; 23:897-907.

9. Natale DA, Umek RM, Kowalski D. Ease of DNA unwinding is a conserved property of yeast replication origins. Nucleic Acids Res 1993; 21:555-60.

10. Nieduszynski CA, Knox Y, Donaldson AD. Genome-wide identification of replication origins in yeast by comparative genomics. Genes Dev 2006; 20:1874-9.

11. Bell SP, Dutta A. DNA replication in eukaryotic cells. Annu Rev Biochem 2002; 71:333-74.

12. Kreitz S, Ritzi M, Baack M, Knippers R. The human origin recognition complex protein 1 dissociates from chromatin during S phase in HeLa cells. J Biol Chem 2001; 276:6337-42.

13. DePamphilis ML. The 'ORC cycle': a novel pathway for regulating eukaryotic DNA replication. Gene 2003; 310:1-15.

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

15. Mendez J, Stillman B. 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; 20:8602-12.

16. Hofmann JF, Beach D. cdt1 is an essential target of the Cdc10/Sct1 transcription factor: requirement for DNA replication and inhibition of mitosis. EMBO J 1994; 13:425-34.

17. Takeda DY, Parvin JD, Dutta A. Degradation of Cdt1 during S phase is Skp2-independent and is required for efficient progression of mammalian cells through S phase. J Biol Chem 2005; 280:23416-23.

18. Yoshida K. Geminin organizes the molecular platform to balance cellular proliferation and differentiation. Front Biosci 2007; 12:2984-92.

19. Tada S, Li A, Maiorano D, Mechali M, Blow JJ. Repression of origin assembly in metaphase depends on inhibition of RLF-B/Cdt1 by geminin. Nat Cell Biol 2001; 3:107-13.

20. Madine M, Laskey R. Geminin bans replication licence. Nat Cell Biol 2001; 3:E49-50.

21. Lutzmann M, Maiorano D, Mechali M. A Cdt1-geminin complex licenses chromatin for DNA replication and prevents rereplication during S phase in Xenopus. EMBO J 2006; 25:5764-74.

22. De Marco V, Gillespie PJ, Li A, Karantzelis N, Christodoulou E, Klompmaker R, van Gerwen S, Fish A, Petoukhov MV, Iliou MS, Lygerou Z, Medema RH, Blow JJ, Svergun DI, Taraviras S, Perrakis A. Quaternary structure of the human Cdt1-Geminin complex regulates DNA replication licensing. Proc Natl Acad Sci U S A 2009; 106:19807-12.

23. Randell JC, Bowers JL, Rodriguez HK, Bell SP. Sequential ATP hydrolysis by Cdc6 and ORC directs loading of the Mcm2-7 helicase. Mol Cell 2006; 21:29-39.

24. Tye BK. MCM proteins in DNA replication. Annu Rev Biochem 1999; 68:649-86.

25. Blow JJ, Dutta A. Preventing re-replication of chromosomal DNA. Nat Rev Mol Cell Biol 2005; 6:476-86.

26. Zou L, Stillman B. Formation of a preinitiation complex by S-phase cyclin CDK-dependent loading of Cdc45p onto chromatin. Science 1998; 280:593-6.

27. Pacek M, Tutter AV, Kubota Y, Takisawa H, Walter JC. Localization of MCM2-7, Cdc45, and GINS to the site of DNA unwinding during eukaryotic DNA replication. Mol Cell 2006; 21:581-7.

28. Liu E, Li X, Yan F, Zhao Q, Wu X. Cyclin-dependent kinases phosphorylate human Cdt1 and induce its degradation. J Biol Chem 2004; 279:17283-8.

29. Fujita M. Cdt1 revisited: complex and tight regulation during the cell cycle and consequences of deregulation in mammalian cells. Cell Div 2006; 1:22.

30. Stillman B. Cell cycle control of DNA replication. Science 1996; 274:1659-64.

31. Zou L, Stillman B. Assembly of a complex containing Cdc45p, replication protein A, and Mcm2p at replication origins controlled by S-phase cyclin-dependent kinases and Cdc7p-Dbf4p kinase. Mol Cell Biol 2000; 20:3086-96.

