Univerzita Karlova v Praze

Přírodovědecká fakulta

Studijní program: Speciální chemicko-biologické obory

Studijní obor: Molekulární biologie a biochemie organismů

Bc. Martin Petr

Interakce promotoru a terminátoru v eukaryotické transkripci pomocí RNA polymerázy II

Promoter-terminator interactions in eukaryotic RNA polymerase II transcription

Bakalářská práce

Školitel: prof. RNDr. Pavel Hozák, DrSc.

Ústav molekulární genetiky, AV ČR

Praha, 2012

Prohlášení

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně a že jsem uvedl všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze, 27. 8. 2012

Martin Petr

Poděkování

Velice děkuji svému školiteli, prof. RNDr. Pavlu Hozákovi, DrSc., za příležitost pracovat v jeho laboratoři a také za důvěru, kterou tím ve mě vložil.

Můj velký dík patří také Tomášovi za množství připomínek k obsahu této práce. Především mu ale budu vždy zavázán za trpělivost, se kterou se mi během mého premiérového působení v laboratoři věnoval. Nekončící záplavu mých dotazů a nešikovných přešlapů přijímal vždy s obdivuhodným klidem hodným zenového mistra.

Děkuji svým rodičům za to, že mé nenadálé otočení kormidla od informatiky směrem k biologii přijali s pochopením a že mě v mém dalším studiu i nadále podporovali a povzbuzovali.

Abstract

Gene loops are chromatin structures formed by juxtaposition of distal genomic regions. Since these regions are often involved in transcription cycle control, gene loops therefore provide another mechanism of regulation of gene expression. This thesis summarizes recent findings about gene loops, focusing specifically on loops formed by interactions between promoter and terminator regions of genes transcribed by the eukaryotic RNA polymerase II. Different cases of gene loops discovered in several yeast genes, the mammalian *BRCA1* tumor suppressor and the HIV-1 integrated provirus are described, including mechanisms that possibly lead to the formation of these structures. Since gene loops and interactions between promoter and terminator in yeast have been linked to the transcriptional memory, their involvement in this phenomenon is discussed. Finally, as *BRCA1* and HIV-1 are directly linked to serious human diseases, the potential significance of alterations of gene loops in the development of various pathological conditions is presented.

Keywords: gene loops, chromatin loops, chromatin conformation, transcriptional memory, transcription, gene expression, regulation

Abstrakt

Genové smyčky jsou chromatinové struktury, k jejichž vzniku dochází při fyzické interakci vzdálených oblastí DNA. Jelikož se tyto oblasti často podílejí na kontrole transkripčního cyklu, genové smyčky tvoří další úroveň regulace genové exprese. Tato práce shrnuje aktuální znalosti o genových smyčkách, přičemž se zaměřuje speciálně na smyčky vzniklé interakcemi promotorových a terminátorových oblastí genů transkribovaných pomocí eukaryotické polymerázy II. Jsou rozebrány smyčky objevené u kvasinkových genů, u savčího tumor supresoru *BRCA1* a také u integrovaného genomu viru HIV-1, a to včetně mechanismů, jež pravděpodobně vedou k jejich vzniku. Jelikož byla nalezena spojitost mezi genovými smyčkami a transkripční pamětí u kvasinek, je popsán jejich fyziologický význam při tomto jevu. V souvislosti s přímým spojením *BRCA1* a HIV-1 se závažnými lidskými chorobami je dále navržen obecný význam poruch ve funkci genových smyček při vzniku závažných patologických stavů.

Klíčová slova: genové smyčky, chromatinové smyčky, konformace chromatinu, transkripční paměť, transkripce, genová exprese, regulace

Table of contents

List of abbreviations

3C	capturing chromosome conformation
kb	kilobase
mRNA	messenger RNA
poly(A)	polyadenylation
pre-mRNA	precursor mRNA
pta1-td	degron-tagged Pta1 strain
rpb1-1	temperature sensitive RPB1 mutant strain
ssu72-td	degron-tagged Ssu72 strain
sua7-1	TFIIB mutant strain
BRCA1	breast cancer susceptibility gene 1
ChIP	chromatin immunoprecipitation
CNS	conserved noncoding sequence
CTD	C-terminal domain
DNA	deoxyribonucleic acid
DPE	downstream promoter element
GTF	general transcription factor
HeLa	immortal cell line
HIV-1	human immunodeficiency virus type 1
Inr	initiator core promoter
LTR	long terminal repeat
Mlp1	myosin-like protein 1
MCF7	breast cancer cell line
MSD	major splice donor site
NPC	nuclear pore complex
ORF	open reading frame
Pta1	cleavage/polyadenylation factor
PCR	polymerase chain reaction
PIC	preinitiation complex
RNA	ribonucleic acid
RNAP	RNA polymerase
RPB1	RNAP II subunit containing CTD
RT PCR	real-time polymerase chain reaction
Ssu72	transcription/RNA-processing factor
SPA	synthetic polyadenylation site
TBP	TATA-binding protein
TFIIB	general transcription factor IIB
TFIID	general transcription factor IID
TFIIE	general transcription factor IIE
TFIIF	general transcription factor IIF
TFIIH	general transcription factor IIH

1. Introduction

The research of gene expression activation has for a long time focused primarily on studies of transcription initiation mechanisms and promoter selection. The commonly accepted model of transcription of protein-coding genes in eukaryotes is quite straightforward. It involves the initial recruitment of RNA polymerase II (RNAP II) to the transcription initiation site by a set of DNA-binding and protein-binding general transcription factors assembled stepwise to form a preinitiation complex (PIC). After elongation of the nascent pre-mRNA from a linear DNA template, the transcription terminates after encountering a terminator sequence, with RNAP II dissociating from the template. An important feature of this model is that in order to reinitiate the transcription, *de novo* assembly of the PIC and subsequent recruitment of RNAP II is required. However, there has been a growing amount of evidence during the last decade challenging this straightforward model of RNAP II transcription cycle and suggesting that the whole process is more intricate and complex.

In vitro studies have indicated that a subset of transcription initiation factors remains at the promoter region after the RNAP II leaves the promoter. These factors form a so-called "scaffold" complex which then facilitates a rapid reassembly of the complete initiation complex upon reactivation of transcription (Yudkovsky et al., 2000).

Moreover, many studies have suggested the involvement of higher order chromatin structures called chromatin loops or gene loops in eukaryotic transcription. These topological arrangements were discovered to bring distal regulatory regions into proximity, causing activation or repression of genes. Gene loops have been observed in various species and cell types and have been shown to work over wide range of distances. For example, long-range parental-specific gene loops formed by interactions between imprinted *H19* and *Igf2* genes influence the expression of these genes over a distance of ~100 kb (Murell et al., 2004). In the case of mouse β-globin gene, locus control regions were observed in proximity with the β-globin promoter over a distance of 40-60 kb in expressing tissue, whereas this interaction was not observed in non-expressing tissue (Tolhuis et al., 2002). These structures therefore provide an explanation for the long known, but not understood mechanism of long-range actions of various distal regulatory DNA elements, such as enhancers or silencers.

This thesis will focus specifically on the role of physical interactions between promoter and terminator regions of eukaryotic genes transcribed by RNAP II. After describing the basic model of transcription cycle, main methods primarily used for analyzing distal genomic interactions and gene looping will be covered. The recent knowledge about gene loops and their function in transcriptional memory in yeast *Saccharomyces cerevisiae* and evidence about the juxtaposition of promoter and terminator regions in mammals will then be presented. In conclusion, the future perspectives and the importance of gene looping in pathological conditions will be discussed.

2. Eukaryotic transcription by RNAP II

Transcription is the process of RNA synthesis from the DNA template. It is the first step of gene expression, during which the information stored in the molecule of DNA is converted into the final gene product, protein or RNA. Enzyme that catalyzes the transcription reaction is called RNA polymerase. RNA polymerases are multisubunit enzymes that vary in size, composition and function. In bacteria and archaea, all genes are transcribed by a single RNA polymerase. In eukaryotes, transcription is catalyzed by three RNA polymerases (RNAP I, RNAP II and RNAP III) that are all similar in their structure and features, but each transcribes a different set of genes. While RNAP I and III transcribe genes encoding ribosomal RNAs, transfer RNAs and small nuclear RNAs, the RNAP II is responsible for transcription of all protein-coding eukaryotic genes.

