Tesis Doctoral

Biología Funcional y Genómica

Instituto de Biología Funcional y Genómica

Prakhar Bisht

"Mechanisms modulating transcribed chromatin topology and suppressing replication fork instability"

Salamanca, 2017

TABLE OF CONTENT

ABSTRACT **1**

RESUMEN **3**

I. INTRODUCTION/ *INTRODUCCIÓN*

CHAPTER 1 **4**
1. Yeast as a study model 4
1.2 Cell cycle checkpoints in *S. cerevisiae* 5
2. DNA Replication 6
2.1 DNA topology during replication and its regulation 8
2.1.1 DNA topology 9
2.1.2 DNA topoisomerases; the enzymes responsible for modulating DNA topology 11
2.1.3 Importance of DNA topology modulation during DNA replication 14
2.2 Natural impediments to DNA Replication 17
2.2.1 DNA binding proteins 17
2.2.2 Replication-Transcription collisions 19

CHAPTER 2 **33**
3. DNA Secondary structures 33
3.1 Consequences of secondary structure formation 34
3.2 G-quadruplex; as secondary structures 35
3.2.1 Structures of G-quadruplexes 36
3.2.2 Effects of g-quadruplex on DNA replication 37
3.2.3 Role of G4 unwinding helicases 37
3.2.4 CEB25 a G-quadruplex forming human minisatellite 38
3.2.5 C-MYC; three-tetrad G-quadruplexes in gene promoters 39

II. OBJECTIVES / *OBJETIVOS* **42**

VI. MATERIALS AND METHODS / *MATERIALES Y MÉTODOS* **78**
7.1 Strains and plasmids 78
7.2 Growing media 80
7.3 List of buffers 81
7.4 P.C.R 83
7.5 High efficiency LiAc transformation 84
7.6 Growth conditions, cell cycle arrest 85
7.7 Serial dilution and spot assay 86
7.8 TCA protein extraction 86
7.9 SDS PAGE and Western Blot analysis 86

7.10 FACS analysis 87
7.11. Co-immunoprecipitation assay 87
7.12 Neutral/Neutral 2D gel electrophoresis analysis 88
7.13 Southern Blot and hybridization 91
7.14 Chloroquine two-dimensional gel electrophoresis technique 92

VII. BIBLIOGRAPHY / *BIBLIOGRAFÍA* **96**

VIII. ACKNOWLEDGMENTS **108**

FIGURE INDEX

Figure 1.1. Representation of the budding yeast cell cycle.

Figure 1.2. Schematic model of the initiation of DNA synthesis.

Figure 1.3. Examples of topological domains.

Figure 1.4. Twisting and writhing in the tubing model of DNA.

Figure 1.5. Mechanism of DNA topoisomerase I for solving supercoiling.

Figure 1.6. Generalized mechanism for the action of DNA topoisomerase II.

Figure 1.7. Model for topology of the replicating chromosome.

Figure 1.8. Formation of pre-catenated DNA at the termination of replication.

Figure 1.9. Organization of rRNA repeats in *S. cerevisiae*.

Figure 1.10. The rRNA gene transcript.

Figure 1.11. Schematic representation of a gene transcription regulation by RNA polymerase II.

Figure 1.12. The FACT complex enables RNA polymerase II transcription of nucleosome-packed DNA.

Figure 1.13. Stepwise representation of stages in mRNA maturation.

Figure 1.14. The THO/ TREX complex.

Figure 1.15. THO complex role in mRNP biogenesis at the interface between transcription and export of mRNA from the nucleus to cytoplasm.

Figure 1.16. Directionality of co-directional and head on collision encounters of replication and transcription machineries.

Figure 1.17. Outcomes of transcription-replication collisions.

Figure 2.1. Non B-form DNA secondary structures.

Figure 2.2. Schematic illustration of a G-quadruplex structures.

Figure 2.3. G-quadruplex structures during DNA replication.

Figure 2.4. The CEB25 G-quadruplex structure depicted by two-dimensional nuclear magnetic resonance spectroscopy (2D NMR).

Figure 2.5. Schematic representation of NMR structure of *C-MYC* minisatellite forming G-quadruplex.

Figure 7.1 Schematic representation of the replication intermediates detectable by 2D gel analysis.

Figure 7.2. Schematic representation of the second dimension preparation.

Figure 7.3. Schematic representation of the migration of topoisomers in two dimension chloroquine gels.

TABLE INDEX

Table 1. Subfamilies of DNA topoisomerases.

Table 7.1. Genotype of strains generated and used in this study.

Table 7.2: Plasmid used in this study.

Table 7.3. Plasmids used for PCR**.**

.

INTRODUCTION

CHAPTER 1

1. YEAST AS A STUDY MODEL

From the past two decades the budding yeast *Saccharomyces cerevisiae* has been a preferred model system for molecular genetics research. It has the basic cellular mechanics of DNA replication, recombination, cell division and metabolism; which are well conserved from yeast to higher eukaryotes, including mammals. Also, the budding yeast genome, with a haploid set of 16 chromosomes, has been totally sequenced twenty years ago. The use of *S. cerevisiae* as a study model has the advantage that the cell cycle is relatively concise with respect to higher eukaryotes and can be easily monitored by the aid of a microscope due to the characteristic cell shape in the different cell cycle phases **(Figure 1.1)**. Likewise, the life cycle of *S. cerevisiae* makes it an amicable genetic system. It is a unicellular organism that can grow in a diploid state as well as in a haploid state with two different mating types (Mat a and Mat α). Mat a and Mat α haploid cells can mate in order to produce a diploid cell. These diploid cells can undergo meiosis under conditions unfavorable to cell growth, leading to the formation of four haploid spores in which the inheritance of the genetic information follows a mendelian pattern. Hence, serving a great system to segregate and study various genetic traits.

1.2. Cell cycle checkpoints in *S. cerevisiae*:

The cell cycle is the progression of events where a cell has to duplicate and divide its cellular content into two daughter cells in such a way that each cell has the information and machinery required to reiterate the same process. In eukaryotic division, all the genetic material is accurately replicated and segregated to the two daughter cells. The DNA replication process and sister chromatid segregation takes place in two discrete phases. These are known as S-phase; the DNA synthesis phase and M-phase; the mitosis phase. These two phases are further separated by two gap phases, known as G1 and G2.

The control of eukaryotic cell growth and division is monitored by molecular mechanisms known as “checkpoints”, which determine a proper timing for all cellular processes. Passage through a checkpoint from one phase of cell cycle to the next requires a coordinated function of proteins that monitor cell growth and DNA integrity. Uninhibited cell propagation and division of damaged DNA can contribute to genomic instability and tumorigenesis.

The three major cell cycle checkpoints known are:

- **G1 (Restriction) checkpoint**: Monitors cell to reach the required size so as to undergo a new round of DNA synthesis.
-
- **The spindle assembly checkpoint**: Ensures that DNA replication is complete, chromosomes are aligned on the metaphase plate to facilitate correct sister chromatid segregation.

- **G2-M DNA damage checkpoint**: Ensures that the cells have completed DNA synthesis and is ready to entry into the proliferative Mitotic (M) phase.

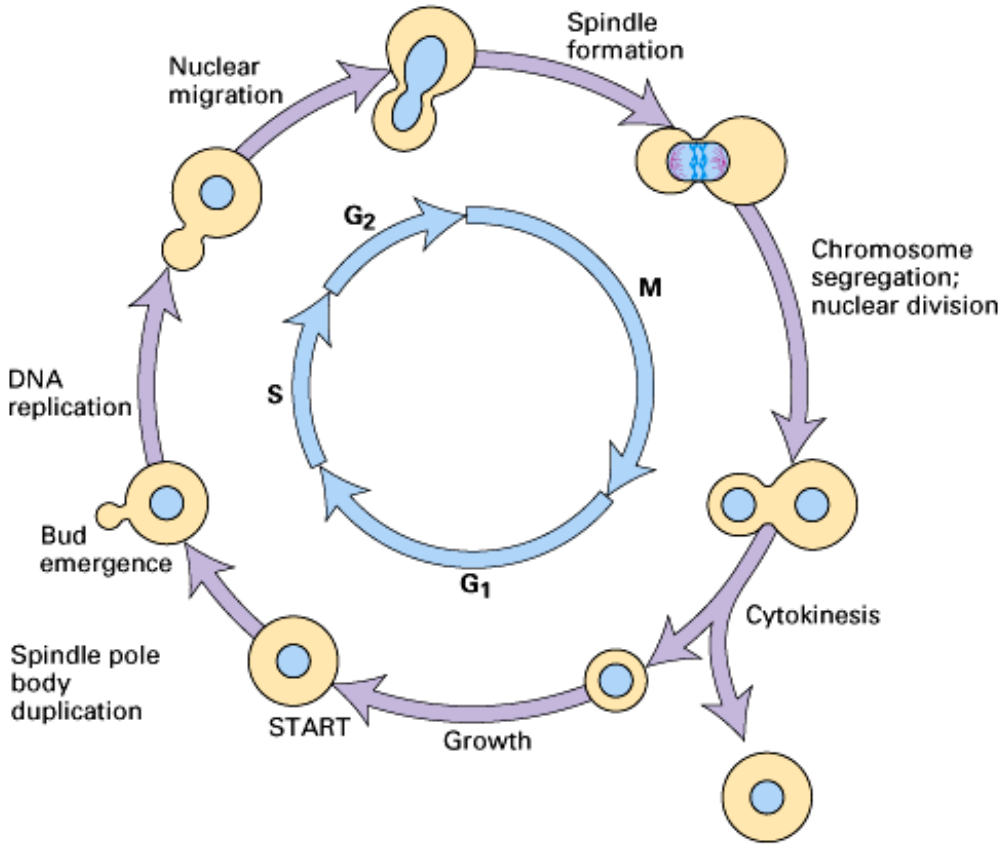


Figure 1.1. Representation of the budding yeast cell cycle. The cell cycle is divided in four stages: G1, S, G2 & M. In G1 the cells remain unbudded while bud appearance marks the onset of S-phase. In G2, cells grow further in size to enter into the M phase (Mitosis phase); which is evident by the presence of bi-lobed structure of the cells (adapted from http://www.pha.jhu.edu/~ghzheng/old/webct/note7_3.htm).

2. DNA REPLICATION

Maintenance of life requires duplication of the genetic information stored in the DNA of each living organism. In eukaryotic cells, before a cell divides it has to duplicate its genome so that two identical copies of its DNA content can be partitioned into daughter cells. DNA replication is initiated at thousands of origins on the DNA, each giving rise to two replisomes that travel away from the initiation site in opposite directions.

During DNA replication the double helical structure requires the separation of the two individual DNA strands. This process is known as the strands melting or unwinding and is vital to allow DNA polymerases carry out DNA synthesis. Local unwinding of the duplex takes place at specific sites within the chromosomes known as replication origins. Unlike prokaryotes, where DnaA is the only protein catalyzing the unwinding, in eukaryotes this process requires involvement of a multi-protein complex with the Origin Recognition Complex (ORC) that binds to DNA replication origins (Bell,S.P;A.Dutta 2002), along with two additional proteins, CDC6 and CDT1 **(Figure 1.2).** This multi-protein complex is responsible for the recruitment of the core component of the putative replicative helicase, the MCM2-7 protein complex, in an ATP-dependent manner. This phenomenon is presently known as "licensing" or pre-RC (replication complex) formation. *In vitro* studies have shown that the budding yeast MCM2-7 complex reconstituted from pre-RC is loaded on DNA as a salt-resistant head-to-head double hexamer complex in an ATP-dependent manner (Evrin, C.; Clarke, P 2009). This configuration is considered to be essential for initiation of a pair of bidirectional replication forks.

Recruitment of additional factors onto the pre-RC, i.e. the CDC45 protein and the GINS complex, holsters the formation of the pre-Initiation Complex (pre-IC) whose assembly requires Cyclin-Dependent Kinase (CDK) activity. Consequently, this reaction is specifically inhibited by CDK inhibitors such as Sic1 **(Figure 1.2)**. It is believed that CDK activity is important to phosphorylate key factors of pre-IC in order to activate them. CDC45 allows the assembling of initiation complexes by recruitment of DNA polymerases α and δ at replication origins (Mimura. S et al 1998) (Walter. J et al 2000). Recent studies have illustrated that the MCM2-7 complex physically interacts with CDC45 and the GINS complex, thus forming the CMG complex (MacNeill. S et al 2010). This complex is believed to constitute the active form of the replicative helicase. Structural studies identified that the CDC45 protein binding to the MCM2-7 complex stabilizes the ring-shaped structure of this helicase onto the DNA (Costa. A et al 2010). The CDC7 protein kinase along with its regulatory subunits Dbf4 or Drf1 (also called DDK, for Dbf4- or Drf1-Dependent Kinase), phosphorylates MCM2-7 subunits and stimulates MCM2-7 function probably assisting a conformational change in the helicase complex (Sheu. Z et al 2010) (Randell et al 2010). In *S. cerevisiae*, two additional proteins, SLD2 and SLD3, have also been shown to be required at the initiation of DNA synthesis. Their phosphorylation by CDKs is an essential step in formation of pre-ICs (Tanaka. S et al 2007) (Zegerman. P et al 2007) which is considered to determine the complex formation between Dpb11/TopBP1, SLD2, GINS and DNA polymerase ε, and its subsequent recruitment on DNA.