32. Tanaka S, Diffley JF. Deregulated G1-cyclin expression induces genomic instability by preventing efficient pre-RC formation. Genes Dev 2002; 16:2639-49.

33. Djeliova V, Russev G, Anachkova B. Dynamics of association of origins of DNA replication with the nuclear matrix during the cell cycle. Nucleic Acids Res 2001; 29:3181-7.

34. Lengronne A, Schwob E. The yeast CDK inhibitor Sic1 prevents genomic instability by promoting replication origin licensing in late G(1). Mol Cell 2002; 9:1067-78.

35. DePamphilis ML. Eukaryotic DNA replication: anatomy of an origin. Annu Rev Biochem 1993; 62:29-63.

36. Kornberg A. Control of initiation of the Escherichia coli chromosome. Cold Spring Harb Symp Quant Biol 1991; 56:275-8.

37. Vashee S, Cvetic C, Lu W, Simancek P, Kelly TJ, Walter JC. Sequence-independent DNA binding and replication initiation by the human origin recognition complex. Genes Dev 2003; 17:1894-908.

38. Kanda T, Kamiya M, Maruo S, Iwakiri D, Takada K. Symmetrical localization of extrachromosomally replicating viral genomes on sister chromatids. J Cell Sci 2007; 120:1529-39.

39. Schepers A, Ritzi M, Bousset K, Kremmer E, Yates JL, Harwood J, Diffley JF, Hammerschmidt W. Human origin recognition complex binds to the region of the latent origin of DNA replication of Epstein-Barr virus. EMBO J 2001; 20:4588-602.

40. Harvey KJ, Newport, J. . CpG methylation of DNA restricts prereplication complex assembly in Xenopus egg extracts. Mol Cell Biol 2003; 23:6769-79.

41. Dijkwel PA, Wang S, Hamlin JL. Initiation sites are distributed at frequent intervals in the Chinese hamster dihydrofolate reductase origin of replication but are used with very different efficiencies. Mol Cell Biol 2002; 22:3053-65.

42. Altman AL, Fanning, E. . Defined sequence modules and architectural element cooperate to promote initiation at an ectopic mammalian chromosomal replication origin. Mol Cell Biol 2004; 24:38-50.

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

44. Wang L, Lin, C.M., Brooks, S., Cimbora, D., Groudine, M., Aladjem, M.I. . The human beta-globin replication initiation region consists of two modular independent replicators. Mol Cell Biol 2004; 24:3373-86.

45. Abdurashidova G, Deganuto M, Klima R, Riva S, Biamonti G, Giacca M, Falaschi A. Start sites of bidirectional DNA synthesis at the human lamin B2 origin. Science 2000; 287:2023-6.

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

47. Abdurashidova G, Danailov MB, Ochem A, Triolo G, Djeliova V, Radulescu S, Vindigni A, Riva S, Falaschi A. Localization of proteins bound to a replication origin of human DNA along the cell cycle. EMBO J 2003; 22:4294-303.

48. Paixao S, Colaluca IN, Cubells M, Peverali FA, Destro A, Giadrossi S, Giacca M, Falaschi A, Riva S, Biamonti G. Modular structure of the human lamin B2 replicator. Mol Cell Biol 2004; 24:2958-67.

49. Keller C, Ladenburger EM, Kremer M, Knippers R. The origin recognition complex marks a replication origin in the human TOP1 gene promoter. J Biol Chem 2002; 277:31430-40.

50. Cadoret JC, Meisch F, Hassan-Zadeh V, Luyten I, Guillet C, Duret L, Quesneville H, Prioleau MN. Genome-wide studies highlight indirect links between human replication origins and gene regulation. Proc Natl Acad Sci U S A 2008; 105:15837-42.

51. Courbet S, Gay S, Arnoult N, Wronka G, Anglana M, Brison O, Debatisse M. Replication fork movement sets chromatin loop size and origin choice in mammalian cells. Nature 2008; 455:557-60.