The transcription of protein-coding genes by RNAP II consists of three successive phases – initiation, elongation and termination. The initiation starts with the RNAP II binding to a specific DNA region called the promoter, which is a sequence critical for correct positioning of the enzyme at the transcription start site. However, the recruitment of RNAP II to the transcription start site depends on additional proteins called general transcription factors (GTFs). These GTFs assemble in a stepwise fashion by DNA-protein and protein-protein interactions into the preinitiation complex (PIC), which is then escorting the RNAP II to the promoter site. Five general transcription factors, TFIIB, TFIID, TFIIE, TFIIF and TFIIH, are required *in vivo* to initiate transcription from most eukaryotic protein-coding genes.

PIC formation is initiated by the TATA-binding protein (TBP), a subunit of TFIID that binds to the TATA box promoter element. Although TATA box is a widespread element among yeast promoters, most eukaryotic promoters lack this sequence. In these cases, other TFIID subunits initiate the PIC assembly by binding to other promoter elements, including the initiator (Inr) or the downstream promoter element (DPE). However, TBP is always involved in the process of PIC assembly, regardless of the sequence specificity. After TFIIB associates with the TBP, the DNA-TBP-TFIIB complex is able to recruit RNAP II, which is escorted to the promoter by TFIIF. The assembly of PIC is then completed by binding of TFIIE and TFIIH (for review see Buratowski, 1994).

The progression of the RNAP II through the transcription cycle is dependent on covalent modifications of a unique structure in its C-terminal domain (CTD). This structure consists of multiple repetitions of a conserved heptapeptide sequence: Tyr1-Ser2-Pro3-Thr4-Ser5-Pro6-Ser7. Changes in the phosphorylation status of Ser2 and Ser5 cause transitions between the initiation-competent and elongation-competent form of the RNAP II. Specifically, it has been shown that RNAP II is recruited to the promoter when its CTD is unphosphorylated. Phosphorylation of this terminal domain is required for an event called "promoter clearance", in which the RNAP II leaves the promoter and proceeds to the elongation phase (for review see Zhang et al., 2012).

TFIIH general transcription factor and CTD kinase 1 complex phosphorylate the CTD (Cho et al., 2001). In turn, Ssu72, a part of the 3'-end processing complex, was shown to dephosphorylate the CTD and thus participate in the conversion of the RNAP II to the initiation-competent dephosphorylated form (Krishnamurthy et al., 2004).

The functional significance of phosphorylation and dephosphorylation of CTD is not limited only to transitions between initiation and elongation phases. Prior to its release from the RNAP II, the nascent RNA molecule has to undergo a number of processing changes – capping, cleavage and polyadenylation. The CTD serves as a "platform" for attachment of various pre-mRNA processing enzymes, dependending on its phosphorylation status (for review see Zhang et al., 2012). Shortly after the elongation starts, a 7-methylguanylate cap is added to the 5'-end of the nascent RNA by a capping enzyme temporarily bound to the CTD. Cleavage and polyadenylation are initiated by encountering a conserved poly(A) signal region towards the 3'-end of the nascent pre-mRNA. This sequence is recognized by a CTD-bound 3'-end processing

complex which cleaves the pre-mRNA and synthesizes up to 250 adenosine residues at its 3'-end, thus creating a poly(A)-tail.

After encountering the poly(A) signal on the template strand, the elongation still continues for more than 1000 bases downstream from this site. After cleavage and polyadenylation of the nascent pre-mRNA are finished, the transcription is terminated and the RNAP II leaves the DNA template.

3. Techniques for studying gene loops

A number of methods have been developed or adapted to detect and analyze higher order chromatin structures. Two of the most important techniques required for understanding and interpretation of results presented in this thesis are briefly described below.

Capturing chromosome conformation

Capturing chromosome conformation (3C) is the most widely used method for studying physical interactions between multiple genomic loci. It allows the detection and quantification of the frequency of interaction between distant parts of DNA by producing specific ligation products (Dekker et al., 2002). Although originally it has been developed to analyze long range genomic interactions spanning hundreds of kilobases, it was later adapted even to study contacts between regions as close as 1kb (Ansari and Hampsey, 2005; Singh and Hampsey, 2007).

Its principle relies on the formaldehyde cross-linking of protein interactions which mediate contacts between distal genomic loci. After chromatin extraction, partial purification and digestion by restriction enzyme, cross-linked DNA fragments are isolated. The digested DNA is then religated in diluted solution. If the regions of interest are indeed physically interacting, they are ligated to each other. This occurs with greater frequency in contrast to random intermolecular ligation because intramolecular ligation reaction is a more favorable in diluted solution.

The physical interaction of two distal DNA fragments is proved by PCR amplification. As seen in the figure 1, a segment that forms a gene loop is spanned by a pair of two divergent primers (denoted A and B in figure 1). If the analyzed segment indeed adopts a looping conformation, the intramolecular ligation of the cross-linked DNA fragments causes the primers

to become convergent and after a PCR amplification a specific A-B PCR product can be observed.

Indirect ChIP

Indirect ChIP (schematically presented in figure 1) is based on chromatin immunoprecipitation. ChIP in general is used for locating DNA-binding sites of a particular protein. After chemical cross-linking of the DNA-binding proteins to the DNA itself and subsequent isolation and fragmentation of chromatin, the protein-bound DNA fragments are immunoprecipitated and identified by PCR. However, in contrast to ChIP, in indirect ChIP the indirectly associated DNA fragments and not the directly bound DNA fragments are analyzed by PCR.

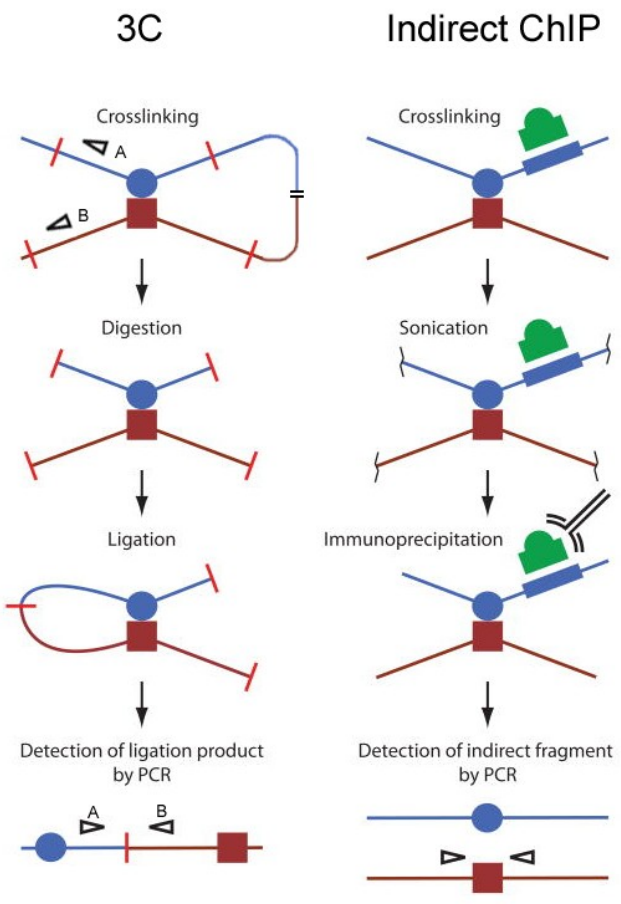

Figure 1. Capturing chromosome conformation and indirect ChIP. (Taken and adapted from Kadauke and Blobel, 2009.)

4. Promoter-terminator interactions in yeast

One of the first insights leading to the discovery of distal genomic interactions between promoter and terminator regions was the discovery of a physical interaction between the 3'-end pre-mRNA processing factor Ssu72 and a general transcription factor TFIIB (Sun and Hampsey, 1996; Dichtl et al., 2002; Krishnamurthy et al., 2004). Moreover, Ssu72 protein was found to be associated not only to the terminator regions of yeast genes but to the promoter regions as well (Nedea et al., 2003). These results suggested that complexes and sequences participating in distal regions during transcription may be brought into physical proximity resulting in a formation of a looped chromatin conformation.