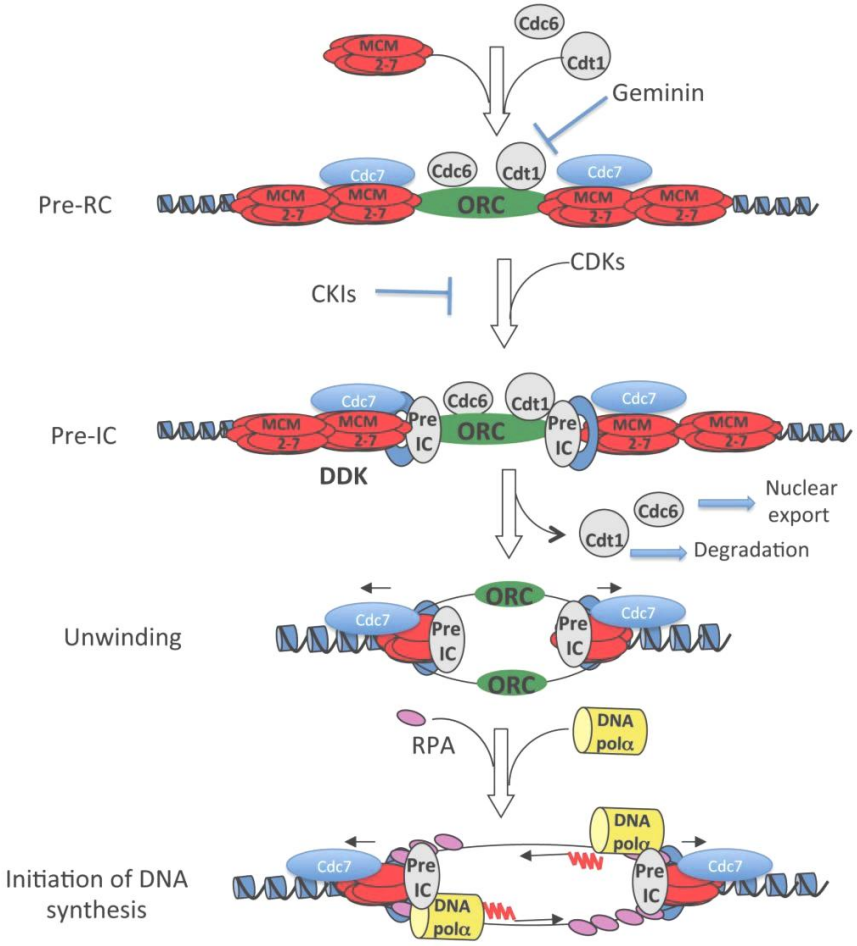


Figure 1.2. Schematic model of the initiation of DNA synthesis. DNA replication is categorically divided in Pre-RC (replication complex), Pre-IC (initiation complex), unwinding and DNA synthesis. For simplicity the pre-IC consist of multiple proteins in one complex i.e. Cdc45, the GINS complex, MCM10, TopBP1, SLD2 and SLD3. A ring-like complex of proteins (blue) made of Claspin, Timeless and Tipin make the Fork Protection Complex that links the GMC helicase to the DNA polymerases. RNA primers (red waves) are made by the catalytic activity of DNA primase and elongated by the catalytic activity of DNA polymerase α (black arrows linked to red waves) (Adapted from Bénédicte Recolin et al 2014)

On activation, the helicase unwinds the DNA thereby generating long stretches of single-stranded DNA (ssDNA), which acts as a substrate for recruiting the heterotrimeric complex RPA (Replication Protein A). RPA binding stimulates the DNA polymerase activity on ssDNA, which is generated by the helicase itself during the unwinding reaction (Jiang. X et al 2006). RPA-loaded ssDNA facilitates the recruitment of DNA polymerase α/primase holoenzyme **(Figure 1.2),** which consists of five distinct polypeptides. This complex has the unique ability to initiate DNA synthesis *de novo,* i.e. in the absence of a 3' hydroxyl (OH) end. Due to its associated

primase activity, it synthesizes short 10 bases-long ribonucleotides stretches at melted replication origins that are then extended by the catalytic subunit of DNA polymerase α. These replication intermediates act as substrates for another DNA polymerase, Pol δ, whose processive activity then allows an efficient replication of the whole genome. This is aided by the recruitment of PCNA; a replication protein that functions as DNA polymerase δ cofactor, which strongly stimulates its polymerase activity. PCNA is known to form a homotrimeric ring which is loaded around the DNA in a reaction catalyzed by RF-C (Replication Factor C). This pentameric protein complex regulates the opening of the PCNA ring in an ATP-dependent manner and then reseals it around the DNA. Since the DNA molecule is asymmetric, DNA synthesis occurs continuously on one strand i.e. leading strand and discontinuously on the opposite strand i.e. lagging strand. It has been also suggested that another polymerase, DNA polymerase ε, required for chromosomal DNA replication, functions on the leading strand DNA synthesis by extending the initial primer generated by DNA polymerase α/primase.

When the three DNA polymerases are recruited onto the DNA the replisome is formed. The DNA replication machinery duplicates the genome with an average speed of 3000 bp per minute (Hubermann J.A. et al 1966), making very few or no mistakes. On the lagging strand, additional factors are required in the maturation and sealing of DNA chains. Factors like the Fen1 exonuclease remove the RNA primers from the Okazaki fragments while DNA ligase I seals the gaps left between Okazaki fragments. Since DNA replication is bidirectional, two replisomes simultaneously originate from one single replication origin polymerizing DNA chains spreading outwards. Hence, forming the basic unit of a replication domain called the replicon (Mechali. M et al 2001). Likewise, circumstances occur where two replisomes will encounter each other when two replicons have been completely replicated. At this stage, the replisome has to be unloaded from chromatin. The mechanisms responsible for unloading of the replisome are till date not well understood. Although, it has been reported that the MCM-BP protein, may function as the unloader of the MCM2-7 complex form replicated chromatin, but the molecular mechanism of this reaction remains still to be understood (Nishiyama et al 2011). Recently, it has been proposed that the E3 ligase SCF^{Dia2} drives the ubiquitylation of the CMG helicase on its Mcm7 subunit and recruits the Cdc48 segregase, associated with the ubiquitylated CMG, leading rapidly to the disassembly of the helicase from the newly replicated DNA. (Maric M et al 2014)

2.1. DNA topology during replication and its regulation:

Topological aspects of DNA structure arise primarily from the fact that the two DNA strands are repeatedly intertwined. Untangling these two strands, which occurs in all major genetic processes such as DNA replication by DNA helicases imposes a series of constraints. Hence, to understand the consequences of torsional stress introduced by the replisome, it is first important to understand in brief what, how and by whom DNA topology is monitored.

2.1.1. DNA topology:

A DNA molecule consists of two complementary polynucleotide chains that are interwound, acquiring a double helix structure. In physiological conditions it is found in a B-DNA conformation, where a right-handed helix with a period of approximately 10.5 base pairs (bp) per turn is formed. For all natural DNA forms, the free end rotation is either restricted or forbidden altogether. When the ends of DNA segment are unable to rotate within a closed space it is known as a topological domain. A simplest example of such is the covalently closed circular DNA, found in plasmids and viruses (**Figure 1.3.a**). Although eukaryotic chromosomes are linear overall, they too consist of large DNA loops firmly attached to the nuclear matrix or protein bodies (**Figure 1.3.b, c**). Hence, the principle of basic topology for circular DNAs, for study purposes, can be generalized for both these classes.

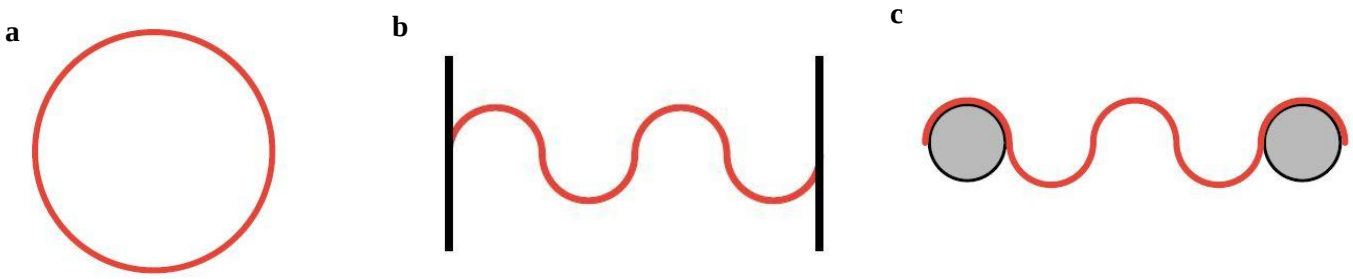


Figure 1.3. Examples of topological domains. (a) Circular DNA, (b) linear DNA attached to the membrane, (c) linear DNA attached to protein aggregates (adapted from Sergei M Mirkin et al 2001)

The fundamental parameter for understanding topology of a covalently closed circular DNA is called the linking number (**Lk**). There are two fundamental features of the Lk. First, Lk is always an integer. Second, Lk cannot be changed due to any deformation of the DNA strands, i.e. it is invariant. Another parameter of a circular DNA is called "Twist or Tw" (**Figure 1.4.b**). Tw signifies how the individual strands of DNA coil around the axis of the DNA helix. In a right-handed helix DNA with ≈10.5 base pairs (bp) per turn, Tw is mostly a positive value for any naked DNA. "Writhing (**Wr**)" is another important characteristic of circular DNA, is a measure of the coiling of the helix axis in space, i.e. the shape of the DNA molecule as a whole (**Figure 1.4.c**). With the aid of these parameters following equation can be formalized:

$$\mathbf{Lk = Tw + Wr}$$

It implies that for a given closed-circular molecule, since Lk is invariant, any change in the twist of the DNA must be accompanied by an equal and opposite change in the writhe, and vice versa. If number of base pairs per DNA turn is designated as "helical repeats or ***h***", then the Lk value of the **N**-base pair long circular DNA molecule corresponds to;

$$\mathbf{Lk_o = Tw_o = N/}\boldsymbol{h}$$

When the Wr is zero and Tw is equal to Lk a molecule is in a relaxed state, which resembles the planer circular DNA (**Figure 1.4.a**) and its Lk is taken as $\mathbf{Lk_o}$. A DNA molecule whose Lk value differs from the Lk_o is supercoiled. A quantitative measure of DNA supercoiling is the linking difference (**ΔLk**);

$$\mathbf{\Delta Lk = Lk - Lk_o}$$

When the value of ΔLk is negative, the corresponding DNA is negatively supercoiled. Which means that negatively supercoiled DNA is somewhat unwound compared with the relaxed DNA (**Figure 1.4.c**). When the ΔLk value is positive, DNA is positively supercoiled, and is somewhat overwound compared with the relaxed DNA.

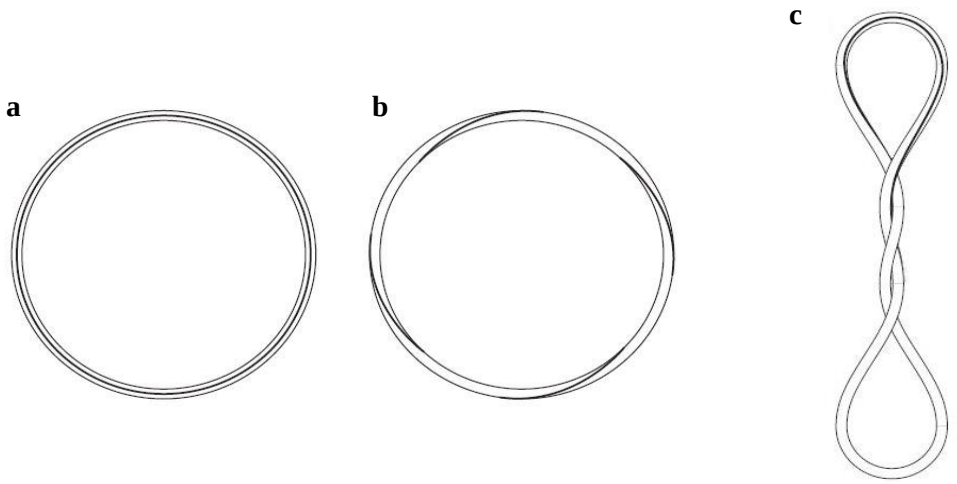


Figure 1.4. Twisting and writhing in the tubing model of DNA. a) 'Relaxed' DNA; Lk = Lk_o. b) The DNA is untwisted 4 times along its axis. By applying ΔLk = Tw+ Wr; ΔLk= −4, since ΔTw = −4, Wr = 0. c) DNA interwound by introducing four turns of writhe i.e. ΔLk = −4, since ΔTw ≈ 0, Wr ≈ −4 (Adapted from DNA Topology book Bates.A, Maxwell.A 2005).