52. Berezney R, Dubey DD, Huberman JA. Heterogeneity of eukaryotic replicons, replicon clusters, and replication foci. Chromosoma 2000; 108:471-84.

53. Nakamura H, Morita T, Sato C. Structural organizations of replicon domains during DNA synthetic phase in the mammalian nucleus. Exp Cell Res 1986; 165:291-7.

54. Raghuraman MK, Winzeler EA, Collingwood D, Hunt S, Wodicka L, Conway A, Lockhart DJ, Davis RW, Brewer BJ, Fangman WL. Replication dynamics of the yeast genome. Science 2001; 294:115-21.

55. VanderVliet PC, ed. DNA replication in Eukaryotic Cells New York: Cold Spring Harbor Laboratory Press 1996.

56. Wei X, Samarabandu J, Devdhar RS, Siegel AJ, Acharya R, Berezney R. Segregation of transcription and replication sites into higher order domains. Science 1998; 281:1502-6.

57. Mechali M. DNA replication origins: from sequence specificity to epigenetics. Nat Rev Genet 2001; 2:640-5.

58. Danis E, Brodolin K, Menut S, Maiorano D, Girard-Reydet C, Mechali M. Specification of a DNA replication origin by a transcription complex. Nat Cell Biol 2004; 6:721-30.

59. Kohzaki H, Murakami Y. Transcription factors and DNA replication origin selection. Bioessays 2005; 27:1107-16.

60. Malott M, Leffak M. Activity of the c-myc replicator at an ectopic chromosomal location. Mol Cell Biol 1999; 19:5685-95.

61. Watanabe Y, Fujiyama A, Ichiba Y, Hattori M, Yada T, Sakaki Y, Ikemura T. Chromosome-wide assessment of replication timing for human chromosomes 11q and 21q: disease-related genes in timing-switch regions. Hum Mol Genet 2002; 11:13-21.

62. Gilbert DM. In search of the holy replicator. Nat Rev Mol Cell Biol 2004; 5:848-55.

63. McNairn AJ, Gilbert DM. Epigenomic replication: linking epigenetics to DNA replication. Bioessays 2003; 25:647-56.

64. Aladjem MI. Replication in context: dynamic regulation of DNA replication patterns in metazoans. Nat Rev Genet 2007; 8:588-600.

65. Schubeler D, Scalzo D, Kooperberg C, van Steensel B, Delrow J, Groudine M. Genome-wide DNA replication profile for Drosophila melanogaster: a link between transcription and replication timing. Nat Genet 2002; 32:438-42.

66. Cosgrove MS. Histone proteomics and the epigenetic regulation of nucleosome mobility. Expert Rev Proteomics 2007; 4:465-78.

67. Iizuka M, Stillman, B. . Histone acetyltransferase HBO1 interacts with the ORC1 subunit of the human initiator protein. J Biol Chem 1999; 274:23027-34.

68. Burke TW, Cook JG, Asano M, Nevins JR. Replication factors MCM2 and ORC1 interact with the histone acetyltransferase HBO1. J Biol Chem 2001; 276:15397-408.

69. Somoza JR, Skene, R.J., Katz, B.A., Mol, C., Ho, J.D., Jennings, A.J., Luong, C., Arvai, A., Buggy, J.J., Chi, E., Tang, J., Sang, B.C., Verner, E., Wynands, R., Leahy, E.M., Dougan, D.R., Snell, G., Navre, M., Knuth, M.W., Swanson, R.V., McRee, D.E., Tari, L.W. Structural snapshots of human HDAC8 provide insights into the class I histone deacetylases. . Structure (Camb) 2004; 12:1325-34.

70. Dokmanovic M, Marks, P.A. . Prospects: Histone Deacetylase Inhibitors. J Cell Biochem 2005; 96:293-304.

71. Kemp MG, Ghosh M, Liu G, Leffak M. The histone deacetylase inhibitor trichostatin A alters the pattern of DNA replication origin activity in human cells. Nucleic Acids Res 2005; 33:325-36.