Dependence of gene looping on active transcription

First thorough and detailed evidence supporting the hypothesis of gene loops and juxtaposition of promoter and terminator regions, initially presented by O'Sullivan, 2004, was provided by Ansari and Hampsey, 2005 and Singh and Hampsey, 2007. By using a mutant strain *rpb1-1* expressing a temperature sensitive form of RNAP II that functions normally at permissive temperature but stops transcribing at 37°C (Nonet et al., 1987), chromatin conformations of various genes were analyzed. After cultivation at permissive temperature, 3C analysis of isolated chromatin revealed a clear looping PCR profile in *rbp1-1* and wild-type strains, indicating the juxtaposition of promoter and terminator regions. However, when shifted to a non-permissive temperature when no transcription occurs, the signal disappeared from *rpb1-1* strain while it was still present in the wild-type (Ansari and Hampsey, 2005; Singh and Hampsey, 2007). Similar results were obtained by testing a gene-specific repression of both *BUD3* and *SEN1* transcription by putting them under control of *GAL1* promoter. In this case the looping occurred only when the cells were grown in the presence of the transcription-inducing galactose and was not detected when exposing cells to glucose (Ansari and Hampsey, 2005). This suggested that gene loops not only do exist but their formation correlates with active transcription. Moreover, the precise site of chromatin juxtaposition was mapped specifically to promoter and terminator regions only. No other part of each of the studied ORFs seemed to be participating in gene loops (Ansari and Hampsey, 2005; Singh and Hampsey, 2007).

Factors required for the formation of gene loops

The first factor investigated for its possible role in gene looping was Ssu72, a phosphatase known to be an important part of the 3'-end processing complex and also an interaction partner of the THIIB general transcription factor (Sun and Hampsey, 1996; Dichtl et al., 2002; Krishnamurthy et al., 2004). Another factor investigated for its possible influence in gene looping was Pta1, also a component of the 3'-end processing machinery and a known interaction partner of the Ssu72 protein (Nedea et al., 2003; Krishnamurthy et al., 2004).

To examine the roles of both proteins in loop formation, two strains expressing degron-tagged forms of Ssu72 and Pta1 were used. In these strains under permissive conditions, protein levels of Ssu72 and Pta1 are as in the wild-type. However, the shift to 37°C causes a rapid depletion of the tagged proteins by a proteolysis-mediated degradation (Dohmen et al., 1994). 3C analysis of chromatin extracted from *ssu72-td* and *pta1-td* cells after 1 hour of incubation in 37°C revealed a strong reduction of PCR signal which was in contrast to wild-type strains, where this signal remained unchanged. Importantly, this change was not caused by an indirect effect of globally impaired transcription as RNA levels remained unchanged (Ansari and Hampsey, 2005). These results show that 3'-end processing complex is necessary for establishing contacts between promoter and terminator regions of yeast genes.

The same negative impact on the formation of gene loops was observed in *sua7-1* yeast strain that expresses a mutated homolog of the general transcription factor TFIIB. Interestingly, although 3C revealed that the formation of loops was clearly impaired in *sua7-1*, no negative effect on transcription levels has been observed, suggesting that the role of TFIIB in this process is transcription independent (Singh and Hampsey, 2007). In the wild-type strain, ChIP revealed the association of TFIIB to terminator region of *PMA1* and *BLM10*, in addition to its binding to the promoter. However, in *sua7-1* mutant, the association of TFIIB homolog to the terminator but not to the promoter was markedly diminished, further confirming the absence of gene loops (Singh and Hampsey, 2007).

Importantly, because the TATA-binding protein (TBP) did not bind to the 3'-end region, the association of TFIIB described above could not be caused by a presence of a cryptic promoter

(Singh and Hampsey, 2007). TFIIB therefore binds to the terminator region independently of the TBP and transcription initiation.

As Ssu72 is a known binding partner of TFIIB, its role in TFIIB association to terminator regions was analyzed using ChIP assay of chromatin extracted from the *ssu72-td* mutant. After the depletion of the Ssu72 protein, the association of TFIIB to the terminator, but not to the promoter region, was diminished (Singh and Hampsey, 2007). It is therefore clear that TFIIB association to the terminator depends on the 3'-end processing complex which Ssu72 is a part of. Since 3'-end processing complex is recruited to the 3'-end of the nascent pre-mRNA, an initial round of transcription seems to be necessary for the juxtaposition of promoter and terminator and formation of gene loop.

A model of gene loops in yeast

Taken together, these results provide a clear evidence for the existence of gene loops in yeast. After analyzing a wide range of different genes it can be assumed that these higher order chromatin structures may be a common phenomenon of the yeast RNAP II transcription.

The current model of gene loop formation is depicted in figure 2. Upon transcription initiation, RNAP II, TFIIB and TFIIF leave the promoter region, whereas the rest of the pre-initiation complex (PIC) remains associated with the promoter in the form of a "scaffold" complex (Yudkovsky et al., 2000). Ssu72, discovered to be a crucial factor in the formation of gene loops, plays possibly two roles in this process. Firstly, it has been shown to serve as a RNAP II CTD Ser5 phosphatase (Krishnamurthy et al., 2004) and is therefore partly responsible for the conversion of an elongation-competent phosphorylated form of the RNAP II to its initiation-competent dephosphorylated form. Secondly, since it has been discovered to physically interact with TFIIB (Sun and Hampsey, 1996; Dichtl et al., 2002) and is also required for TFIIB localization to terminator regions during gene looping (Singh and Hampsey, 2007), it could mediate TFIIB association with RNAP II at the promoter-terminator juxtaposition. However, this possibility yet remains to be confirmed. Since the "scaffold" contains all components necessary for the initiation of transcription except RNAP II, TFIIB and TFIIF, this model provides an elegant mechanism for possible rapid transcription re-initiation.

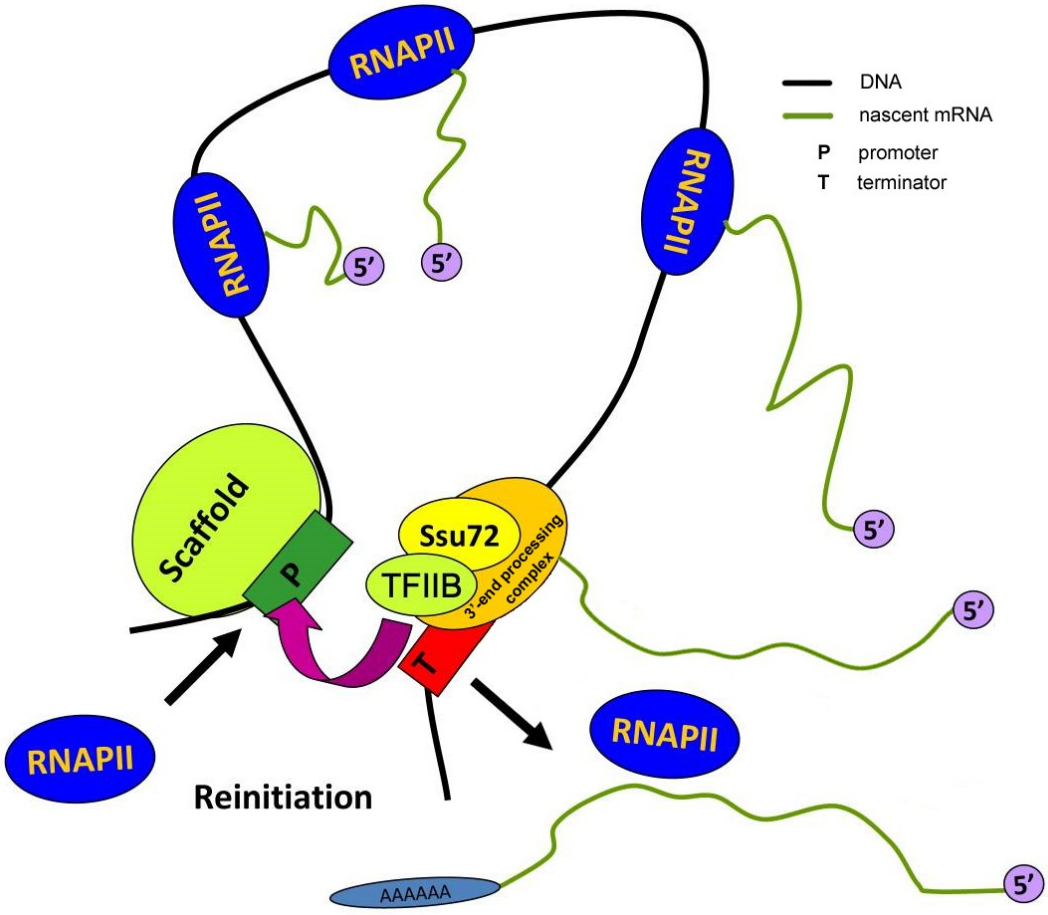

Figure 2. A model of transcription dependent gene looping. (Taken and adapted from Ansari and Hampsey, 2005.)