Likewise, another useful parameter of supercoiled DNA is the "superhelical density (**σ**)", defined as:

$$\boldsymbol{\sigma} = \mathbf{\Delta Lk/\ Lk_o}$$

At times it is easier to use "σ", rather than ΔLk, to compare supercoiling between different DNAs. This is because "σ" is normalized for DNA length and it estimates the number of supercoils per helical turn of DNA. For circular DNAs isolated from living organisms the

absolute value of "σ" may vary from - 0.02 to - 0.09, i.e. there are 2–9 supercoils per 100 helical turns of DNA. (Calladine CR and Drew HR et al 1997)

2.1.2. DNA topoisomerases; the enzymes responsible for modulating DNA topology:

DNA topoisomerases are a class of ubiquitous enzymes which alter DNA supercoiling by catalyzing the winding and unwinding of DNA double helix. They do this by introducing a cut that breaks the DNA backbone, through which one of the DNA strands passes through another, thus relaxing or coiling the DNA before it re-seals the breaks (Champoux, J. J et al 2001). In the strand-breakage reaction by a DNA topoisomerase, the tyrosyl oxygen of the enzyme attacks the phosphorus backbone of DNA, causing a phosphotyrosine covalent linkage while breaking a DNA phosphodiester bond simultaneously. Resealing of the DNA strand is carried out by a second transesterification, which is basically the reverse of the first reaction. In this the oxygen of the DNA hydroxyl group attacks the phosphorus of the phosphotyrosine link, breaking the covalent bond between the protein and DNA, and thus re-forming the DNA backbone (Champoux, J. J et al 2001). By these series of reactions a transient enzyme-mediated gateway is created for the DNA strands to pass over one another to mediate release or gain of coiling.

DNA topoisomerases fall into two categories: type I and type II. These classes have further subfamilies which are structurally and mechanistically similar (**Table 1**).

Type I DNA Topoisomerases:

- Break one strand of a DNA helix.
- Are ATP independent (except reverse gyrase)
- Primarily relax positively supercoiling (over-wound) and/or negative supercoiling (under-wound) DNA, whereas reverse gyrase can introduce positive supercoils.
- Their mechanism requires either a transient nick allowing strand passage by 'enzyme bridging' (Type IA) or a transient cleavage leading to its 'strand rotation' (Type IB). In both mechanisms, one DNA strand passes the other and hence relaxes (unwinds) the supercoiling, followed by rejoining the ends of the broken strand (Figure 1.5.a)
- Play an important role in DNA replication and transcription (eukaryotic DNA topoisomerase I), and recombination (eukaryotic DNA topoisomerase III).

Type II DNA Topoisomerases:

- Break two strands of a DNA helix.
- Are ATP dependent.
- Are responsible for relaxing DNA (yeast topoisomerase II), as well as introducing negative (bacterial DNA gyrase) where both belong to Class IIA.
- Their mechanism of action involves passing an intact DNA helix through the gap made by the breaking both strands of a DNA helix and then resealing the strands (**Figure 1.6**).
- Play an important role in chromosome condensation (eukaryotic DNA topoisomerase II) and during segregation of daughter chromosomes in cell division (eukaryotic DNA topoisomerase II and prokaryotic DNA topoisomerase IV).

Subfamily	Representative members
IA	•Bacterial DNA topoisomerases I and III Yeast DNA topoisomerase III •*Drosophila melanogaster* DNA topoisomerases IIIα and IIIβ •Mammalian DNA topoisomerases IIIα and IIIβ
IB	•Eukaryotic DNA topoisomerase I •Mammalian mitochondrial DNA topoisomerase I
IIA	•Bacterial gyrase, DNA topoisomerase IV •Yeast DNA topoisomerase II •*Drosophila* DNA topoisomerase II •Mammalian DNA topoisomerases IIα and IIβ
IIB	•*Sulfolobus shibatae* DNA topoisomerase VI (subunit A homologous to yeast Spo11)

Table 1. Subfamilies of DNA topoisomerases. (Adapated from James C Wang et al 2002)

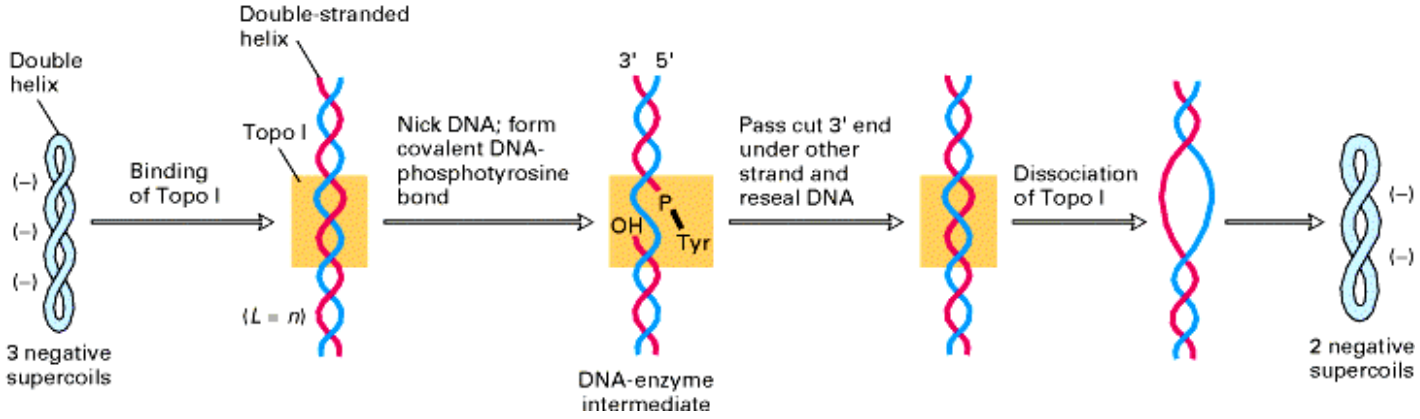


Figure 1.5. Mechanism of DNA topoisomerase I for solving supercoiling. A complex is formed between the catalytic tyrosine of topoisomerase I and a phosphate on a DNA strand. Topoisomerase I introduces a nick in one strand, allows the unwinding by strand rotation or passage to relax supercoiling and reseals the strands afterwards (Adapted from http://helicase.pbworks.com/w/page/17605720/Tyler-Huff)

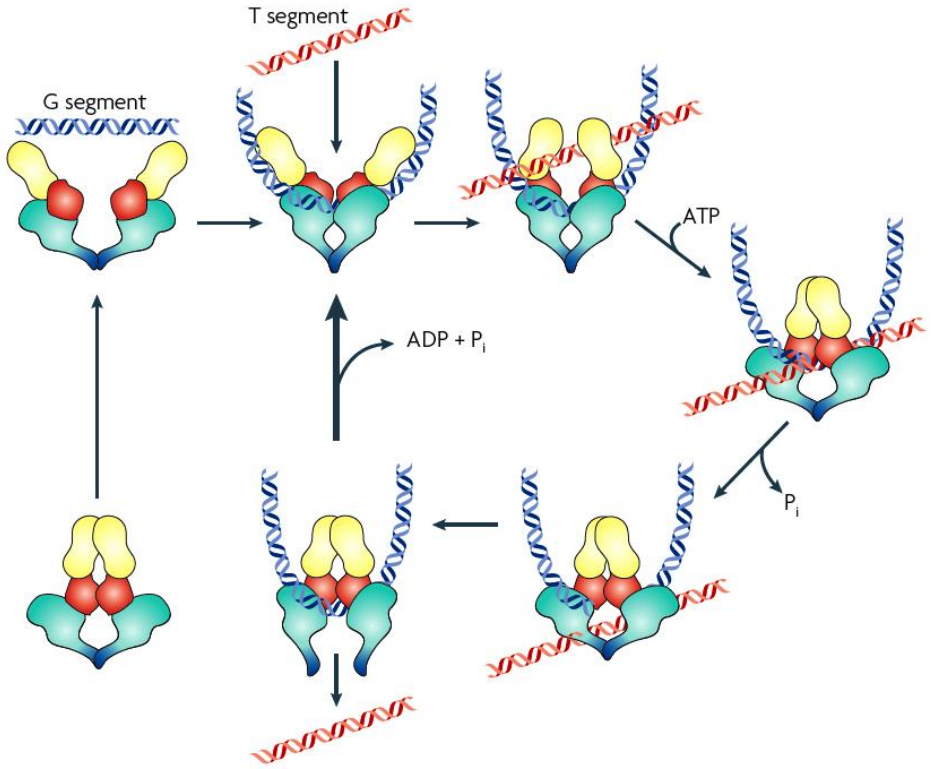


Figure 1.6. Generalized mechanism for the action of DNA topoisomerase II. The enzyme active site opens and binds the first segment of DNA designated the G (Gate) segment. It then binds a second segment, the T (Transported) segment. After cleavage of the G segment, the T segment is passed through the protein and released on the opposite side. A Mg^{2+} ion is required for cleavage of the G segment, which results in the formation of phosphotyrosine bonds between the cleaved ends of the DNA strand and active site tyrosine residues. ATP binding causes a conformational change that closes the enzyme active site. ATP hydrolysis may also assist in passage of the T segment (Adapted from J. L. Nitiss et al 2009)

2.1.3. Importance of DNA topology modulation during DNA replication:

Changes in DNA topology are imposed during the process of initiation, elongation, and termination of DNA replication, which are discussed below;

a) Replication initiation:

In all living beings DNA replication begins at specific segments known as origins of replication. To initiate replication the origin region needs to be unwound before the DNA synthesis could take place. Hence, it is intuitive to think that negative supercoiling is likely to favour the initiation of this process. In *E. coli*, DnaA when bound at the origins organizes into a nucleoprotein complex which is more stable if *oriC* is present on a negatively supercoiled DNA (Fuller, R.S et al 1983). On similar grounds, the importance of supercoiling in initiation is observed in the yeast autonomously replicating sequences (ARS), which in presence of negatively supercoiled DNA unwinds readily. *In vivo*, this ease of unwinding is correlated with their efficiency as replication origins (Umek, R.M et al 1988). Although topoisomerases may not play a direct role in replication initiation complexes but by maintaining the supercoiled state of the template DNA they play a crucial role on the efficiency of initiation. For example, in bacteria requirement of DNA gyrase at the initiation step of replication is thought to relate for capacity to introduce negative supercoils and not its specific role at the origins.

b) Replication Elongation:

During replication elongation the fork proceeds unwinding the DNA template and this elimination of crossings in the parental duplex is compensated by the accumulation of positive supercoiling ahead of the fork. This is because fork rotation might be partly hindered *in vivo*, due to size or it being attached to fixed structures within the cell, limiting the transfer of torsional stress to newly synthesized duplexes to form precatenane structures (discussed below). In *E. coli*, if unresolved, positive supercoiling might impede elongation. Thus, DNA gyrase is the topoisomerase responsible to avoid that by introducing negative supercoils that overall relax DNA overwinding. Although, it has been also pointed out that topoisomerase IV can also carry out a similar function, where DNA gyrase activity is limiting (Khodursky, A.B. et al 2000), indicating that the two enzymes may share overlapping roles. There are times when the elongating forks converge towards topological domain barrier (as mentioned above) and there is lesser room to accommodate the positive supercoils building up ahead of the forks. An outcome of this could be that the positive supercoils diffuse behind the fork by rotation and create interwinding in the daughter strands **(Figure 1.7)** (Postow, L. et al 2001). The formation of these interwindings is called ‘precatenanes’. The name comes from the fact that if unresolved before replication is complete they will form catenanes. Since the precatenanes may contain nicks or single stranded breaks, DNA topoisomerase III in *E.coli* could resolve them and support further replication fork progression (Hiasa H.et al 1994) (Nurse P.et al 2003). However, if till the

replication termination these precatenanes are unresolved, type II DNA topoisomerases can carry out their removal so as to avoid formation of catenated DNA molecules (discussed below) (Zechiedrich, E.L.et al 1995).

In *S. cerevisiae* and *S. pombe*, the elongation of DNA synthesis is affected substantially by the inactivation of both DNA topoisomerases I and II, but not by single inactivation of either (Kim, R.A. et al 1989). Both these enzymes can solve positive supercoiling, which implies that in eukaryotes, both DNA topoisomerase I and II can act to relieve the torsional stress generated during replication elongation. In agreement with this notion, both yeast Top1 and Top2 associate to replication forks (Bermejo et al 2007). Although, type IA enzyme i.e. yeast topoisomerase III does not seem to assist during elongation in yeast.