72. Ghosh M, Kemp M, Liu G, Ritzi M, Schepers A, Leffak M. Differential binding of replication proteins across the human c-myc replicator. Mol Cell Biol 2006; 26:5270-83.

73. Aggarwal BD, Calvi BR. Chromatin regulates origin activity in Drosophila follicle cells. Nature 2004; 430:372-6.

74. Takei Y, Swietlik M, Tanoue A, Tsujimoto G, Kouzarides T, Laskey R. MCM3AP, a novel acetyltransferase that acetylates replication protein MCM3. EMBO Rep 2001; 2:119-23.

75. Collins N, Poot RA, Kukimoto I, Garcia-Jimenez C, Dellaire G, Varga-Weisz PD. An ACF1-ISWI chromatin-remodeling complex is required for DNA replication through heterochromatin. Nat Genet 2002; 32:627-32.

76. Zhou J, Chau, C.M., Deng, Z., Shiekhattar, R., Spindler, M.P., Schepers, A., Lieberman, P.M. . Cell cycle regulation of chromatin at an origin of DNA replication. EMBO J 2005; 24:1406-17.

77. Fyodorov DV, Blower MD, Karpen GH, Kadonaga JT. Acf1 confers unique activities to ACF/CHRAC and promotes the formation rather than disruption of chromatin in vivo. Genes Dev 2004; 18:170-83.

78. Becker PB, Horz W. ATP-dependent nucleosome remodeling. Annu Rev Biochem 2002; 71:247-73.

79. Varga-Weisz PD, Becker PB. Regulation of higher-order chromatin structures by nucleosome-remodelling factors. Curr Opin Genet Dev 2006; 16:151-6.

80. Abdurashidova G, Radulescu S, Sandoval O, Zahariev S, Danailov MB, Demidovich A, Santamaria L, Biamonti G, Riva S, Falaschi A. Functional interactions of DNA topoisomerases with a human replication origin. EMBO J 2007; 26:998-1009.

81. Pearson CE, Zorbas H, Price GB, Zannis-Hadjopoulos M. Inverted repeats, stem-loops, and cruciforms: significance for initiation of DNA replication. J Cell Biochem 1996; 63:1-22.

82. Zannis-Hadjopoulos M, Price GB. Eukaryotic DNA replication. J Cell Biochem 1999; Suppl 32-33:1-14.

83. Zannis-Hadjopoulos M, Yahyaoui W, Callejo M. 14-3-3 cruciform-binding proteins as regulators of eukaryotic DNA replication. Trends Biochem Sci 2008; 33:44-50.

84. Lewis EB. A gene complex controlling segmentation in Drosophila. Nature 1978; 276:565-70.

85. McGinnis W, Krumlauf R. Homeobox genes and axial patterning. Cell 1992; 68:283-302.

86. Kaufman TC, Seeger MA, Olsen G. Molecular and genetic organization of the antennapedia gene complex of Drosophila melanogaster. Adv Genet 1990; 27:309-62.

87. Bateson W. Material for the study of variation: treated with special regard to discontinuity in the origin of species. (Macmillan, New York) 1894.

88. Condie BG, Capecchi MR. Mice with targeted disruptions in the paralogous genes hoxa-3 and hoxd-3 reveal synergistic interactions. Nature 1994; 370:304-7.

89. Hafen E, Levine M, Gehring WJ. Regulation of Antennapedia transcript distribution by the bithorax complex in Drosophila. Nature 1984; 307:287-9.

90. Gehring WJ, Qian YQ, Billeter M, Furukubo-Tokunaga K, Schier AF, Resendez-Perez D, Affolter M, Otting G, Wuthrich K. Homeodomain-DNA recognition. Cell 1994; 78:211-23.

91. Chan SK, Jaffe L, Capovilla M, Botas J, Mann RS. The DNA binding specificity of Ultrabithorax is modulated by cooperative interactions with extradenticle, another homeoprotein. Cell 1994; 78:603-15.