5. Transcriptional memory in yeast

The existence of gene loops and promoter-terminator juxtaposition has been thoroughly analyzed and confirmed. Nevertheless, the functional significance of looping in transcription has remained unknown, notably because the disruption of looping did not seem to have any influence on transcription levels, as noted above (Ansari and Hampsey, 2005; Ansari and Hampsey, 2007). However, the discovery of the promoter and terminator juxtaposition with its possible facilitation of transcription re-initiation provided an elegant solution to the unresolved phenomenon of transcriptional memory in yeast.

The term “transcriptional memory” has been coined to describe the ability of yeast genes to “store” information about their previous transcriptional activity. This effect enables cells to rapidly reinitiate transcription following a certain period of repression. Yeast *GAL* genes are induced in the presence of galactose as the only source of carbon and repressed when exposed to glucose. It has been observed that although the *GAL* induction is relatively slow and takes up to 2 hours until it reaches a peak of its transcription activity, its full re-induction after a period of repression is just a matter of minutes (Brickner et al., 2007; Kundu et al., 2007; Zacharioudakis et al., 2007). Several factors have been discovered to contribute to the transcriptional memory, including a H2A.Z histone variant (Brickner et al., 2007), Gal1 catabolic enzyme (Zacharioudakis et al., 2007) or SWI/SNF chromatin remodeling complex (Kundu et al., 2007). However, the exact molecular mechanism by which cells maintain a memory of the recent transcriptional activity remained unknown.

The role of looping in transcriptional memory

The role of gene loops in transcriptional memory was determined by measuring activation kinetics of the *GAL10* gene (Lainé et al., 2009), which has been confirmed to form gene loops upon its transcriptional activation (Singh and Hampsey, 2007).

Remarkably, the time-scale 3C analysis of *GAL10* during the phase of glucose repression revealed loops persisting for more than 6 hours after the moment the repression of transcription had started (Lainé et al., 2009).

Activation kinetics of *GAL10* was measured in *sua7-1* strain expressing a mutated homolog of TFIIB known to impair gene looping but maintaining a normal transcription phenotype (Singh and Hampsey, 2007). RT PCR did not reveal any significant difference in mRNA accumulation kinetics between the wild-type and *sua7-1* following an initial galactose induction. Both required more than 1 h for full activation. However, 3C assay of cells exposed to the same conditions revealed gene loops in the wild-type only (Lainé et al., 2009). This suggested that the initial round of transcription of *GAL10* does not depend on the presence of gene loops. Importantly, after galactose re-induction following a period of glucose repression, the maximal transcription activity in the wild-type was reached in about 2 min. This was in contrast with *sua7-*

1 strain where the time required for maximal reactivation was similar to the situation after initial galactose induction (Lainé et al., 2009).

Taken together these results proved that *GAL10* transcription displays rapid reactivation kinetics in strains with the ability to form gene loops. In contrast, cells unable to form gene loops have lost the ability of rapid transcription re-induction.

The role of the nuclear pore complex in transcriptional memory

Another factor discovered to play an important role in transcriptional memory is an interaction of actively transcribed genes with nuclear pore complex (NPC). There have been several reports of genes localizing to the nuclear periphery upon their transcription activation, for example *GAL* regulon genes, *INO1*, *HXK1* and α-factor induced genes (Luthra et al. 2007; Brickner and Walter, 2004; Taddei et al., 2006; Cassolari et al. 2004). Moreover, translocation of *GAL1* and *INO1* to the nuclear periphery has been reported to remain unchanged even during the state of repression and this peripheral retention has been implied to enhance rates of gene expression by facilitating its rapid re-induction (Brickner et al., 2007).

The interaction between transcriptionally active genes and the NPC is mediated by myosin-like protein 1 (Mlp1), a perinuclear protein located in the filamentous mesh of the NPC basket (Casolari et al., 2005; Luthra et al., 2007). Apart from this function, Mlp1 was shown to be involved in mRNA quality control by nuclear retention of incorrectly spliced mRNAs (Galy et al., 2004; Vinciguera et al., 2005).

Myosin-like protein 1 in a looping-mediated memory effect

To investigate the possible role of Mlp1 in transcription and looping mediated transcriptional memory, a series of experiments was performed in the wild-type and mlp1Δ strains. The target of these experiments was *HXK1*, a gene known to translocate to NPC upon transcriptional activation (Taddei et al., 2006).

Consistently with a number of yeast genes analyzed previously, the 3C analysis of *HXK1* in the wild-type showed looping structures during its transcription activity and even 1 h after glucose repression (O'Sullivan 2004; Ansari and Hampsey, 2005; Singh and Hampsey, 2007; Tan-Wong et al., 2009). In contrast to wild-type, 3C of mlp1Δ strain revealed a loss of promoter-terminator juxtaposition as early as 5 min after an exposures to glucose (Tan-Wong et al., 2009).

ChIP assay of Mlp1 revealed its transcription dependent association with 5'-end and 3'-end region of *HXK1* gene in wild-type cells which remained present for up to 1 h after glucose repression. This binding correlated with the presence of gene loops (Tan-Wong et al., 2009).

These results support the idea of the involvement of Mlp1 in the maintenance of promoter-terminator juxtaposition during transcriptional repression. To examine the direct involvement of Mlp1 in transcriptional memory, RT PCR time-course analysis of mRNA accumulation kinetics was performed. While wild-type cells displayed an expected rapid accumulation of *HXK1* mRNA during re-induction after 1 h of glucose repression, this memory effect was completely absent in the mlp1Δ. However, after only 10 min of repression, when gene loops still persist even in the mlp1Δ strain, mRNA accumulation kinetics in wild-type and mlp1Δ strains were identical (Tan-Wong et al., 2009).

In conclusion these data indicate that Mlp1 is required for the maintenance of transcriptional memory and stability of promoter-terminator interactions in *HXK1* gene during repression, but it does not participate in the initial formation of gene loops. This was further confirmed by repeating the Mlp1 ChIP analysis of *HXK1* gene in the *sua7-1* strain. While Mlp1 ChIP of the wild-type yeast revealed a 5'-end and 3'-end binding of *HXK1* gene during active transcription and after 1 hour of glucose repression, this binding was completely lost in the mutant *sua7-1* strain (Tan-Wong et al., 2009). This proves that Mlp1 attachment to 5'-end and 3'-end regions of *HXK1* occurs after the juxtaposition is formed and that it probably attaches to the base of the gene loop where the promoter-terminator interaction takes place.

Even though the physiological role of gene loops in transcriptional memory in yeast has been thoroughly investigated and confirmed, the ability of a gene to form this topological structure does not necessarily implicate its ability to display a memory effect. *INO1* translocates to the nuclear periphery upon transcription activation and forms gene loops by promoter-terminator juxtaposition (Brickner et al., 2007). However, in contrast to *HXK1*, it does not maintain this chromatin structure after glucose repression at all and thus is not able "memorize" a previous transcriptional activity (Tan-Wong et al., 2009). ChIP assay revealed that Mlp1 is still bound to the promoter and poly(A) regions of *INO1* gene, but this Mlp1 attachment is lost during transcription repression (Tan-Wong et al., 2009). This suggests that although gene loops are

necessary for the effect of transcriptional memory, their primary function may be in tethering the active genes to the site of mRNA control and nuclear export to cytoplasm.