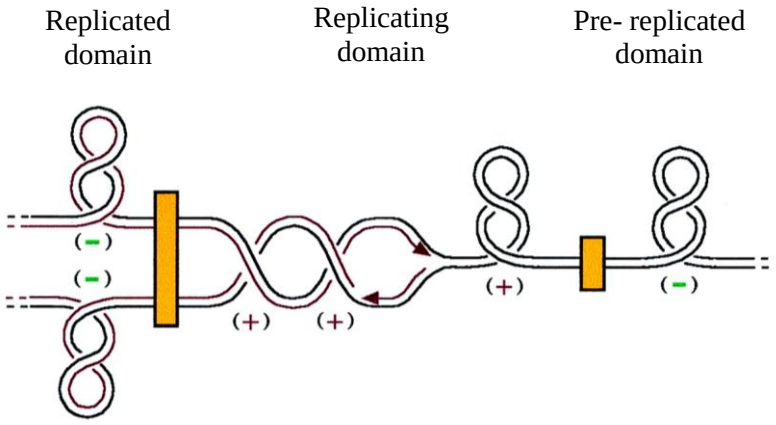


Figure 1.7. Model for topology of the replicating chromosome. A segment of chromosomal DNA is depicted, with black lines as parental strands and red lines as nascent strands. In the bacterial chromosome, domain barriers (yellow boxes) isolate the topology around the fork from the rest of the chromosome, which is (−) supercoiled by DNA gyrase, as shown in the domains on either side of the replication domain. Replication creates a (+) ΔLk in the replicating domain (*center*), which can cause (+) supercoils ahead of the fork and (+) precatenanes behind it. Thus, either gyrase or topo IV (or Top1 or Top2 in eukaryotes) could support replication by removing (+) supercoils in front of the fork, and topo IV (or Top2 in eukaryotes) could also support replication by removing precatenanes behind the replication fork (Adapted from Postow, L. et al 2001).

c) Replication termination:

During the termination of replication when the two replication forks converge towards each other, the unlinking of the parental DNA strands might be impeded by the accumulation of a

high density of supercoiling at the remaining unreplicated DNA strands **(Figure 1.8.a)**. In such circumstances, torsional stress is thought to be transferred behind converging forks to allow the last steps on DNA synthesis and leading to the accumulation of pre-catenated DNA (**Figure 1.8.b**). The formation of such structures has been documented as intermediates in replication in a number of systems (Wasserman, S.A. et al 1986). In principle, such pre-catenated DNA molecules may be resolved by both type IA or type II topoisomerases. Although, depending on catenated product so formed, the activity of the topoisomerases can be distinguished. As mentioned above, type IA topoisomerase III in *E.coli* or eukaryotic Top3 require a nick, or a single-stranded gap as a substrate to remove precatenanes (Nurse P.et al 2003). While type II enzymes have been well characterized to be perform the decatenation of daughter chromosomes, and thus contribute to proper segregation of the genetic material (Steck, T.R. et al 1984) (Zechiedrich, E.L.et al 1995).

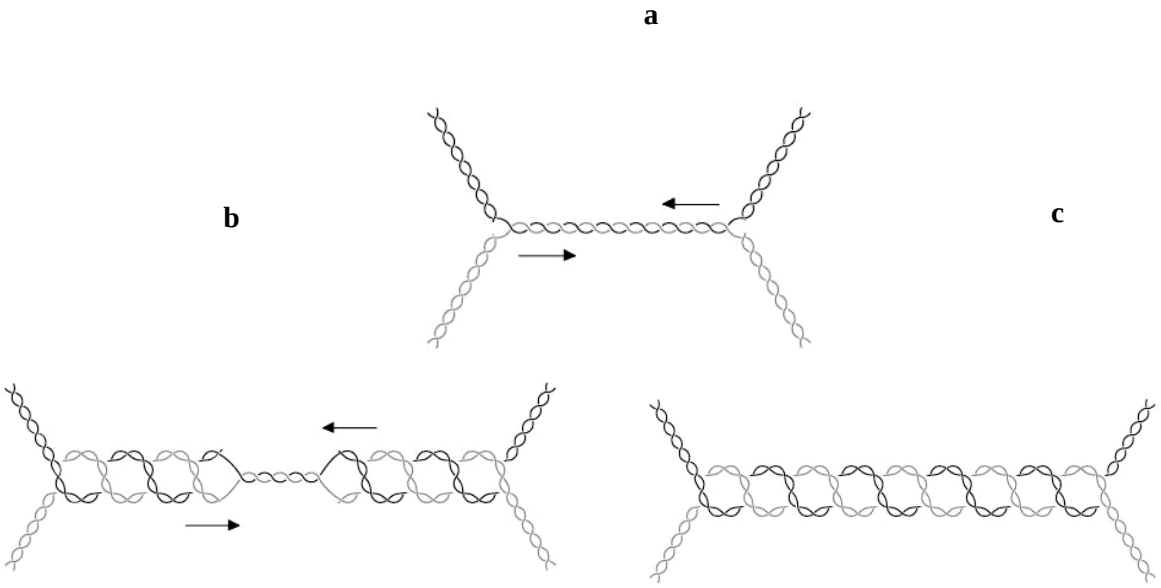


Figure 1.8. Formation of pre-catenated DNA at the termination of replication. a) At the terminus of replication, converging replication forks, leads to the interwinding of daughter molecules and the formation of pre-catenanes (b). c) Upon completion of replication, the products are catenated DNA which can be resolved by type II topoisomerases (Adapted from DNA Topology book Bates.A, Maxwell.A 2005)

In *S. cerevisiae* it has been shown the catenanes accumulate in closed circular plasmids upon inactivation of DNA topoisomerase II (DiNardo, S. et al 1984), signifying that the enzyme is necessary for the segregation of chromosomes at the termination of replication. Also in *S. pombe* it has been elucidated that this enzyme is required during mitosis for proper chromosomal segregation (Yanagida, M. et al 1987).

In summary, the of DNA replication is linked to changes in DNA topology that if not acted upon by DNA topoisomerases can lead to stalling of fork progression and counteract chromosome segregation, giving rise to chromosomal breaks and genomic instability.

2.2. Natural Impediments to DNA Replication:

Natural impediments responsible for fork stalling can either be "programmed" or occur "accidentally". The first case is exemplified by termination of replication in bacteria, halt of replication in eukaryotes by ribosomal operons, pausing sites at termination regions or replication slow zones and mating-type locus in fission yeast (Ekaterina V. Mirkin et al 2007). The pausing sites for forks may also include unusual DNA secondary structures, certain non-histone protein-DNA complexes and collisions with transcription. On the basis of the extent of hindrance caused to the replication forks the above-mentioned sites can be broadly classified in two classes. "Replication fork barriers" are those where replication stalling is programmed to control DNA progression for protecting genome integrity. While the second class, forks are usually capable of bypassing the obstacle after temporarily pausing, hence given the name "Replication fork pausing sites". The pausing event can last from several seconds, like in yeast centromeres, to almost half an hour, like at sites of EBNA-1 (Epstein-Barr nuclear antigen 1) protein binding in Epstein-Barr virus (EBV) DNA replication.

In synchrony to my work, I discuss further the best-characterized sites in the budding yeast model known to form natural obstacles for the approaching replication forks. A class of such impediments, i.e. formation of unusual DNA secondary structures will be dealt with in detail as Chapter 2 of this introduction.

2.2.1. DNA Binding Proteins:

a) Eukaryotic ribosomal barriers

The eukaryotic rRNA Replication fork barrier (RFB) was first discovered in *Saccharomyces cerevisiae*, situated near the 3' end of the 35S ribosomal RNA gene (rDNA) (Brewer, B. J., and W. L. Fangman. 1988) (Linskens, M. H., and J. A. Huberman. 1988). The budding yeast rRNA locus is an arrangement of tandem repeats of the 35S RNA gene (precursor for 18S, 5.8S, and 25S rRNAs), which is transcribed by RNA polymerase I. Next to it is a first non-transcribed spacer (NTS1). The 5S rRNA gene, transcribed by the RNA polymerase III, is followed by a second non-transcribed spacer (NTS2). The 5S rRNA includes an origin that fires bi-directionally. Budding yeast contains sites with multiple arrays of rDNA repeats, e.g. 150 rDNA repeats in chromosome XII, which are highly transcribed. Thus, it would be exceedingly common to have head-on collisions between the replication forks and transcription bubbles. One

way to avoid such collisions is to impose a directionality of the replisome with respect to the transcription progression. This is mediated by rDNA replication Fork Barriers (RFBs), which favor replication & transcription to proceed co-directionally through the rDNA loci (Linskens, M.H et al 1988)

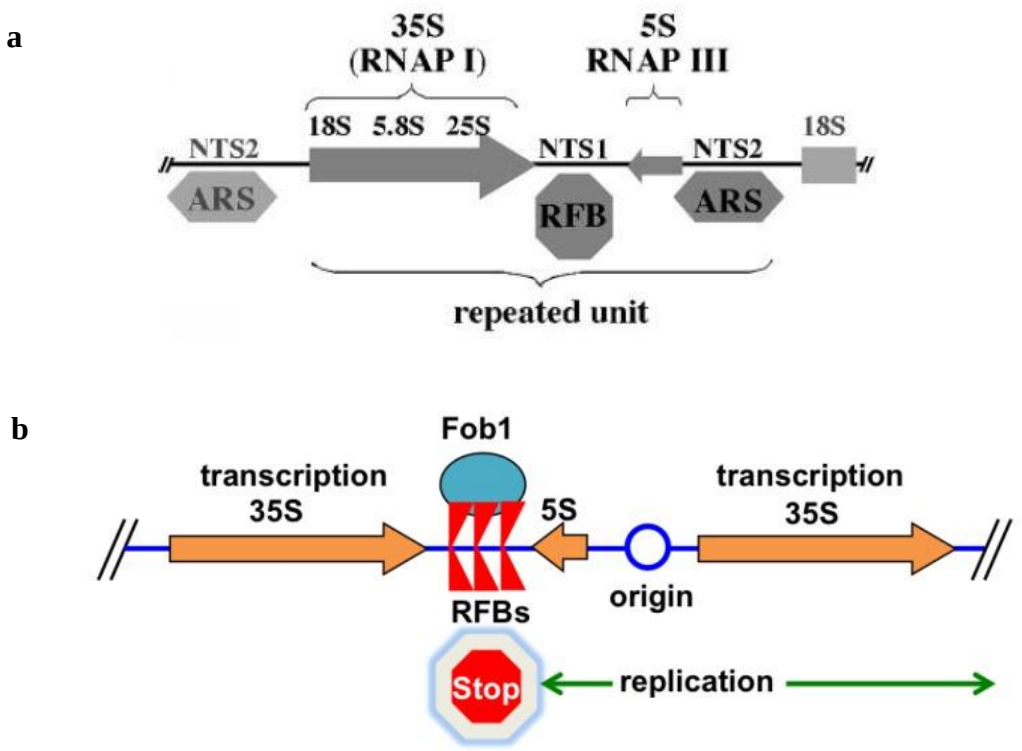


Figure 1.9. Organization of rRNA repeats in *S. cerevisiae*. a) The locus consists of multiple units of direct repeats, each of which contains a 35S RNA transcribed by RNA Pol I, NTS1, 5S transcribed by RNA Pol III and NTS2 (adapted from Ekaterina V. Mirkin et al 2007). b) Fob1 binds RFB sequences to block replication forks progressing in the opposing direction to 35S ribosome gene transcription. Directionality of replication is shown by green arrow. 'Stop' sign signifies the barrier to the fork movement from the 3' direction (adapted from Adam R. Leman et al 2013)

A polar block of replication is induced by the protein Fob1 (**Figure 1.9.b**). Fob1 binds RFBs located at the 3' non-transcribed region of each rDNA gene (Kobayashi, T et al 1996). On binding, it wraps the DNA around itself to significantly modify the local state of chromatin and thus block replication from the 3' direction (i.e. forks progressing co-directionally to 35S transcription) (Kobayashi, T et al 2003). Not only does Fob1 binding induce a polar replication fork pausing but it also causes increased recombination events at the RFB regions (Johzuka, K et al 2002). Two proteins that are known to act to stabilize stalled replication forks, Tof1 and Csm3, have been shown to protect forks arrested at the Fob1-mediated barrier too. Their role in combination with Fob1 is to counteract the action of Rrm3, a DNA helicase essential for

replication through the rDNA RFB region and other natural pausing sites. It has been postulated that Rrm3 removes Fob1 to facilitate fork progression (Mohanty, B.K et al 2006) Since, Rrm3 was also found to physically interact with the replication fork clamp loader PCNA (Schmidt, K.H et al 2002) (Azvolinsky, A et al 2006) it might favor the polar passage of replication forks through fork block sites throughout the genome. Thus, it can be said that the ribosomal barriers have possible role in protecting the rRNA genes from the head-on transcription-replication collisions.

b) Non-histone protein-DNA complexes in budding yeast

Studies have shown that during chromosomal replication in the *S. cerevisiae*, fork stalling takes place at many genomic loci where non-histone proteins were bound to DNA besides the ribosomal barrier. The first insight of such inhibition of replication by non-histone bound protein-DNA complexes comes fro the observation that replication forks halt at centromeres in *S. cerevisiae* (Greenfeder, S et al 1992). Even on altering the directionality of replication through the centromeric regions, they found that the pausing was independent of its orientation. With yeast plasmids, they later confirmed that the pause sites strongly correlated centromeric DNA binding to non-histone proteins where they are tightly packed. On comparison of the amount of stalled versus normal replication intermediates they estimated that the time of pausing was 0.1 to 0.2 min.