92. Ekker SC, Jackson DG, von Kessler DP, Sun BI, Young KE, Beachy PA. The degree of variation in DNA sequence recognition among four Drosophila homeotic proteins. EMBO J 1994; 13:3551-60.

93. Joshi R, Passner JM, Rohs R, Jain R, Sosinsky A, Crickmore MA, Jacob V, Aggarwal AK, Honig B, Mann RS. Functional specificity of a Hox protein mediated by the recognition of minor groove structure. Cell 2007; 131:530-43.

94. Kuziora MA, McGinnis W. A homeodomain substitution changes the regulatory specificity of the deformed protein in Drosophila embryos. Cell 1989; 59:563-71.

95. Zeng W, Andrew DJ, Mathies LD, Horner MA, Scott MP. Ectopic expression and function of the Antp and Scr homeotic genes: the N terminus of the homeodomain is critical to functional specificity. Development 1993; 118:339-52.

96. Mann RS, Morata G. The developmental and molecular biology of genes that subdivide the body of Drosophila. Annu Rev Cell Dev Biol 2000; 16:243-71.

97. Rauskolb C, Wieschaus E. Coordinate regulation of downstream genes by extradenticle and the homeotic selector proteins. EMBO J 1994; 13:3561-9.

98. Rauskolb C, Peifer M, Wieschaus E. extradenticle, a regulator of homeotic gene activity, is a homolog of the homeobox-containing human proto-oncogene pbx1. Cell 1993; 74:1101-12.

99. Nourse J, Mellentin JD, Galili N, Wilkinson J, Stanbridge E, Smith SD, Cleary ML. Chromosomal translocation t(1;19) results in synthesis of a homeobox fusion mRNA that codes for a potential chimeric transcription factor. Cell 1990; 60:535-45.

100. Shen WF, Chang CP, Rozenfeld S, Sauvageau G, Humphries RK, Lu M, Lawrence HJ, Cleary ML, Largman C. Hox homeodomain proteins exhibit selective complex stabilities with Pbx and DNA. Nucleic Acids Res 1996; 24:898-906.

101. Borrow J, Shearman AM, Stanton VP, Jr., Becher R, Collins T, Williams AJ, Dube I, Katz F, Kwong YL, Morris C, Ohyashiki K, Toyama K, Rowley J, Housman DE. The t(7;11)(p15;p15) translocation in acute myeloid leukaemia fuses the genes for nucleoporin NUP98 and class I homeoprotein HOXA9. Nat Genet 1996; 12:159-67.

102. Chang CP, Shen WF, Rozenfeld S, Lawrence HJ, Largman C, Cleary ML. Pbx proteins display hexapeptide-dependent cooperative DNA binding with a subset of Hox proteins. Genes Dev 1995; 9:663-74.

103. Shen WF, Montgomery JC, Rozenfeld S, Moskow JJ, Lawrence HJ, Buchberg AM, Largman C. AbdB-like Hox proteins stabilize DNA binding by the Meis1 homeodomain proteins. Mol Cell Biol 1997; 17:6448-58.

104. Laurent A, Bihan R, Deschamps S, Guerrier D, Dupe V, Omilli F, Burel A, Pellerin I. Identification of a new type of PBX1 partner that contains zinc finger motifs and inhibits the binding of HOXA9-PBX1 to DNA. Mech Dev 2007; 124:364-76.

105. Godwin AR, Capecchi MR. Hoxc13 mutant mice lack external hair. Genes Dev 1998; 12:11-20.

106. LaRonde-LeBlanc NA, Wolberger C. Structure of HoxA9 and Pbx1 bound to DNA: Hox hexapeptide and DNA recognition anterior to posterior. Genes Dev 2003; 17:2060-72.

107. Galant R, Walsh CM, Carroll SB. Hox repression of a target gene: extradenticle-independent, additive action through multiple monomer binding sites. Development 2002; 129:3115-26.

108. Merabet S, Saadaoui M, Sambrani N, Hudry B, Pradel J, Affolter M, Graba Y. A unique Extradenticle recruitment mode in the Drosophila Hox protein Ultrabithorax. Proc Natl Acad Sci U S A 2007; 104:16946-51.