A model of transcriptional memory in yeast

The present understanding of the role of gene loops in yeast transcriptional memory is summarized in figure 3. Following transcription initiation and an initial round of transcription, the promoter and terminator regions of a gene are juxtaposed (as presented in chapter 4). This topological structure is then translocated to the NPC by a yet unknown mechanism and tethered to the NPC by the myosin-like protein 1, thus anchoring the actively transcribing unit close to the site of pre-mRNA processing, quality control and export. Mlp1 is responsible for stabilization of the promoter-terminator juxtaposition and maintaining its position near the NPC during short-term repression. This stabilization allows a rapid mRNA accumulation after re-induction of transcription. However, after long periods of repression this anchoring breaks, the gene loop leaves the nuclear periphery and the DNA adopts the linear conformation again (Tan-Wong et al., 2009).

Gene looping and subsequent tethering to the NPC as just described is not specific only to genes displaying memory effect, since they occur even in case of genes that lack the ability of rapid re-induction after a period of transcription repression. Transcriptional memory enables a rapid reinduction and accumulation of mRNA several orders of magnitude faster in comparison to initial induction, fulfilling immediate metabolic needs of the yeast cell. However, in case of the majority of genes, this prolonged retention of gene loop at the NPC would not provide any significant advantage for the cell and their promoter-terminator interactions therefore end almost immediately after the transcription ceases.

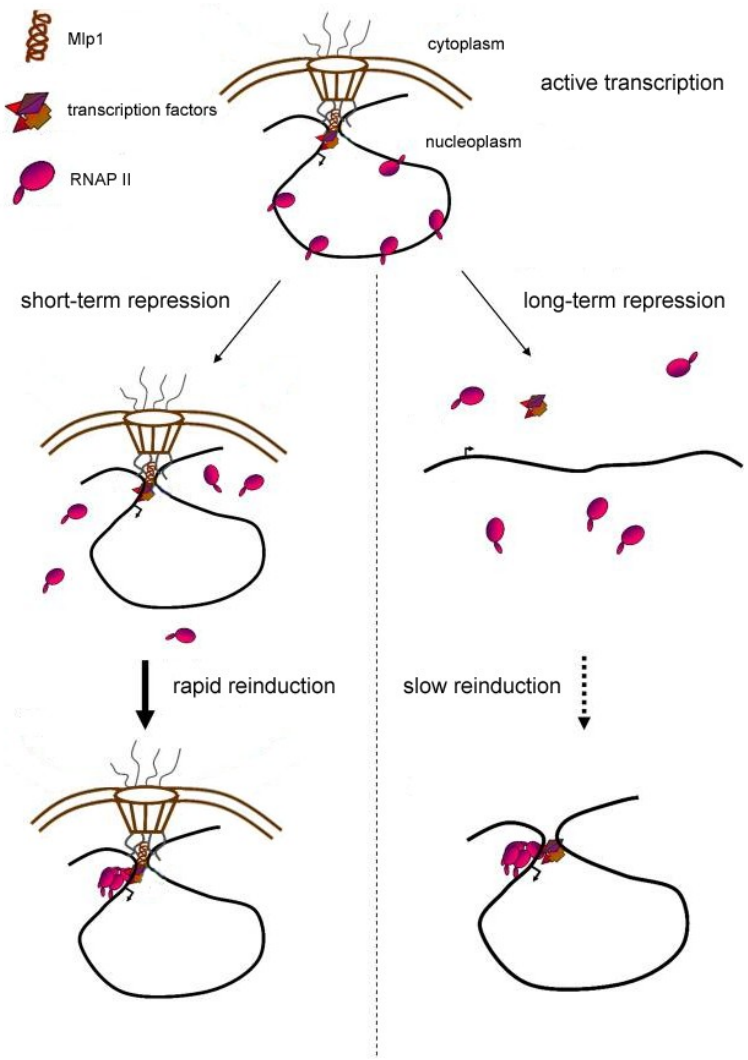

Figure 3. Effect of a short-term and long-term repression on transcriptional memory. (Taken and adapted from Tan-Wong et al., 2009.)

6. Gene loops in the mammalian *BRCA1* gene

BRCA1 (breast cancer susceptibility gene 1) is a gene associated with both sporadic and familial breast cancers and it has been discovered to play an important role in DNA repair, regulation of the cell cycle and apoptosis in response to DNA damage (for review see Yoshida and Miki, 2004). In contrast to familiar breast cancers, where mutation in *BRCA1* is often the cause of the disease, alterations in *BRCA1* sequence have not been observed in sporadic breast cancers

(Futreal et al.,1994; Merajver et al., 1995). Moreover, tumor tissue biopsies revealed significantly lower levels of the *BRCA1* expression in comparison to normal breast tissue, suggesting that not a mutation but a defect in *BRCA1* gene regulation occurs. (Thompson et al., 1995; Magdinier et al., 1998). One possible mechanism for the *BRCA1* downregulation has been suggested to occur through the hypermethylation of the *BRCA1* promoter region (Dobrovic and Simpfendorfer, 1997). However, since this epigenetic modification has been observed only in a subset of all analyzed cases, another regulatory mechanism has to be at work in a *BRCA1*-related tumorigenesis.

The expression of *BRCA1* is under complex regulation by two estrogen-stimulated promoters (Xu et al. 1997). Furthermore, the activity of both promoters is regulated by two conserved noncoding sequences (CNS-1 and CNS-2) in intron 2 of the *BRCA1* gene (Wardrop et al. 2005). These two regulatory elements have a opposite effect on *BRCA1* expression. A mutation in the CNS-1 reduces the *BRCA1* expression by 30%, in turn, a mutation in the CNS-2 leads to 200% increase (Wardrop et al. 2005).

3C analysis of *BRCA1* was performed on the DNA isolated from the breast cancer cell line MCF7. Using primers surrounding 5'-end, 3'-end and intron 2 it was shown that the 5'-end promoter region associates with intron 2 and 3'-end of *BRCA1*. Moreover, additional two gene loops between the 5'-end and the region between introns 2 and 3 and between intron 13 and exon 15 were also detected (Tan-Wong et al., 2008).

Estrogen treatment has been shown to cause an elevation of the *BRCA1* expression indirectly through its mitogenic effect on cell proliferation, possibly as a response to associated increase in DNA synthesis (Spillman and Bowcock, 1996, Marks et al., 1997). The effect of *BRCA1* induction on its 3C looping profile after exposure to estrogen was therefore investigated. Relative cross-linking frequency between 5'-end and 3'-end regions of the *BRCA1* were decreased after 5 h of estrogen treatment and after 24 h the signal corresponding to the 5'-end and 3'-end interaction completely disappeared. Interestingly, the loop formed by interaction of the 5'-end region and regulatory sequences in intron 2 as well as two remaining loops remained unchanged (Tan-Wong et al., 2008). Similar activation/repression dependent looping has been previously observed in yeast. However, in this particular case the promoter and terminator interaction seems to occur during transcription repression, in contrast to gene loops found in

yeast that depend on active transcription (Ansari and Hampsey, 2005; Singh and Hampsey 2007). Surprisingly, treating of MCF7 cells with transcription inhibitor following an estrogen induction impaired the formation of gene loops. Thus, even though promoter-terminator juxtaposition occurs during *BRCA1* repression, active transcription is still required for the formation of gene loops (Tan-Wong et al., 2008).