Since Rrm3 helicase was found to be required at rRNA locus in yeast, studies Rrm3-deficient yeast strains identified 1,400 discrete sites of replication pause sites throughout the genome (Ivessa, A.et al 2003) (Ivessa, A. et al 2000). The sites included tRNA genes, centromeres, telomeres, silent mating-type loci, and inactive origins of replication as well as the rRNA locus. All these genomic elements are tightly bound by non-histone protein-DNA complexes. Since stalling of replication forks at the same loci in Rrm3-deficient strains compared to wild-type strains greatly increased, a role for Rrm3 in removing such tightly bound protein from the DNA to assist replication progression was proposed. The need for at least one helicase dedicated to solve these tightly bound protein DNA complexes to avoid fork stalling and breaks formation, indicates the implications of such protein complexes on the replication mechanism on a genome wide scale (Ivessa, A. et al 2000).

2.2.2. Replication-Transcription collisions:

In order to understand the genome wide outcomes of replication transcription interference, a brief insight on the gene transcription machinery, its role in synthesis, processing and export of mRNA is discussed as follows.

a) Gene Transcription:

It is the elaborate process that eukaryotic cells use to duplicate genetic information stored in DNA into units of RNA replica. Unlike prokaryotes, where a single RNA polymerase initiates the transcription of all different types of RNA, eukaryotes employ three variants of RNA polymerases that transcribe different sets of genes. For eukaryotic mRNA synthesis the three polymerases are composed of 10 subunits or more, where each requires a specific set of transcription factors to bring it to the DNA template. The three different types of RNA polymerases are:

RNA polymerase I is situated within the nucleolus, where the ribosomal RNA (rRNA) is transcribed, processed, and assembled to form ribosomes. All the large as well as small subunit RNAs are synthesized by RNA polymerase I. It is considered the most abundantly present polymerase within the cell and also the fastest in synthesizing RNA in comparison to any other polymerases. RNA polymerase I transcribes all of the rRNAs except for the 5S rRNA molecule. The rRNAs are components of the ribosome and are considered structural RNAs as they have a cellular role but are not translated into protein. Each transcript comprises of a copy of three rRNAs, i.e. the 28S and 5.8S large subunit RNAs and 18S RNA the small subunit. The transcription of rRNA is very efficient owing to the fact that each rRNA can only transcribe one ribosome, whereas a single mRNA could make multiple proteins. The primary transcript consists of both the small and large subunits i.e. 28S-5.8S-18S, in that order. The processing of rRNA requires the modification of the specific nucleotides along with the cleavage of the transcripts into the individual RNA components (**Figure 1.10**).

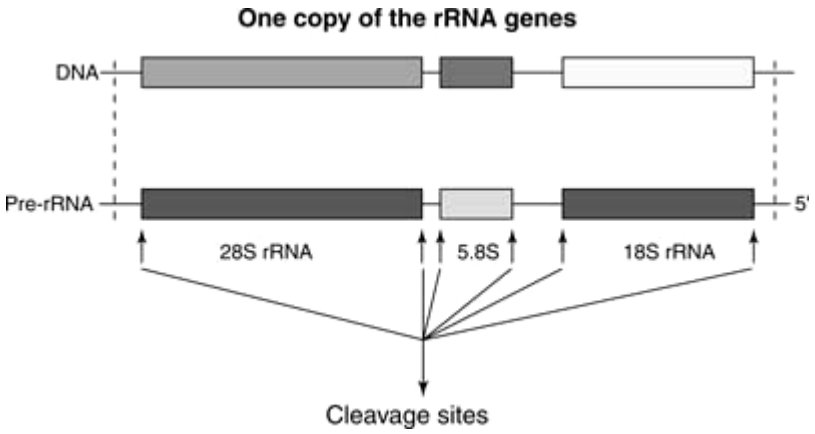


Figure 1.10. The rRNA gene transcript. Single copy of rRNA gene is shown which on being transcribed results in 28S, 5.8S, 18S rRNA subunits. The pre-rRNA then needs to be further cleaved to later form the individual RNA components (adapted from https://www.cliffsnotes.com/study-guides/biology/biochemistry-ii/eukaryotic-genes/eukaryotic-transcription)

RNA polymerase II is located within the nucleus and transcribes messenger RNA and a few other small cellular RNAs. Class II promoters can be defined by their extreme sensitivity to α-

amanitin (a mushroom poison; which imparts sensitivity to the three polymerases differently). Like in prokaryotes, class II promoters contain two conserved sequences, namely the CAAT and TATA boxes. The TATA box is found 25-30 base pairs upstream of the transcription start site, i.e. located at approximately -30 relative to the initiation (+1) site (**Figure 1.11.a**). The TATA box is the core promoter element which is the binding site for a transcription factor known as TATA-binding protein (TBP). The TBP itself is a subunit of the Transcription Factor II D (TFIID). When TFIID binds with the TATA box via TBP, five more transcription factors and RNA polymerase II from a complex in a stepwise manner around the TATA box called the pre-initiation complex (**Figure 1.11.a**). Another transcription factor called the Transcription Factor II H (TFIIH) is responsible for separating opposing strands of double-helical DNA and provides access to the RNA polymerase to use single-stranded DNA as a template.

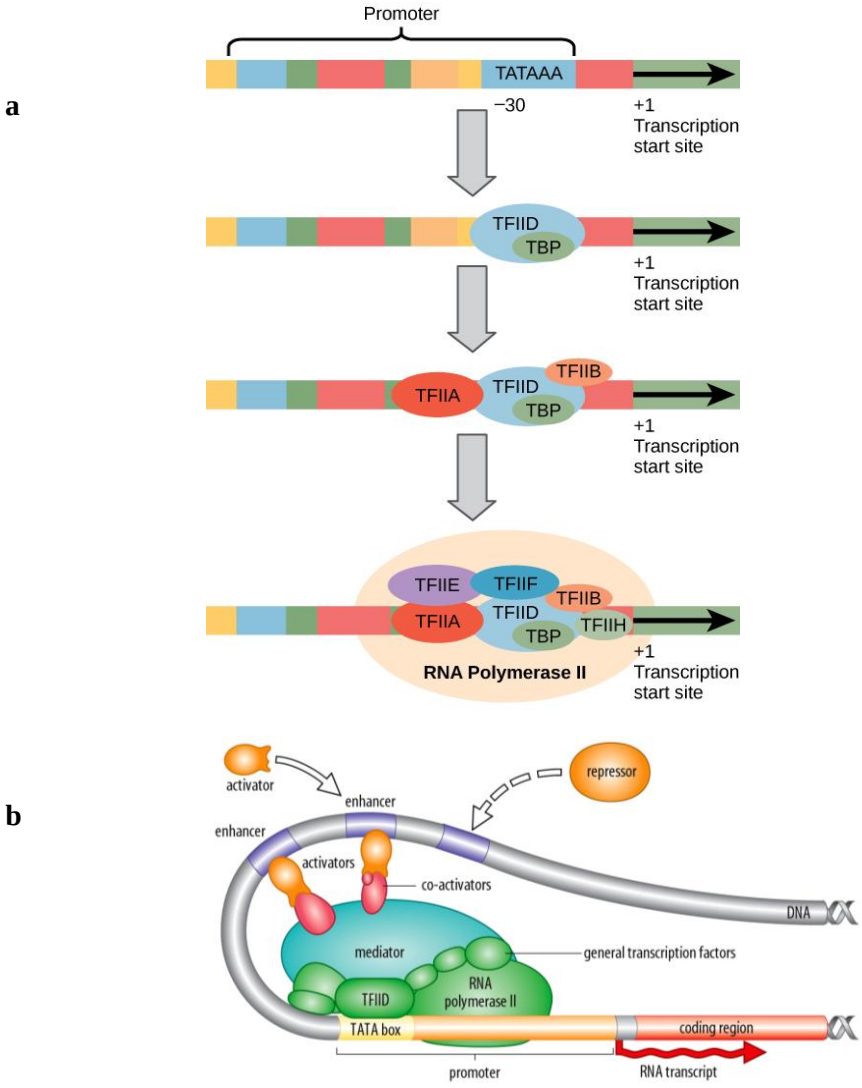


Figure 1.11. Schematic representation of a gene transcription regulation by RNA polymerase II. a) Step wise depiction of transcription factors that recognize the promoter. RNA polymerase II then binds and forms the transcription initiation complex (adapted from https://www.boundless.com/biology) b)

Transcription factors; activators, repressors and enhancers loading to propel the elongation by the RNA polymerase II (adapted from http://www.mun.ca/biology/desmid/brian/BIOL3530/DEVO_10/devo_10.html)

However, the pre-initiation complex alone drives only a basal or low level of transcription. Other proteins known as the 'activators' and 'repressors', along with other associated co-activators or co-repressors, are responsible for determining the transcription rate (**Figure 1.11.b**). The activator proteins increase the transcription rate whereas the repressor proteins decrease the transcription rate. While activators can interact directly or indirectly with the core machinery of transcription through enhancer binding, repressors need recruitment of co-repressor complexes causing transcriptional repression by chromatin condensation of enhancer regions (discussed below). Situations may occur when repressors may function by allosteric competition against a determined activator to repress gene expression. Such DNA-binding motifs having affinity for both activators and repressors may induce a physical competition for the binding site occupancy. If the repressor has a higher affinity than the activator, transcription would effectively be blocked by the repressor.

The binding sites for both these transcription factors are regulated by cis-regulatory elements called 'enhancers' **(Figure 1.11.b)**. They range from 200 bp to 1 kb in length and can be present at proximal 5' upstream to the promoter or within the first intron of the gene regulated, or at intergenic regions far away from the locus. By DNA looping the active enhancer a bound regulatory protein activates transcription by direct protein-protein interaction with RNA polymerase II at the promoter region. (Levine, M et al 2010) This promoter-enhancer complex forms the basis for interaction between transcription factors and RNA polymerase II to trigger its movement and escape from the promoter.

RNA polymerase III governs the transcription of 5S rRNA, tRNA and small nuclear pre-RNAs. The unique feature of its promoter is that it can be located inside the coding region of the gene itself, unlike to all other promoters. Pol III-transcribed genes like tRNA have internal promoters where the key elements are located within the transcribed regions which are identified by transcription factor IIIC (TFIIIC) (Schramm L.et al 2002). TFIIIC recruits TFIIIB, which is composed of the subunits BRF1, BDP1 and TATA box binding protein (TBP) (Schramm L.et al 2000). The TFIIIB then is responsible for the recruitment of RNA Pol III and initiates transcription.

Transcription elongation requires the transcription machinery to move histones out every time it encounters a nucleosome. This is facilitated by the protein dimer called FACT i.e. 'facilitates chromatin transcription'. FACT disassembles the nucleosome ahead of a transcribing RNA polymerase II by removing two of the eight histones i.e. a single dimer of H2A and H2B histones (Kireeva, M. L. et al 2002) (**Figure 1.12**). Ubiquitylation of H2B-K123 leads to the removal by

the FACT histone chaperone complex (Pavri, R. et al 2006). This presumably unwinds the DNA around that nucleosome such that the RNA polymerase II can transcribe through it. Interestingly, it has been shown that histone chaperone Nap1 in yeast also stabilizes a hexameric histone i.e. the H3–H4 tetramer and one H2A–H2B dimer complex (Kuryan, B. G et al 2012) **(Figure 1.12)**. Soon after, FACT reassembles the nucleosome behind RNA polymerase II by rearranging the missing histones to it.

RNA polymerase II in total consists of a 12-protein subunit complex. In general, all RNA polymerases travel along the DNA template in the 3' to 5' direction and synthesize new RNA strands in the 5' to 3' direction by adding new nucleotides to the 3' end the RNA strand. During the unwinding of the DNA helix, RNA strand synthesis takes place in a transcription bubble of about 25 unwound DNA base pairs. Out of which only 8 nucleotides of newly-synthesized RNA remain base paired to the template DNA. The rest of the RNA molecule remains unattached to the template to allow rewinding of the DNA template.

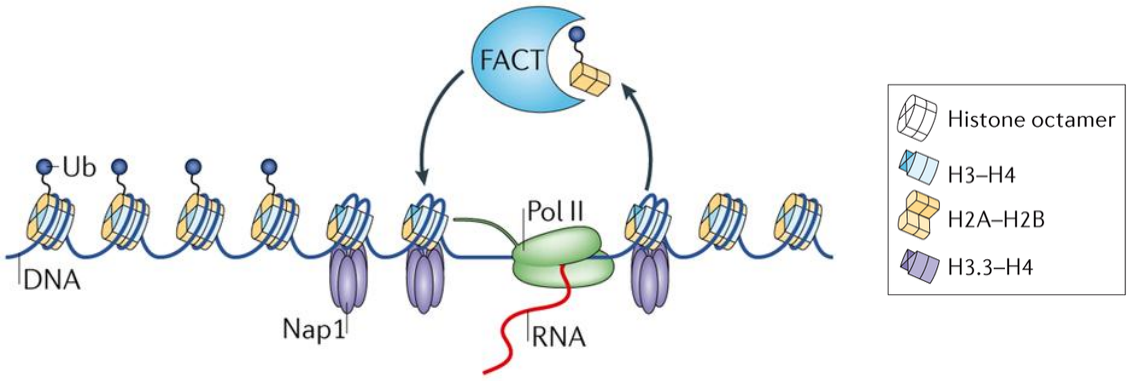


Figure 1.12. The FACT complex enables RNA polymerase II transcription of nucleosome-packed DNA. Passage of the elongating RNA polymerase II is aided by the FACT (in blue) histone chaperone complex, which targets ubiquitylated (Ub) H2A–H2B dimmers (in orange) for removal. The resultant hexameric nucleosome is stabilized by the histone chaperone Nap1 (in dark violet) in yeast. The box on side depicts the structural representation of the histone subunits (adapted from S.Venkatesh et al 2015)

The termination step is dependent on the type of eukaryotic RNA polymerase.