109. Berthelsen J, Zappavigna V, Ferretti E, Mavilio F, Blasi F. The novel homeoprotein Prep1 modulates Pbx-Hox protein cooperativity. EMBO J 1998; 17:1434-45.

110. Gebelein B, McKay DJ, Mann RS. Direct integration of Hox and segmentation gene inputs during Drosophila development. Nature 2004; 431:653-9.

111. Mann RS, ed. Hox genes, 2009.

112. Gauwerky CE, Croce CM. Chromosomal translocations in leukaemia. Semin Cancer Biol 1993; 4:333-40.

113. Ziemin-van der Poel S, McCabe NR, Gill HJ, Espinosa R, 3rd, Patel Y, Harden A, Rubinelli P, Smith SD, LeBeau MM, Rowley JD, et al. Identification of a gene, MLL, that spans the breakpoint in 11q23 translocations associated with human leukemias. Proc Natl Acad Sci U S A 1991; 88:10735-9.

114. Afonja O, Smith JE, Jr., Cheng DM, Goldenberg AS, Amorosi E, Shimamoto T, Nakamura S, Ohyashiki K, Ohyashiki J, Toyama K, Takeshita K. MEIS1 and HOXA7 genes in human acute myeloid leukemia. Leuk Res 2000; 24:849-55.

115. Kamps MP, Baltimore D. E2A-Pbx1, the t(1;19) translocation protein of human pre-B-cell acute lymphocytic leukemia, causes acute myeloid leukemia in mice. Mol Cell Biol 1993; 13:351-7.

116. Rawat VP, Cusan M, Deshpande A, Hiddemann W, Quintanilla-Martinez L, Humphries RK, Bohlander SK, Feuring-Buske M, Buske C. Ectopic expression of the homeobox gene Cdx2 is the transforming event in a mouse model of t(12;13)(p13;q12) acute myeloid leukemia. Proc Natl Acad Sci U S A 2004; 101:817-22.

117. Luo L, Kessel M. Geminin coordinates cell cycle and developmental control. Cell Cycle 2004; 3:711-4.

118. Falaschi A, Abdurashidova G, Biamonti G. DNA replication, development and cancer: a homeotic connection? Crit Rev Biochem Mol Biol; 45:14-22.

119. Gregoire D, Brodolin K, Mechali M. HoxB domain induction silences DNA replication origins in the locus and specifies a single origin at its boundary. EMBO Rep 2006; 7:812-6.

120. Aberdam D, Negreanu V, Sachs L, Blatt C. The oncogenic potential of an activated Hox-2.4 homeobox gene in mouse fibroblasts. Mol Cell Biol 1991; 11:554-7.

121. Dear TN, Sanchez-Garcia I, Rabbitts TH. The HOX11 gene encodes a DNA-binding nuclear transcription factor belonging to a distinct family of homeobox genes. Proc Natl Acad Sci U S A 1993; 90:4431-5.

122. Del Bene F, Wittbrodt J. Cell cycle control by homeobox genes in development and disease. Semin Cell Dev Biol 2005; 16:449-60.

123. Buske C, Feuring-Buske M, Antonchuk J, Rosten P, Hogge DE, Eaves CJ, Humphries RK. Overexpression of HOXA10 perturbs human lymphomyelopoiesis in vitro and in vivo. Blood 2001; 97:2286-92.

124. Lam DH, Aplan PD. NUP98 gene fusions in hematologic malignancies. Leukemia 2001; 15:1689-95.

125. Romana SP, Radford-Weiss I, Ben Abdelali R, Schluth C, Petit A, Dastugue N, Talmant P, Bilhou-Nabera C, Mugneret F, Lafage-Pochitaloff M, Mozziconacci MJ, Andrieu J, Lai JL, Terre C, Rack K, Cornillet-Lefebvre P, Luquet I, Nadal N, Nguyen-Khac F, Perot C, Van den Akker J, Fert-Ferrer S, Cabrol C, Charrin C, Tigaud I, Poirel H, Vekemans M, Bernard OA, Berger R. NUP98 rearrangements in hematopoietic malignancies: a study of the Groupe Francophone de Cytogenetique Hematologique. Leukemia 2006; 20:696-706.