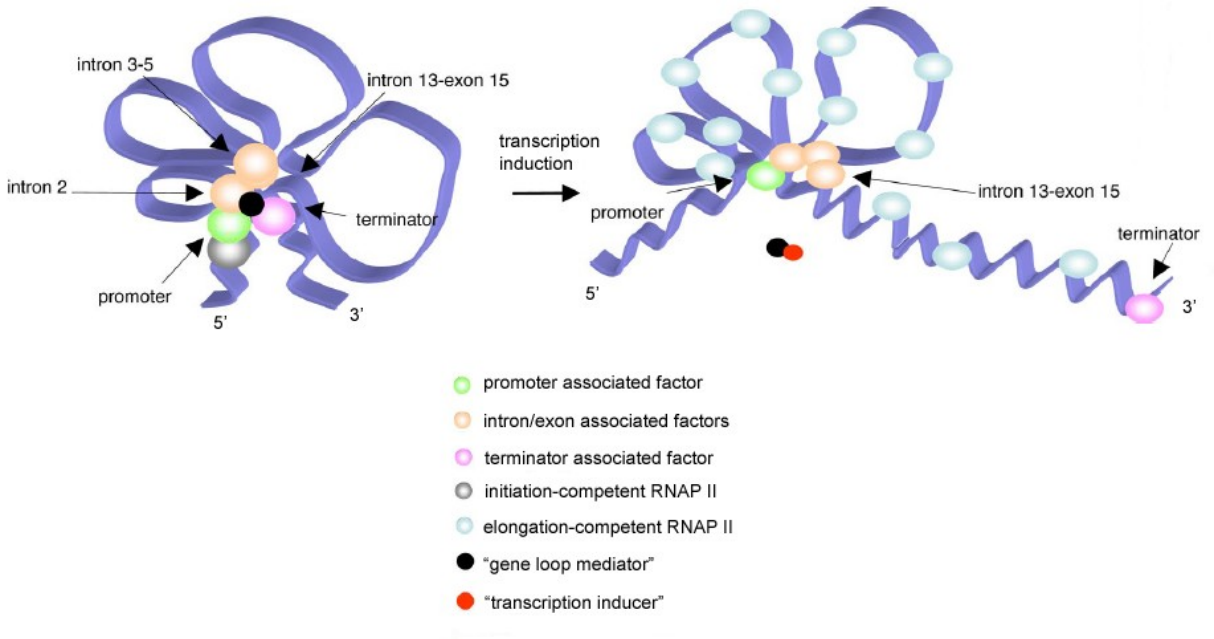

Figure 4. A proposed model of the *BRCA1* gene looping. (Taken and adapted from Tan-Wong et al., 2008.)

The proposed model of a loop-based transcriptional regulation of *BRCA1* involves two different chromatin conformations, schematically depicted in figure 4. In the repressed state the gene forms a so-called "four-leaf clover" structure with promoter and terminator regions brought into proximity along with regulatory sequences in intron 2, sequences in regions spanning introns 3 and 5 and sequences in regions spanning intron 13 and exon 15. The activity of a "gene loop mediator" factor (protein or RNA) has been suggested to mediate the recruitment of all required regulatory factors to corresponding DNA regions, bringing them into proximity and thus forming a "four-leaf clover" conformation. The dependency of the promoter-terminator juxtaposition on active transcription may be explained by a need of a prior transcription and translation of the "gene loop mediator". Upon *BRCA1* induction, the whole chromatin structure is relaxed to a "three-leaf clover" conformation by releasing the promoter-terminator juxtaposition and thus

making the promoter accessible for RNAP II to initiate transcription. It has been proposed, that a "transcription inducer" which tethers the "gene loop mediator" away from the gene loop is required for this change in chromatin conformation and for transcription activation (Tan-Wong et al. 2008).

Different looping profiles were observed in various breast cancer cell lines correspondingly with previous observations of reduced mRNA levels in samples of breast tumor tissues (Tan-Wong et al. 2008; Thompson et al., 1995; Magdinier et al., 1998). These changes in chromatin conformations and therefore in interactions between regulatory elements may affect regulation of *BRCA1* expression. In this regard, further investigation of the role of conserved noncoding sequences in intron 2 (Wardrop et al. 2005) and other possible elements participating in the *BRCA1* looping may provide important insights into regulations of this gene. Deeper understanding of the gene loops and their influence on the *BRCA1* gene expression may provide a crucial tool for future diagnosis and treatment of *BRCA1*-related cancers.

7. Gene loops in the HIV-1 provirus

A common feature of many retroviruses, including the human immunodeficiency virus type 1 (HIV-1), is the presence of long terminal repeats (LTR). These sequences result from the conversion of a single stranded RNA viral molecule into its double stranded DNA proviral form integrated into host's genome by reverse transcription. Two identical LTRs, consisting of three regions: U5, R and U3, flank the itegrated provirus. However, although being identical, each of these LTRs serves a different purpose. While the 5'LTR promotes transcription initiation, the 3'LTR promotes 3'-end processing of HIV-1 transcripts, cleavage and polyadenylation (Böhnlein et al., 1989).

In the 5'LTR, U3 sequence contains RNAP II promoter and enhancer elements critical for the *in vivo* transcription of the virus (Böhnlein et al., 1989). Since the basal transcription of HIV-1 provirus is relatively low, the Tat (*trans*-activator) protein encoded by virus is required for maximal HIV-1 transcription. This protein augments RNAP II initiation and elongation by binding to the hairpin structure TAR (*trans*-activation response element), encoded by the R

region of the 5'LTR and located near the 5'-end of the viral mRNA (Feng and Holland, 1988; Ratnasabapathy et al. 1990).

In the 3'LTR, R and U5 regions promote cleavage and polyadenylation of the transcribed HIV-1 mRNA (Böhnlein et al., 1989). Importantly, since R and U5 regions in the 5'LTR are also capable of promoting 3'-end processing, major splice donor site (MSD) downstream of the 5'LTR is required to inhibit premature processing activity, which causes production of short and defective transcripts (Ashe et al. 1997). In addition, U3 region of the 3'LTR can also serve as a Tat-inducible promoter, similarly to its duplicate in 5'LTR. However, because the activity of the 5'LTR promoter region significantly exceeds that of the 3'LTR, the promoter/enhancer activity of this downstream region is suppressed (Klaver and Berkhout, 1994; Cullen et al., 1984).

Chromatin conformation of the HIV-1 provirus was analyzed in the U1 cell line derived from chronically infected promonocyte cell line U937 (Folks et al., 1987). Since the basal HIV-1 full-length mRNA production in this cell line is very low due to the defective Tat, an exogenous Tat expression or an exposure to TPA phorbol ester are required to induce proviral transcription (Emiliani et al., 1998; Lusic et al., 2003).

Upon Tat-induced and TPA-induced transcription, 3C and sequencing revealed interactions of the 5' U3 promoter with the 3' U5 poly(A) signal and between both LTRs and an adjacent 5' proximal sequence containing the MSD. These interactions, albeit very weak, were also observed in non-induced cells, consistent with low basal HIV-1 transcription. Moreover, after blocking the ongoing TPA-induced transcription by flavopridol treatment, the 3C looping signal was diminished back to the control levels, suggesting that the formation of HIV-1 gene loops depends on active transcription (Perkins et al., 2008).

The possible dependence of gene loops formation on specific HIV-1 sequences was examined using HeLa cells transfected with proviral plasmid. Since 3C of this plasmid revealed the same genomic interactions that were observed in the U1 cell line, this approach enabled direct manipulation of HIV-1 regulatory sequences (Perkins et al., 2008).

After introducing an artificial point mutation to the MSD region, the previously observed HIV-1 looping was completely lost and the plasmid adopted a simple circular conformation. Since MSD is required for blocking the activity of R and U5 regions of the 5'LTR, this structural

change correlates with a switch from full-length HIV-1 mRNA transcription, terminating at the 3'LTR poly(A) signal, to the production of short and defective mRNAs prematurely terminated after encountering the 5'LTR poly(A). Furthermore, a mutation in the 3' poly(A) signal in the 3'LTR also impaired the 5'-3'LTRs juxtaposition, while 5'LTR-MSD interaction remained unaffected. Interestingly, replacing the 5' U3 promoter with the cytomegalovirus (CMV) promoter or replacing the 3'LTR with a synthetic poly(A) site (SPA), while leaving the 5'LTR unchanged, did not affect HIV-1 provirus looping (Perkins et al., 2008).