The rDNA gene bears a specific sequence (11 bp in humans and 18 bp in mice) that is recognized by a termination protein called TTF-1 (Transcription Termination Factor for RNA polymerase I). On binding at its recognition sequence, TTF-1 blocks further transcription. Due to which the RNA polymerase I disengages from the template DNA strand and thus releases its newly-synthesized RNA (El Hage et al. 2008).

RNA polymerase II which is responsible for transcribing protein-encoding, structural RNA and regulatory RNA genes lacks any specific signals or sequence that directs it to termination. RNA polymerase II has the capacity to transcribe RNA anywhere from a few base pairs to thousands past the actual end of the gene. However, the transcript is cleaved at an internal site before RNA polymerase II finishes transcription. Due to cleavage, the remainder of the transcript becomes open for digestion by a 5' exonulcease called Xrn2 (in humans), while the 3' end of the transcript is still being transcribed. When the 5'-exonulease activity encounters the RNA polymerase II, it digests away all the overhanging RNA and hence disengages the polymerase from the template DNA strand. Thus, causing termination of that round of transcription (Hannah E. Mischo.et al 2013).

In protein-coding genes, the cleavage sites which are responsible for the termination signal for the emerging pre-mRNA contain an upstream AAUAAA sequence and a downstream GU-rich sequence. These two specific sequences are separated by about 40-60 nucleotides in the emerging RNA. Once both of the sequences are transcribed, a protein called CPSF (in humans) binds the AAUAAA sequence while a protein called CstF (in humans) binds the GU-rich sequence. On binding these sequences the two proteins form a protein complex that leads to the cleavage of the nascent pre-mRNA by CPSF at a site 10-30 nucleotides downstream of the AAUAA site. The Poly (A) polymerase enzyme (PAP), which catalyzes the addition of a 3' poly-A tail on the pre-mRNA, is also part of the complex that formed with CPSF and CstF.

RNA polymerase III that transcribes the tRNA, 5S rRNA, and structural RNAs genes has a less deducible termination signal. The RNAs transcribed by RNA polymerase III contains short stretches of four to seven uracils at their 3' end, which somehow facilitate RNA polymerase III to both release the nascent RNA and disengage from the template DNA strand.

mRNA processing is the phenomenon where the pre-mRNA undergoes extensive processing before it is ready to be translated. Post synthesis the mRNAs undergoes extensive modifications after synthesis, which includes capping, polyadenylation, and splicing. **5' Capping** of the pre-mRNA takes place while it is still being synthesized. A 7-methylguanosine cap is added to the 5' end of the growing transcript by a 5'-to-5' phosphate linkage. This moiety prevents the nascent mRNA strand from being degraded. Besides, the 5' capping also helps initiation factors involved in protein synthesis to be recognized for translation by the ribosomes (**Figure 1.13**). A **3' Poly-A tail** is attached to the pre-mRNA while the RNA polymerase II is still transcribing downstream of the end of a gene. The pre-mRNA is first cleaved by an endonuclease-containing protein complex, which cuts between the AAUAAA consensus sequence and a GU-rich sequence. This cleavage then detaches the functional pre-mRNA from the rest of the transcript that is still attached to the RNA polymerase II. Another catalytic enzyme called poly (A) polymerase (PAP), which is also a part of the same protein complex that cleaves the pre-mRNA, immediately adds a chain of approximately 200 'A' nucleotides, called the poly (A) tail, towards the 3' end of the cleaved pre-mRNA. The poly (A) tail so added protects the mRNA from degradation and facilitates the export of the mature mRNA to the cytoplasm. Further it is also required for

binding of proteins involved in initiating translation (**Figure 1.13**). **Pre-mRNA splicing** is the process by which the non-coding region of the RNA transcripts are removed or spliced out in order to produce a mature mRNA sequence. The eukaryotic gene pool comprises of protein-coding sequences called exons and the intervening sequences (non-coding) called introns. Intron sequences within the mRNA do not encode any functional proteins.

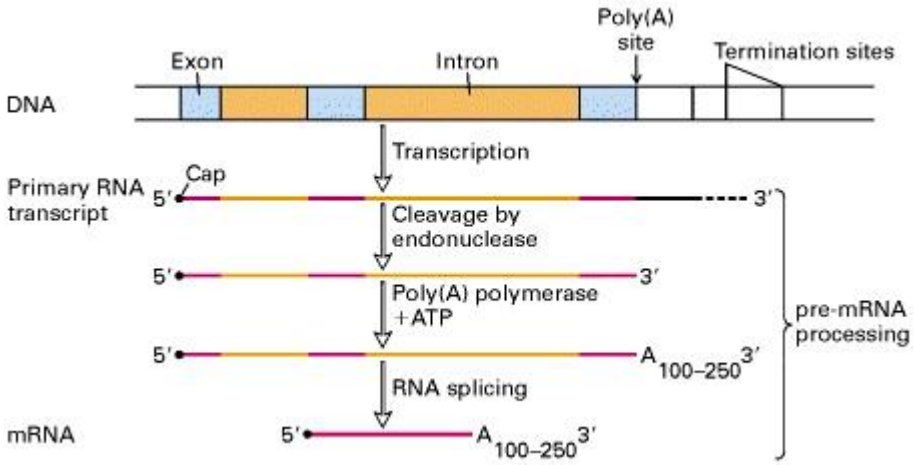


Figure 1.13. Stepwise representation of stages in mRNA maturation. The DNA as a template is transcribed by RNA oolymerase II to yield pre-mRNA. The mRNA is then cleaved by specific endonuclease to allow termination of transcription. Sequentially, 5' capping and poly (A) tail is addition at the 3' end (100-250) takes place. By RNA splicing the introns are removed and hence mature mRNA is formed.

.

b) mRNP biogenesis and export:

After the pre-mRNA has been transcribed from DNA in the nucleus, it is processed into a mature ribonucleoprotein (mRNP) particle, which is then exported from the nucleus to the cytoplasm. The THO complex, a nuclear protein complex conserved well in eukaryotes, is required in the biogenesis of mRNP particles and works as the interface between transcription and RNA export. THO is composed of Tho2 (180 kDa), Hpr1 (90 kDa), Mft1 (45 kDa), and Thp2 (30 kDa) proteins (Chávez S et al 2000). Null mutations in all THO components lead to similar phenotypes of transcription impairment, hyper-recombination and defective RNA export, which establishes that THO is a functional and physical complex. Although, for efficient co-transcriptional mRNA export it is essential to recruit two more factors i.e. Yra1 and Sub2 (Strässer K et al 2002) (**Figure 1.14**) where the latter is an RNA-dependent ATPase involved in mRNA export. These proteins when together form a large complex called TREX i.e. transcription-export complex. Sub2 was found to be a high copy suppressor of Hpr1 deletion, it indicating a connection between THO and RNA export. Disrupting any of the tightly coupled

steps for mRNA export to the cytoplasm causes the activation of the RNA surveillance pathway, which subsequently leads to degradation of non-active mRNA molecules (Houseley J et al 2006). Similarly, perturbing THO subunits in yeast also leads to impairment of mRNP formation, which in turn causes defects in transcription elongation and export. Disruption of Hpr1 subunit causes mRNA entrapment at the transcription sites giving rise to the formation of RNA-DNA hybrids within R-loop structures. The R-loop formation results in hyper-recombination and genomic instability if these molecules remain unresolved (Huertas P et al 2003). Lack of the Mtf1 subunit also results in accumulation of transcribed but not matured and exported RNA, which, along with transcriptionally active chromatin, pieces of RNA export machinery and nuclear pore complexes (NPC), forms large aggregates called heavy chromatin (Rougemaille M et al 2006). Several other proteins have also been described to interact with the THO such as Mex67, Gbp2 and Hrb1 (the serine-arginine rich like proteins) and the Prp19 complex, with is involved in splicing and transcription elongation.

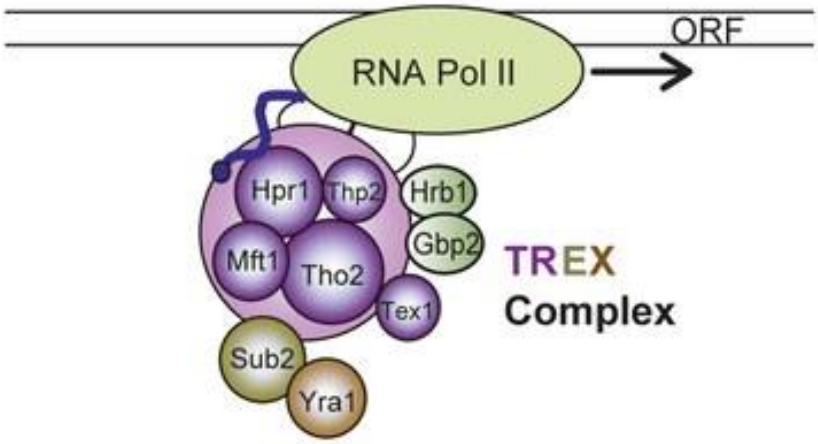


Figure 1.14. The THO/ TREX complex. The THO complex includes the Hpr1, Tho2, Mft1 and Thp2 (violet color). This along with Sub2 and Yra1 is called a TREX complex. The RNA Polymerase II (grey color) is transcribing and forming the mRNA (blue color line). Gbp2 and Hrb1 are also shown in grey. The whole mRNA with protein bound complex is called mRNP. (Adapted from https://www.uni-giessen.de/fbz/fb08/Inst/biochem/straesser/research/research-straesser)

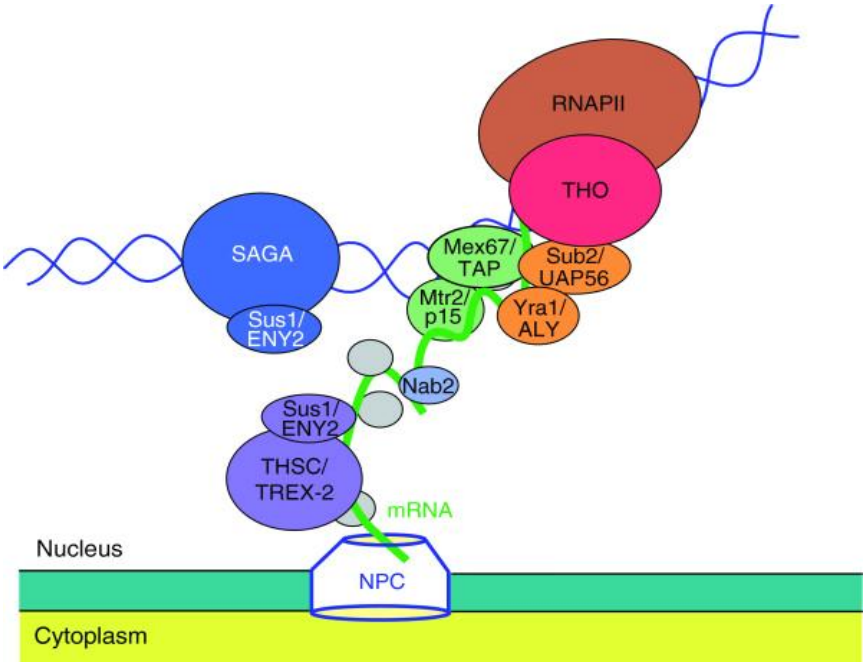


Figure 1.15. THO complex role in mRNP biogenesis at the interface between transcription and export of mRNA from the nucleus to cytoplasm. Proteins shown with their yeast name / human homologs respectively as per Nab2, Mtr2/p15, Sub2/UAP56, Mex67/TAP and Yra1/ALY. While the yeast name / *Drosophila* homologs as per Sus1/ENY2. Protein complexes are shown in capital letters: THO, THSC (or TREX-2) and SAGA. Proteins that interact with each other or between which a physical connection has been reported are in the same color. Sus1 can act as a subunit of both THSC and SAGA complexes. Unlabeled proteins in gray represent other factors important for mRNP biogenesis and export. NPC, nuclear pore complex; RNAPII, RNA polymerase II. (adapted from Sonia Jimeno et al 2010)

In yeast, the THO complex could be one of the first players to act during transcription elongation to facilitate a correct mRNP formation. It paves the path for recruitment of other factors such as Sub2 or Mex67 (Zenklusen D et al 2002). Yra1, another RNA binding protein which interacts with Sub2 and the Mex67-Mtr2 export factor, can also favor the mRNP to the nuclear pore complex. THO also aids in recruiting Mex67 to the mRNP through Hpr1 (**Figure 1.15**). This interaction is regulated by Rsp5, an ubiquitin ligase which polyubiquitinates Hpr1 protein (Hobeika M et al 2009). Close to the nuclear pore complex another complex called the THSC complex (also termed TREX-2) may have an unknown role in mRNP biogenesis and export. It is intriguing that mutations in THSC also result in transcription elongation impairment, faulty RNA export and even transcription related hyper-recombination as seen by mutating THO complex. (Aguilera A et al 2005).