126. Fujino T, Suzuki A, Ito Y, Ohyashiki K, Hatano Y, Miura I, Nakamura T. Single-translocation and double-chimeric transcripts: detection of NUP98-HOXA9 in myeloid leukemias with HOXA11 or HOXA13 breaks of the chromosomal translocation t(7;11)(p15;p15). Blood 2002; 99:1428-33.

127. Taketani T, Taki T, Shibuya N, Kikuchi A, Hanada R, Hayashi Y. Novel NUP98-HOXC11 fusion gene resulted from a chromosomal break within exon 1 of HOXC11 in acute myeloid leukemia with t(11;12)(p15;q13). Cancer Res 2002; 62:4571-4.

128. La Starza R, Trubia M, Crescenzi B, Matteucci C, Negrini M, Martelli MF, Pelicci PG, Mecucci C. Human homeobox gene HOXC13 is the partner of NUP98 in adult acute myeloid leukemia with t(11;12)(p15;q13). Genes Chromosomes Cancer 2003; 36:420-3.

129. Raza-Egilmez SZ, Jani-Sait SN, Grossi M, Higgins MJ, Shows TB, Aplan PD. NUP98-HOXD13 gene fusion in therapy-related acute myelogenous leukemia. Cancer Res 1998; 58:4269-73.

130. Pineault N, Abramovich C, Ohta H, Humphries RK. Differential and common leukemogenic potentials of multiple NUP98-Hox fusion proteins alone or with Meis1. Mol Cell Biol 2004; 24:1907-17.

131. Cantile M, Franco R, Tschan A, Baumhoer D, Zlobec I, Schiavo G, Forte I, Bihl M, Liguori G, Botti G, Tornillo L, Karamitopoulou-Diamantis E, Terracciano L, Cillo C. HOX D13 expression across 79 tumor tissue types. Int J Cancer 2009; 125:1532-41.

132. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD, Lander ES.

Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 1999; 286:531-7.

133. Lawrence HJ, Rozenfeld S, Cruz C, Matsukuma K, Kwong A, Komuves L, Buchberg AM, Largman C. Frequent co-expression of the HOXA9 and MEIS1 homeobox genes in human myeloid leukemias. Leukemia 1999; 13:1993-9.

134. Pineault N, Buske C, Feuring-Buske M, Abramovich C, Rosten P, Hogge DE, Aplan PD, Humphries RK. Induction of acute myeloid leukemia in mice by the human leukemia-specific fusion gene NUP98-HOXD13 in concert with Meis1. Blood 2003; 101:4529-38.

135. Ayton PM, Cleary ML. Molecular mechanisms of leukemogenesis mediated by MLL fusion proteins. Oncogene 2001; 20:5695-707.

136. Hisa T, Spence SE, Rachel RA, Fujita M, Nakamura T, Ward JM, Devor-Henneman DE, Saiki Y, Kutsuna H, Tessarollo L, Jenkins NA, Copeland NG. Hematopoietic, angiogenic and eye defects in Meis1 mutant animals. EMBO J 2004; 23:450-9.

137. Abate-Shen C. Deregulated homeobox gene expression in cancer: cause or consequence? Nat Rev Cancer 2002; 2:777-85.

138. Waltregny D, Alami Y, Clausse N, de Leval J, Castronovo V. Overexpression of the homeobox gene HOXC8 in human prostate cancer correlates with loss of tumor differentiation. Prostate 2002; 50:162-9.

139. Cillo C, Cantile M, Mortarini R, Barba P, Parmiani G, Anichini A. Differential patterns of HOX gene expression are associated with specific integrin and ICAM profiles in clonal populations isolated from a single human melanoma metastasis. Int J Cancer 1996; 66:692-7.