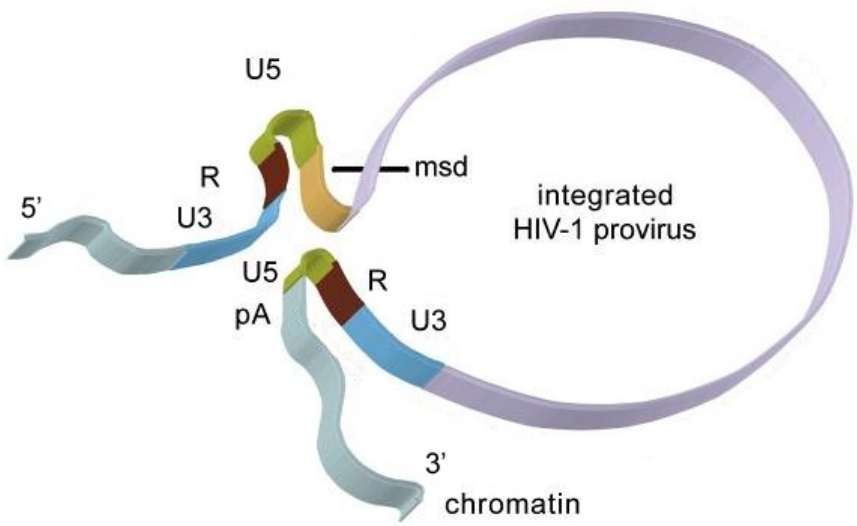

Figure 5. Interactions between 5' and 3' LTRs causing the looping conformation of the HIV-1 provirus. (Taken and adapted from Perkins et al., 2008.)

Taken together, these results show that the proviral genome forms two loops under ongoing transcription as depicted in figure 5. One loop is formed by juxtaposition of the 5' U3 and R regions with the 3' U5 poly(A) signal region, second loop juxtaposes both LTRs with the 5' proximal MSD sequence. The latter interaction is consistent with the involvement of MSD in blocking the cleavage and polyadenylation activity of 5'LTR poly(A) signal in favor of the 3'LTR poly(A) (Ashe et al., 1997). Active transcription is not the only prerequisite of HIV-1 looping, although it seems that ongoing transcription is crucial for a long-term maintenance of HIV-1 gene loops. A point mutation in the MSD that activates the 5'LTR poly(A) signal, as well as an inactivation of the 3' poly(A) signal both abolished HIV-1 looping conformation. This suggests that HIV-1 gene loops strictly depend on competitive interactions between the two

different cleavage and polyadenylation signals and secure the correct transcription of the full-length HIV-1 mRNA. In turn, neither the replacement of the 5' U5 promoter with the CMV promoter nor the replacement of the 3' poly(A) with a synthetic poly(A) site did affect gene loops formation. Therefore, gene loops may not be limited to the HIV-1 retrovirus only.

8. Conclusions

In this thesis the recent knowledge about gene loops and juxtapositions of promoter and terminator regions in a wide range of genes transcribed by RNA polymerase II has been presented.

In yeast, it was shown that an initial round of transcription with the subsequent physical interaction between the 3'-end processing complex and the general transcription factor TFIIB is required for the formation of gene loops which, in some cases, were detected even after a certain period of transcription repression. This persistence of gene loops under repressive conditions was linked to the phenomenon of transcriptional memory, which enables yeast cells to rapidly reach maximum transcription levels in case of transcription re-induction.

Another example of gene loops was reported in the *BRCA1* tumor suppressor, a gene associated with breast cancer. *BRCA1* was showed to respond to estrogen induction with distinct changes in looping conformation. Interestingly, although the looping is transcription dependent, similarly to the case observed in yeast genes, the juxtaposition of the promoter and terminator of *BRCA1* gene occurs only during the repressed state of the gene. Therefore, this interaction may block initiation factors from binding to promoter elements and relaxation of this structure is required to release the promoter which makes the transcription initiation possible.

Finally, the looping conformation of the HIV-1 provirus has been described. In the integrated HIV-1 provirus, the 3'-end and 5'-end long terminal repeats (LTRs) together with a major splicing donor site are brought into close proximity by a yet unknown mechanism. This juxtaposition depends on an ongoing transcription and strictly depends on a competitive interaction between the two cleavage and polyadenylation signal sequences.

Gene loops and their functions in eukaryotic transcription are still little understood and a lot of questions still remains to be answered. Since a wide range of genes have been discovered to display a transcription dependent looping conformation in yeast, it would be of highest interest to find gene loops in higher eukaryotes. What is the precise molecular mechanism behind the formation of these loops? Is there just one general mechanism of gene looping or more of them? Mammalian Ssu72 and TFIIB proteins were shown to physically interact as their yeast homolgs do (St-Pierre et al., 2005), which raises a possibility that the mechanism of gene looping in higher eukaryotes may be similar to the mechanism in yeast. Since many metazoan cells display a rapid response to certain stimuli, they would certainly benefit from the rapid reactivation of response genes by a mechanism similar to yeast transcriptional memory. However, this possibility remains to be investigated.

The general process of RNAP II transcription is well understood. However, further research of gene specific transcription regulation by gene loops may lead to future understanding of mechanisms that give rise to some pathological conditions. It has been suggested that *BRCA1* gene loops may participate in the regulation of *BRCA1* expression. Since there have been numerous reports of a downregulation of *BRCA1* mRNA levels in breast tumor tissues, alterations in normal *BRCA1* looping may be involved in this process.

Gene loops could also provide another explanation for the effect of viral oncogenesis. HIV-1 transcription normally starts from the 5'LTR promoter. However, the disruption of the HIV-1 provirus looping conformation, with a successive loss of suppression of the 3'LTR promoter and enhancer activity, may enhance the transcription of downstream oncogenes.

Overall, gene loops and promoter-terminator juxtapositions provide another mechanism of regulation of gene expression. Since these promoter-terminator interactions have been observed in a wide range of genes in both yeast and mammalian cells, it is possible that this type of transcription control is a wide-spread phenomenon of eukaryotic RNAP II transcription. Last but not least, deeper understanding of the involvement of gene loops in the regulation of gene expression may help to explain the mechanisms behind various pathological conditions.

References

Ansari A, Hampsey M. 2005. A role for the CPF 3'-end processing machinery in RNAP II-dependent gene looping. Genes Dev 19: 2969-78.

Ashe MP, Pearson LH, Proudfoot NJ. 1997. The HIV-1 5' LTR poly(A) site is inactivated by U1 snRNP interaction with the downstream major splice donor site. EMBO J 16: 5752-63.

Böhnlein S, Hauber J, Cullen BR. 1989. Identification of a U5-specific sequence required for efficient polyadenylation within the human immunodeficiency virus long terminal repeat. J Virol 63: 421-4.

Brickner JH, Walter P. 2004. Gene recruitment of the activated *INO1* locus to the nuclear membrane. PLoS Biol 2: e342.

Brickner DG, Cajigas I, Fondufe-Mittendorf Y, Ahmed S, Lee PC, Widom J, Brickner JH. 2007. H2A.Z-mediated localization of genes at the nuclear periphery confers epigenetic memory of previous transcriptional state. PLoS Biol 5: e81.

Buratowski S. 1994. The basics of basal transcription by RNA polymerase II. Cell 77: 1-3. Review.

Casolari JM, Brown CR, Komili S, West J, Hieronymus H, Silver PA. 2004. Genome-wide localization of the nuclear transport machinery couples transcriptional status and nuclear organization. Cell 117: 427-39.

Casolari JM, Brown CR, Drubin DA, Rando OJ, Silver PA. 2005. Developmentally induced changes in transcriptional program alter spatial organization across chromosomes. Genes Dev 19: 1188-98.

Cho EJ, Kobor MS, Kim M, Greenblatt J, Buratowski S. 2001. Opposing effects of Ctk1 kinase and Fcp1 phosphatase at Ser 2 of the RNA polymerase II C-terminal domain. Genes Dev 15: 3319-29.

Cullen BR, Lomedico PT, Ju G. 1984. Transcriptional interference in avian retroviruses—implications for the promoter insertion model of leukaemogenesis. Nature 307: 241-5.

Dekker J, Rippe K, Dekker M, Kleckner N. 2002. Capturing chromosome conformation. Science 295: 1306-11.

Dichtl B, Blank D, Ohnacker M, Friedlein A, Roeder D, Langen H, Keller W. 2002. A role for SSU72 in balancing RNA polymerase II transcription elongation and termination. Mol Cell 10: 1139-50.