Human THO also associates with proteins of the spliceosome (Masuda S et al 2005). Likewise, there is also indication for recruitment of THO in a transcription dependent manner to the chromatin in Drosophila (Kopytova DV et al 2010). Therefore, role of THO in mRNP metabolism may be significant among eukaryotes. Recent evidence that Drosophila THO complex interacts with ENY2, a transcriptional activator that interacts with the SAGA

transcription factor, opens new opportunities of a co-transcriptional action of THO in higher eukaryotes (**Figure 1.15**) (Kopytova DV et al 2010). However, the impact of THO complex may go beyond transcription elongation and RNA metabolism, as any THO protein modifications or functional role might uncover its pattern of activity during the course of its further understanding.

c) Replication transcription interference:

Given that the replication fork and the RNA polymerase share the same DNA template, situations may arise where collisions between the two machineries are inevitable (Brewer, B. J et al 1988) (Nomura, M. et al 1977). As both the processes are polar, they can collide either 'head-on' or 'co-directionally' (**Figure. 1.16**). A head-on collision occurs when a given gene is being transcribed by using the lagging strand as a template, while in co-directional case the gene transcription is taking place using the leading strand as a template. Since, in prokaryotes replication is approximately 20-fold faster than transcription i.e. ~800 nucleotides nt/sec versus 20 to 50 nt/sec (Kornberg, A et al 1992), in both the head-on and in the co-directional scenarios collision would be relevant. Contrarily, in eukaryotes replication and transcription machineries progress with comparable speeds of 17–33 nt/sec (Hiratani, I et al 2008) and 17–72 nt/sec respectively (Darzacq. X et al 2007) (Singh, J et al 2009), making co-directional collisions less likely.

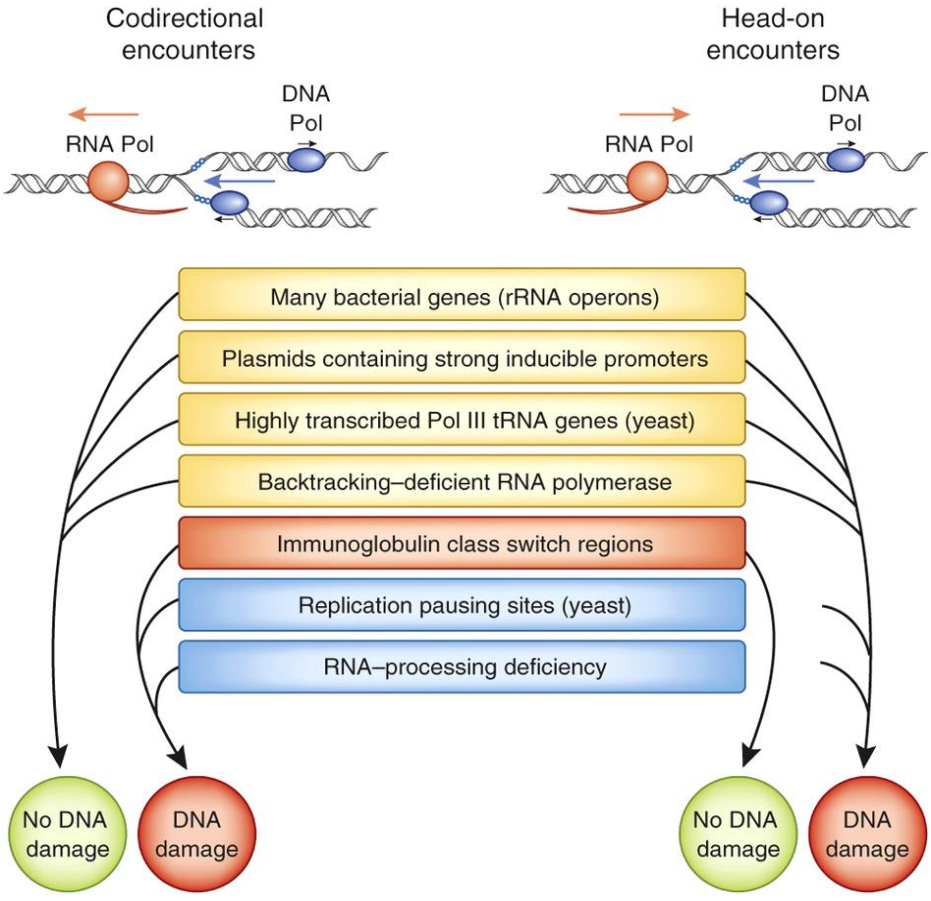


Figure 1.16. Directionality of co-directional and head on collision encounters of replication and transcription machineries. On the basis of genomic location and the state of the transcription complex, the outcomes of transcription-replication encounters can be classified into three main groups: (i) those linked to DNA damage when the transcription machinery encounters the replication fork in the head-on direction (yellow boxes), (ii) those associated with DNA damage when transcription and replication are co-directional (orange box) and (iii) those that may cause genomic instability in either orientation (blue boxes). Orange and blue arrows represent the direction of transcription and replication, fork progression, respectively; small black arrows represent the direction of the DNA polymerases replicating each strand. Pol stands for polymerase. (Adapted from Anne Helmrich et al 2013)

Likewise, initially it was thought that head-on collisions between transcription and replication machineries were more detrimental than the co-directional ones. But lately, a number of studies suggested that it might not hold true always. Given that DNA damage can be caused by head-on or co-directional encounters or even upon collisions in both directions, a summarization of all the scenarios has been postulated as shown in the **Figure 1.16**.

However in yeast, the directionality of replication and transcription processes does not matter for replication fork progression on a genome-wide scale. It has been argued that naturally occurring sites of replication-fork pausing at transcribed regions are not dependent on polarity (Azvolinsky, A et al 2009). Similarly, in RNA-processing defective yeast strains the sites of DNA damage are located to ORFs irrespective of orientation with respect to the closest replication origin (Stirling, P.C et al 2012). Thus, replication forks may meet a transcription-elongation complex from either direction i.e. both co-directional and head-on collisions have a probability to cause DNA breakage, at least when transcription is hampered (**Figure. 1.16**).

d) Genome instability related to replication transcription collisions:

In prokaryotes and eukaryotes, the DSBs formation may be an outcome of two mutually exclusive events of transcription-replication interference i.e. R loops formation and topological stresses. The RNA-DNA hybrid structures displacing a single stranded DNA segment are known as R loops (**Figure. 1.17**). They form when nascent RNA hybridizes with one of the strands in the DNA helix, displacing the other DNA strand and thus forming a loop. The enzyme specific for the removal of R loops by the hydrolytic cleavage is RNase H in prokaryotes and RNase H1 in eukaryotes. This enzyme is found to be necessary to avoid genome instability (Helmrich, A. et al 2011) (Lin, Y et al 2010). Similarly, topological constraints develop when a replication fork comes across a transcribed chromosomal region that cannot rotate freely, due to for example attachment to the nuclear pore (Bermejo, R. et al 2011). If the DNA remains attached to the nuclear pore, the positive supercoils accumulated in front of forks would lead to either fork collapse or reversed forks accumulation. Therefore, replication-transcription clashes cause two

possible outcomes potentially leading to the formation of DNA-damage. However, it is still not clear whether R-loops and topological constraints are independent events or outcomes of a linked mechanism.

RNA-DNA Hybrids:

According to the 'thread-back' model (Westover, K.D et al 2004), the newly synthesized RNA transcript, as is complementary to one of the two DNA strands, can invade and anneal to form a three-stranded nucleic acid molecule containing a RNA-DNA duplex along with a single-stranded DNA (Reaban, M.E et al 1994) (**Figure.1.17**). It is postulated that since during mRNA synthesis there is a transient uncoiling of the DNA duplex, creating negatively supercoiled (underwound) DNA behind the RNA Polymerase II, the newly synthesized mRNA can thus invade and anneal the template to form R-loop structures (Drolet, M et al 2006). Although, R-loops occur at low frequency in normal cells (Helmrich, A. et al 2011) the number and the length may increase when transcription or RNA processing is perturbed (Gan, W. et al 2011) (Tuduri, S. et al 2009) However, disturbing replication, by hindering DNA polymerase progression or by inactivation of replicative topoisomerases, influences the creation or stabilization of R-loops, thus leading to DNA damage (Tuduri, S. et al 2009) (El Hage, A. et al 2010). In yeast THO/TREX mutants, accumulating R-loops behind elongating RNA polymerase II is the major consequence of impaired transcription and increased recombination events (Huertas, P. et al 2003). Interestingly, the suppression of replication stalling at transcribed sites can be achieved by overexpression of RNase H1, the enzyme known to remove R-loops (Aguilera, A et al 2008). It is also proposed that R-loops can increase TAR (Transcription Associated Recombination) in budding yeast by hindering replication fork progression (Prado F, Aguilera A. 2005).

In principle, RNA-DNA hybrids could obstruct fork progression in three non-mutually exclusive ways. R-loops could hamper with DNA replication forks by: (i) Preventing DNA synthesis on the leading or the lagging strand; (*ii*) Obstructing the displacement of the RNAP upon passage of the fork or (*iii*) Favoring the accumulation of DNA lesions on the non-template ssDNA, which would consequently affect DNA synthesis (Aguilera, A et al 2008). In any case, R-loops are thought to be particularly stable to resist the confronted replication machinery and its associated helicases. These structures could also interfere with DNA replication long after transcription has ceased. It has been lately shown that the Sen1/Sentaxin helicase is required in the resolution of R-loops in yeast and human cells, which might be interesting to understand the metabolism of co-transcriptional R loop structures (Mischo, H. E. et al 2011) (Alzu A et al 2012). Other candidates for the elimination of RNA-DNA hybrids are the RecQ helicase RecQ5, which directly interacts with the C-terminal domain of RNA polymerase II and holds significance for maintaining of genome integrity (Aygün, O. et al 2010) (Li, M. et al 2011).

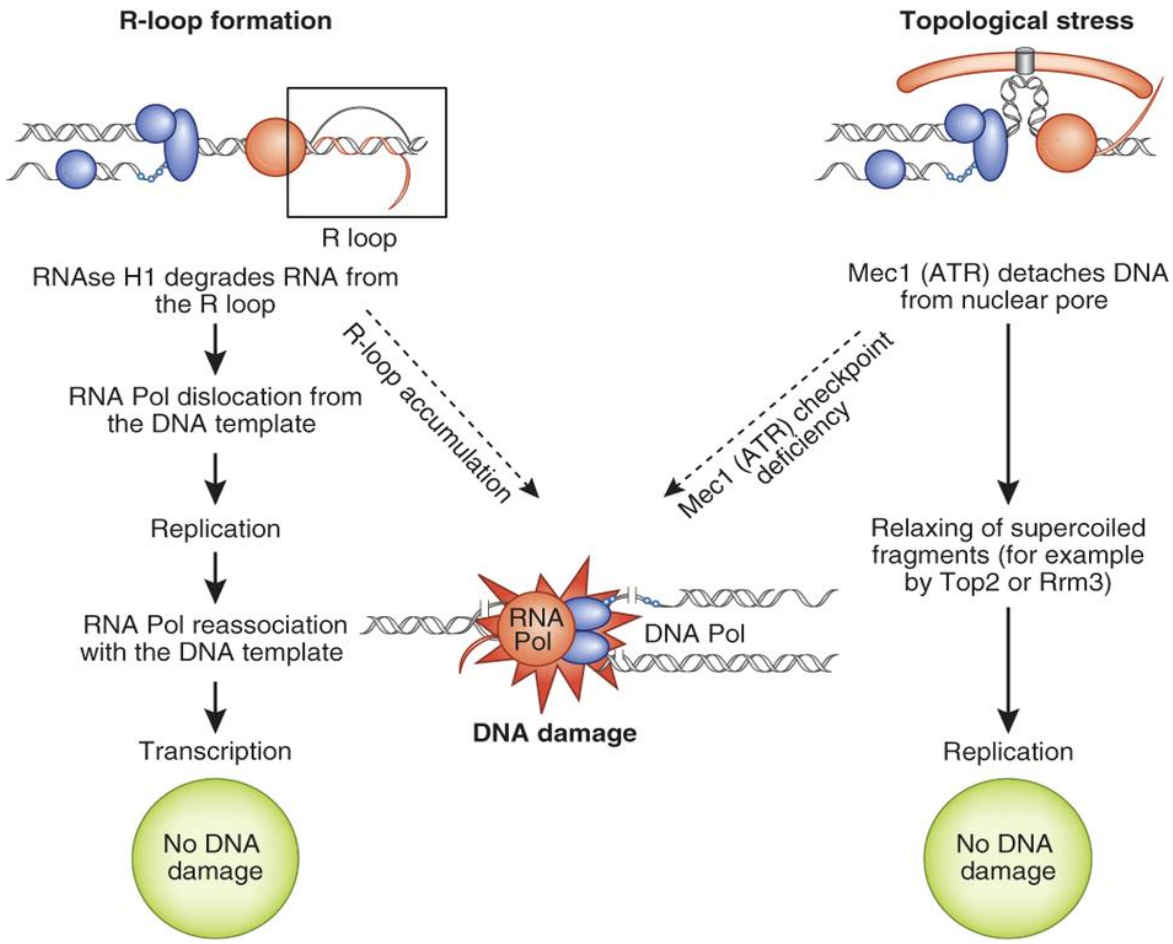


Figure 1.17. Outcomes of transcription-replication collisions. R-loop formation (left) and increased topological stress (right) are the two possible consequences of encounters between transcription and replication machineries. The structure of the R-loop is marked by a box. Under physiological conditions, rescue mechanisms, executed by RNase H1 or by the Mec1 (ATR) checkpoint, prevent DNA damage. Conversely, when rescue mechanisms are either defective or insufficient (dashed arrows), DNA breaks hypermutations are induced. (Adapted from Anne Helmrich et al 2013)