140. Krosl J, Baban S, Krosl G, Rozenfeld S, Largman C, Sauvageau G. Cellular proliferation and transformation induced by HOXB4 and HOXB3 proteins involves cooperation with PBX1. Oncogene 1998; 16:3403-12.

141. Lu Q, Knoepfler PS, Scheele J, Wright DD, Kamps MP. Both Pbx1 and E2A-Pbx1 bind the DNA motif ATCAATCAA cooperatively with the products of multiple murine Hox genes, some of which are themselves oncogenes. Mol Cell Biol 1995; 15:3786-95.

142. Morgan R, Pirard PM, Shears L, Sohal J, Pettengell R, Pandha HS. Antagonism of HOX/PBX dimer formation blocks the in vivo proliferation of melanoma. Cancer Res 2007; 67:5806-13.

143. Aulisa L, Forraz N, McGuckin C, Hartgerink JD. Inhibition of cancer cell proliferation by designed peptide amphiphiles. Acta Biomater 2009; 5:842-53.

144. Dutta A, Bell SP. Initiation of DNA replication in eukaryotic cells. Annu Rev Cell Dev Biol 1997; 13:293-332.

145. de Stanchina E, Gabellini D, Norio P, Giacca M, Peverali FA, Riva S, Falaschi A, Biamonti G. Selection of homeotic proteins for binding to a human DNA replication origin. J Mol Biol 2000; 299:667-80.

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

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

148. Stefanovic D, Stanojcic S, Vindigni A, Ochem A, Falaschi A. In vitro protein-DNA interactions at the human lamin B2 replication origin. J Biol Chem 2003; 278:42737-43.

149. Salsi V, Ferrari S, Ferraresi R, Cossarizza A, Grande A, Zappavigna V. HOXD13 binds DNA replication origins to promote origin licensing and is inhibited by geminin. Mol Cell Biol 2009; 29:5775-88.

150. Diviacco S NP, Zentilin L, Menzo S, Clementi M, Biamonti G, Riva S, Falaschi A, Giacca M. . A novel procedure for quantitative polymerase chain reaction by coamplification of competitive templates. Gene 1992; 122: 313-20.

151. Kang SH, Vieira K, Bungert J. Combining chromatin immunoprecipitation and DNA footprinting: a novel method to analyze protein-DNA interactions in vivo. Nucleic Acids Res 2002; 30:e44.

152. Ladenburger EM, Keller C, Knippers R. Identification of a binding region for human origin recognition complex proteins 1 and 2 that coincides with an origin of DNA replication. Mol Cell Biol 2002; 22:1036-48.

153. Tanaka T KD, Nasmyth K. Loading of an Mcm protein onto DNA replication origins is regulated by Cdc6p and CDKs. . Cell 1997; 90:649-60.

154. Wellik DM . in HOX genes (ed. O Pourquie) p-APE. Hox Genes and Vertebrate Axial Pattern. 2009.

155. Gabellini D, Colaluca IN, Vodermaier HC, Biamonti G, Giacca M, Falaschi A, Riva S, Peverali FA. Early mitotic degradation of the homeoprotein HOXC10 is potentially linked to cell cycle progression. EMBO J 2003; 22:3715-24.

156. Karnani N, Taylor CM, Malhotra A, Dutta A. Genomic study of replication initiation in human chromosomes reveals the influence of transcription regulation and chromatin structure on origin selection. Mol Biol Cell; 21:393-404.

157. Kusic J, Kojic S, Divac A, Stefanovic D. Noncanonical DNA elements in the lamin B2 origin of DNA replication. J Biol Chem 2005; 280:9848-54.

158. Mitkova AV, Biswas-Fiss EE, Biswas SB. Modulation of DNA synthesis in Saccharomyces cerevisiae nuclear extract by DNA polymerases and the origin recognition complex. J Biol Chem 2005; 280:6285-92.

159. Takahashi Y, Rayman JB, Dynlacht BD. Analysis of promoter binding by the E2F and pRB families in vivo: distinct E2F proteins mediate activation and repression. Genes Dev 2000; 14:804-16.