Dobrovic A, Simpfendorfer D. 1997. Methylation of the *BRCA1* gene in sporadic breast cancer. Cancer Res 57: 3347-50.

Dohmen RJ, Wu P, Varshavsky A. 1994. Heat-inducible degron: a method for constructing temperature-sensitive mutants. Science 263: 1273-6.

Emiliani S, Fischle W, Ott M, Van Lint C, Amella CA, Verdin E. 1998. Mutations in the tat gene are responsible for human immunodeficiency virus type 1 postintegration latency in the U1 cell line. J Virol 72: 1666-70.

Feng S, Holland EC. 1988. HIV-1 tat trans-activation requires the loop sequence within tar. Nature 334: 165-7.

Folks TM, Justement J, Kinter A, Dinarello CA, Fauci AS. 1987. Cytokine-induced expression of HIV-1 in a chronically infected promonocyte cell line. Science 238: 800-2.

Futreal PA, Liu Q, Shattuck-Eidens D, Cochran C, Harshman K, Tavtigian S, Bennett LM, Haugen-Strano A, Swensen J, Miki Y, et al. 1994. *BRCA1* mutations in primary breast and ovarian carcinomas. Science 266: 120-2.

Galy V, Gadal O, Fromont-Racine M, Romano A, Jacquier A, Nehrbass U. 2004. Nuclear retention of unspliced mRNAs in yeast is mediated by perinuclear Mlp1. Cell 116: 63-73.

Kadauke S, Blobel GA. Chromatin loops in gene regulation. 2009. Biochim Biophys Acta 1789: 17-25.

Klaver B, Berkhout B. 1994. Comparison of 5' and 3' long terminal repeat promoter function in human immunodeficiency virus. J Virol 68: 3830-40.

Krishnamurthy S, He X, Reyes-Reyes M, Moore C, Hampsey M. 2004. Ssu72 Is an RNA polymerase II CTD phosphatase. Mol Cell 14: 387-94.

Kundu S, Horn PJ, Peterson CL. 2007. SWI/SNF is required for transcriptional memory at the yeast GAL gene cluster. Genes Dev 21: 997-1004.

Lainé JP, Singh BN, Krishnamurthy S, Hampsey M. 2009. A physiological role for gene loops in yeast. Genes Dev 23: 2604-9.

Lusic M, Marcello A, Cereseto A, Giacca M. 2003. Regulation of HIV-1 gene expression by histone acetylation and factor recruitment at the LTR promoter. EMBO J 22: 6550-61.

Luthra R, Kerr SC, Harreman MT, Apponi LH, Fasken MB, Ramineni S, Chaurasia S, Valentini SR, Corbett AH. 2007. Actively transcribed *GAL* genes can be physically linked to the nuclear pore by the SAGA chromatin modifying complex. J Biol Chem 282: 3042-9.

Magdinier F, Ribieras S, Lenoir GM, Frappart L, Dante R. 1998. Down-regulation of *BRCA1* in human sporadic breast cancer; analysis of DNA methylation patterns of the putative promoter region. Oncogene 17 : 3169-76.

Marks JR, Huper G, Vaughn JP, Davis PL, Norris J, McDonnell DP, Wiseman RW, Futreal PA, Iglehart JD. *BRCA1* expression is not directly responsive to estrogen. 1997. Oncogene 14: 115-21.

Merajver SD, Pham TM, Caduff RF, Chen M, Poy EL, Cooney KA, Weber BL, Collins FS, Johnston C, Frank TS. 1995. Somatic mutations in the *BRCA1* gene in sporadic ovarian tumours. Nat Genet 9: 439-43.

Murrell A, Heeson S, Reik W. 2004. Interaction between differentially methylated regions partitions the imprinted genes *Igf2* and *H19* into parent-specific chromatin loops. Nat Genet 36: 889-93.

Nedea E, He X, Kim M, Pootoolal J, Zhong G, Canadien V, Hughes T, Buratowski S, Moore CL, Greenblatt J. 2003. Organization and function of APT, a subcomplex of the yeast cleavage and polyadenylation factor involved in the formation of mRNA and small nucleolar RNA 3'-ends. J Biol Chem 278: 33000-10.

Nonet M, Scafe C, Sexton J, Young R. 1987. Eucaryotic RNA polymerase conditional mutant that rapidly ceases mRNA synthesis. Mol Cell Biol 7: 1602-11.

O'Sullivan JM, Tan-Wong SM, Morillon A, Lee B, Coles J, Mellor J, Proudfoot NJ. 2004. Gene loops juxtapose promoters and terminators in yeast. Nat Genet 36: 1014-8.

Perkins KJ, Lusic M, Mitar I, Giacca M, Proudfoot NJ. 2008. Transcription-dependent gene looping of the HIV-1 provirus is dictated by recognition of pre-mRNA processing signals. Mol Cell 29: 56-68.

Ratnasabapathy R, Sheldon M, Johal L, Hernandez N. 1990. The HIV-1 long terminal repeat contains an unusual element that induces the synthesis of short RNAs from various mRNA and snRNA promoters. Genes Dev 4: 2061-74.

Singh BN, Hampsey M. 2007. A transcription-independent role for TFIIB in gene looping. Mol Cell 27: 806-16.

Spillman MA, Bowcock AM. 1996. *BRCA1* and *BRCA2* mRNA levels are coordinately elevated in human breast cancer cells in response to estrogen. Oncogene 13: 1639-45.

St-Pierre B, Liu X, Kha LC, Zhu X, Ryan O, Jiang Z, Zacksenhaus E. 2005. Conserved and specific functions of mammalian ssu72. Nucleic Acids Res 33: 464-77.

Sun ZW, Hampsey M. 1996. Synthetic enhancement of a TFIIB defect by a mutation in SSU72, an essential yeast gene encoding a novel protein that affects transcription start site selection *in vivo*. Mol Cell Biol 16: 1557-66.

Taddei A, Van Houwe G, Hediger F, Kalck V, Cubizolles F, Schober H, Gasser SM. 2006. Nuclear pore association confers optimal expression levels for an inducible yeast gene. Nature 441: 774-8.

Tan-Wong SM, French JD, Proudfoot NJ, Brown MA. 2008. Dynamic interactions between the promoter and terminator regions of the mammalian *BRCA1* gene. Proc Natl Acad Sci U S A 105: 5160-5.

Tan-Wong SM, Wijayatilake HD, Proudfoot NJ. 2009. Gene loops function to maintain transcriptional memory through interaction with the nuclear pore complex. Genes Dev 23: 2610-24.

Thompson ME, Jensen RA, Obermiller PS, Page DL, Holt JT. 1995. Decreased expression of *BRCA1* accelerates growth and is often present during sporadic breast cancer progression. Nat Genet 9: 444-50.

Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W. 2002. Looping and interaction between hypersensitive sites in the active beta-globin locus. Mol Cell 10: 1453-65.

Vinciguerra P, Iglesias N, Camblong J, Zenklusen D, Stutz F. 2005. Perinuclear Mlp proteins downregulate gene expression in response to a defect in mRNA export. EMBO J 24: 813-23.

Wardrop SL, Brown MA; kConFab Investigators. 2005. Identification of two evolutionarily conserved and functional regulatory elements in intron 2 of the human *BRCA1* gene. Genomics 86: 316-28.

Xu CF, Chambers JA, Solomon E. 1997. Complex regulation of the *BRCA1* gene. J Biol Chem 272: 20994-7.

Yoshida K, Miki Y. 2004. Role of *BRCA1* and *BRCA2* as regulators of DNA repair, transcription, and cell cycle in response to DNA damage. Cancer Sci 95: 866-71. Review.

Yudkovsky N, Ranish JA, Hahn S. 2000. A transcription reinitiation intermediate that is stabilized by activator. Nature 408: 225-9.

Zacharioudakis I, Gligoris T, Tzamarias D. 2007. A yeast catabolic enzyme controls transcriptional memory. Curr Biol 17: 2041-6.

Zhang DW, Rodríguez-Molina JB, Tietjen JR, Nemec CM, Ansari AZ. 2012. Emerging Views on the CTD Code. Genet Res Int 2012: 347214.