Topological constraints:

Another source DNA instability is the topological tension that emerges when transcription and replication machineries meet at loci that are bound to the nuclear periphery (Bermejo, R et al 2012). In yeast cells inactivation of the Mec1 (ATR) replication checkpoint leads to accumulation of reversed forks and hemireplicated intermediates, when replication stress conditions were induced by low doses of hydroxyurea (Sogo et al Science 2002) (**Fig. 1.17**). It was found that overexpression or ablation of endogenous RNase H1 did not influence hydroxyurea sensitivity, thus inference was drawn that the mechanism leading to replication-fork stalling was independent of R-loop formation (Bermejo et al Cell 2011). Instead, under these

conditions fork reversal appears to require both transcription and the localization of the region to the nuclear periphery. In other words, ablation of nuclear pore proteins or the introduction of DNA double-strand breaks between the transcribed unit and the replication fork suppress hydroxyurea sensitivity. Thus, situations where DNA is attached to the nuclear pore and is unable to freely rotate, the topological stress generated in front of the replisome may cause replication-fork stalling. It has been proposed that the Mec1 (ATR) replication-damage checkpoint proteins assist the nuclear pore complex (NPC)-bound DNA region to be removed from the NPC to permit replication fork progression (Bermejo, R. et al 2011). This model holds further significance due the data demonstrating the participation of the Rrm3 helicase in facilitating replication-fork movement through transcribed loci or sites occupied by DNA-binding proteins (Prado, F.et al 2005) (Ivessa, A.S.et al 2003).

CHAPTER 2

3. DNA SECONDARY STRUCTURES:

DNA exists in the native form of a right-handed double helix called the B-conformation. But situations may arise where DNA occasionally forms cruciforms, G-quadruplex, Z-DNA, triplex DNA (H-DNA) and S-DNA (slipped-strand DNA) (**Figure. 2.1**). On the basis of sequence arrangement and symmetry, simple DNA repeats can be classified under three types i.e. inverted repeats (IRs), mirror repeats (MRs), and direct tandem repeats (DTRs). IRs structures are formed by DNA sequences that are Watson-Crick (WC) complements to each other and are equidistant from the center in a DNA strand. MRs structures are also symmetrical, but they are equidistant DNA bases with 'identical' sequence repeats. Lastly, DTRs structures are the non-interrupted iterations of a 'core repeat unit' along the DNA duplex. The classification between these repeat types is not absolute as e.g. the d(AT)*n*/d(TA)*n* sequence can be considered as an IR, MR, or even DTR. The IRs sequences are capable of forming cruciform structures in double stranded DNA or hairpins in single-stranded DNA (Lilley, D. M. J.et al 1989). MRs with homopurine-homopyrimidine are known to form an intramolecular triple-helical DNA structure called the H-DNA (**Figure 2.1.c**) (Mirkin, S. M.et al 1999). On the basis of the strand contributing to form the triplex i.e. pyrimidine or purine, the resultant structures are called H-y or H-r, respectively. The H-y favors the formation under acidic pH whereas the H-r form is stable at physiological pH with the aid of bivalent cations. DTRs, on the other hand can adopt a variety of conformations depending on their base composition. G-quadruplex (G4) is a DTRs structure which is formed due the presence of tandemly arranged runs of guanines (**Figure 2.1.d**) (Gilbert, D. E. et al 1999). G4s are constituted by a four-stranded structure consisting of square co-planar array of four guanine rich DNA that is additionally stabilized by the presence of monovalent ions (Williamson, J. R.et al 1989). DTRs with alternating pyrimidines and purines may also form double helix into a left-handed zigzag pattern called the Z-DNA (**Figure 2.1.b**) (Rich, A.et al 1984). Lastly, DTRs when direct repeats get base-paired with the complementary strand in a misaligned fashion they form the slipped strand DNA structure (S-DNA) (**Figure 2.1.e**) (Sinden, R. R.et al 1999). Due to denaturing and renaturing conditions, the complementary strands can mispair leading to unusual double-helical structure intervened by single-stranded loops.

During the last decade, the secondary structure formation has gained broad attention due to the discoveries stating their link to human hereditary disorders caused by progressive expansions of microsatellites (Pearson, C. E et al 2005, Wells, R. D et al 1998). Consequently, the role of these repeats in replication and stability of DNA has further triggered the interest in the study of unusual DNA structures.

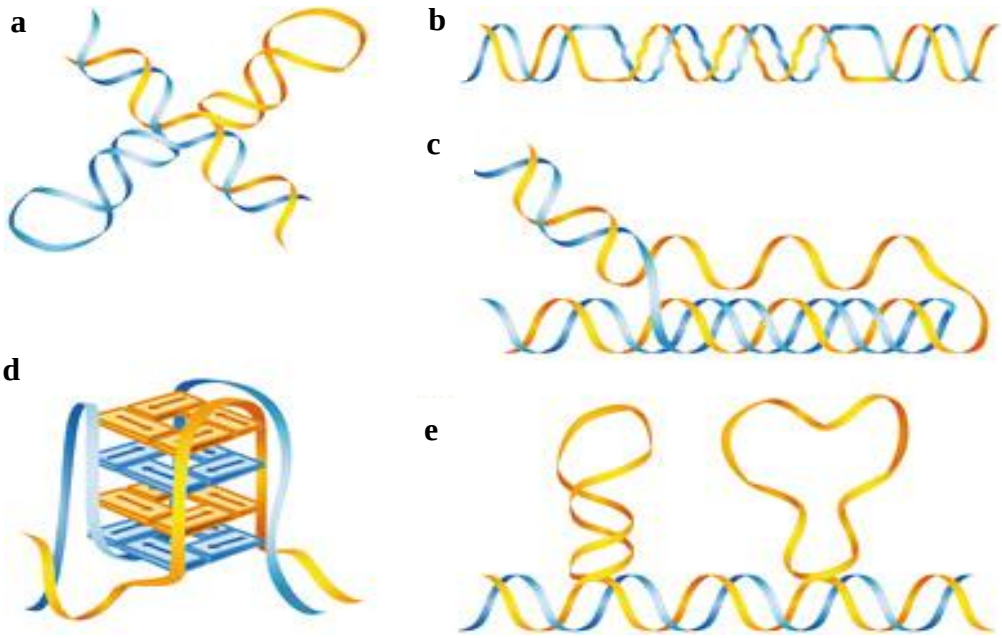


Figure 2.1. Non B-form DNA secondary structures. a) Cruciform DNA structure. b) Left handed helical Z-DNA. c) Triplex DNA (H-DNA). d) G-quadruplex (tetraplex). e) Slipped-strand DNA (S-DNA) (adapted from Junhua Zhao et al 2009)

3.1 Consequences of secondary structure formation:

Growing body of evidence has illustrated that formation of hairpins, triplexes, and G-quadruplex both *in vitro* and *in vivo* leads to DNA replication inhibition. During the course of replication when lagging-strand template transiently become single stranded, it provides a window of opportunity for the formation of unusual structures such as hairpins, triplexes, and G-quadruplexes. These structures interfere with the progression of the lagging-strand polymerase which may lead to blockage of the whole fork ultimately since the two polymerases are linked. Due to the fact that these structures that block progression of the replication fork are initiated by the fork itself, the term 'suicidal sequences' was given to them (Samadashwily, G. M et al 1993).

There are three types of trinucleotide repeats i.e. (CAG)*n*/(CTG)*n*, (CGG)*n*/(CCG)*n*, and (GAA)*n*/(TTC)*n* which are best studied to form these secondary structures. Single-stranded d(CCG)*n*, d(CTG)*n*, d(CGG)*n*, and d(CAG)*n* repeats can fold to form hairpin- like structures stabilized by both WC and non-WC base pairing (Mariappan, S. V et al 1998, Zheng, M. et al 1996). The (CAG)*n*/(CTG)*n* and (CGG)*n*/(CCG)*n* repeats are shown to inhibit replication fork in *E. coli* (Samadashwily, G. M et al 1997), *S. cerevisiae* (Pelletier, R et al 2003), and mammalian cells as well by forming imperfect hairpins or G-quadruplex in the case of (CGG)*n*/(CCG)*n*. Moreover, in both prokaryotes and eukaryotes, these repeats causing replication inhibition is length dependent. For instance, it was seen that a relatively short

(CGG)*n*/(CCG)*n* repeat inhibited DNA replication in yeast whereas a four times longer repeat was required for similar inhibition in bacterial and mammalian cells. A possible explanation could be that for *S. cerevisiae* genome which is highly AT rich, GC rich sequences ((CGG)*n*/(CCG)*n*) repeat are extremely foreign. Additionally, replication stalling by the (CGG)*n*/(CCG)*n* repeat in yeast are orientation independent in the replicon (Pelletier, R et al 2003), which is differently seen in bacteria (Samadashwily, G. M et al 1997) and mammals. Similarly, (GAA)*n*/(TTC)*n* repeat are found to inhibit replication in *S. cerevisiae* in an orientation-dependent manner, specifically with the homopurines strand serves as the lagging strand template i.e. by formation of a triplex (Krasilnikova, M. M et al 2004). Thus, on the basis of strength of inhibition in different systems the repeats base composition can be arranged in the following order: (CGG)*n*/(CCG)*n* > (GAA)*n*/ (TTC)*n* > (CAG)*n*/(CTG)*n*, in co-relation with the tendency to form unusual DNA structures.

3.2 G-quadruplex; as secondary structure:

By the recent results of ENCODE project it has been identified that ~3% of human genome has the potential to form non-canonical secondary structures, which may hold functional importance (Bernstein B.E. et al 2012). One example of such DNA secondary structures is the four-stranded G-quadruplexes, which has gained considerable attention for their emerging role in biological pathways. Back in 1910, (Bang I. et al 1910) it was first observed and the G-tetrad structure was then identified in 1962 (Gellert M. et al 1962). The DNA G-quadruplexes formation was first found in the single-stranded 3'overhang of human telomeres (Sundquist W.I. et al 1989) (Williamson J.R. et al 1989). Later on, they were found to form in the proximal promoter regions of human oncogenes to regulate gene transcription (Chen Y.et al 2012) (Balasubramanian S.et al 2011). Recently, the DNA G-quadruplexes have been associated with replication initiation (Paeschke K.et al 2011) (Besnard E.et al 2012) and newly identified RNA G-quadruplexes have also been found to form in mRNA i.e. in 5'- and 3'-UTR repeats (Christiansen J.et al 1994) (Bugaut A.et al 2012) Many proteins have been found which interact with G-quadruplex structures; which either tends to stabilize G-quadruplexes or unwind and destabilize G-quadruplexes (Yang D.et al 2010) (Fry M.et al 2007). By the aid of G-quadruplex specific antibody, G-quadruplexes have been found in human cells at various sites on chromosomes other than telomeres. There is also an increase in frequency of G-quadruplexes when in the presence of a G-quadruplex-interactive compound (Siddiqui Jain A.et al 2013) (Biffi G.et al 2013). Quarfloxin, which is the first G-quadruplex interactive drug, has already reached phase II clinical trials for the treatment of cancer (Drygin D.et al 2009). Thus, it is of great significance to comprehend the formation, biological role, and potential drug ability of G-quadruplexes (Yang D.et al 2010).

3.2.1 Structures of G-quadruplexes:

G-quadruplexes basically are a stack of guanine tetrads (G-tetrads) (**Figure 2.2**). Within a G-tetrad, four guanine bases are arranged to form a square plane using Hoogsteen hydrogen bonding, in place of the Watson-Crick hydrogen bonding for B-DNA. G-quadruplex structures readily acquire monovalent cations like Na+ or K+ (Sen D.et al 1990), which physiologically stabilize the G-quadruplex by positioning between the G-tetrad planes in coordination with the O6 atoms of the tetrad guanines (**Figure 2.2.a**). G-quadruplexes can be monomeric or multimeric like dimeric or tetrameric (**Figure 2.2. c, d**). Within the G-tetrad, guanine residues have a tendency to adopt either syn or anti-glycosidic conformation (Chen Y.et al 2012) (**Figure 2.2.b**). Likewise, the adjacent DNA strands in a G-quadruplex can have the parallel (same) or anti-parallel (opposite) orientation, where the possible orientations are depicted in the Figure 2.2.d.

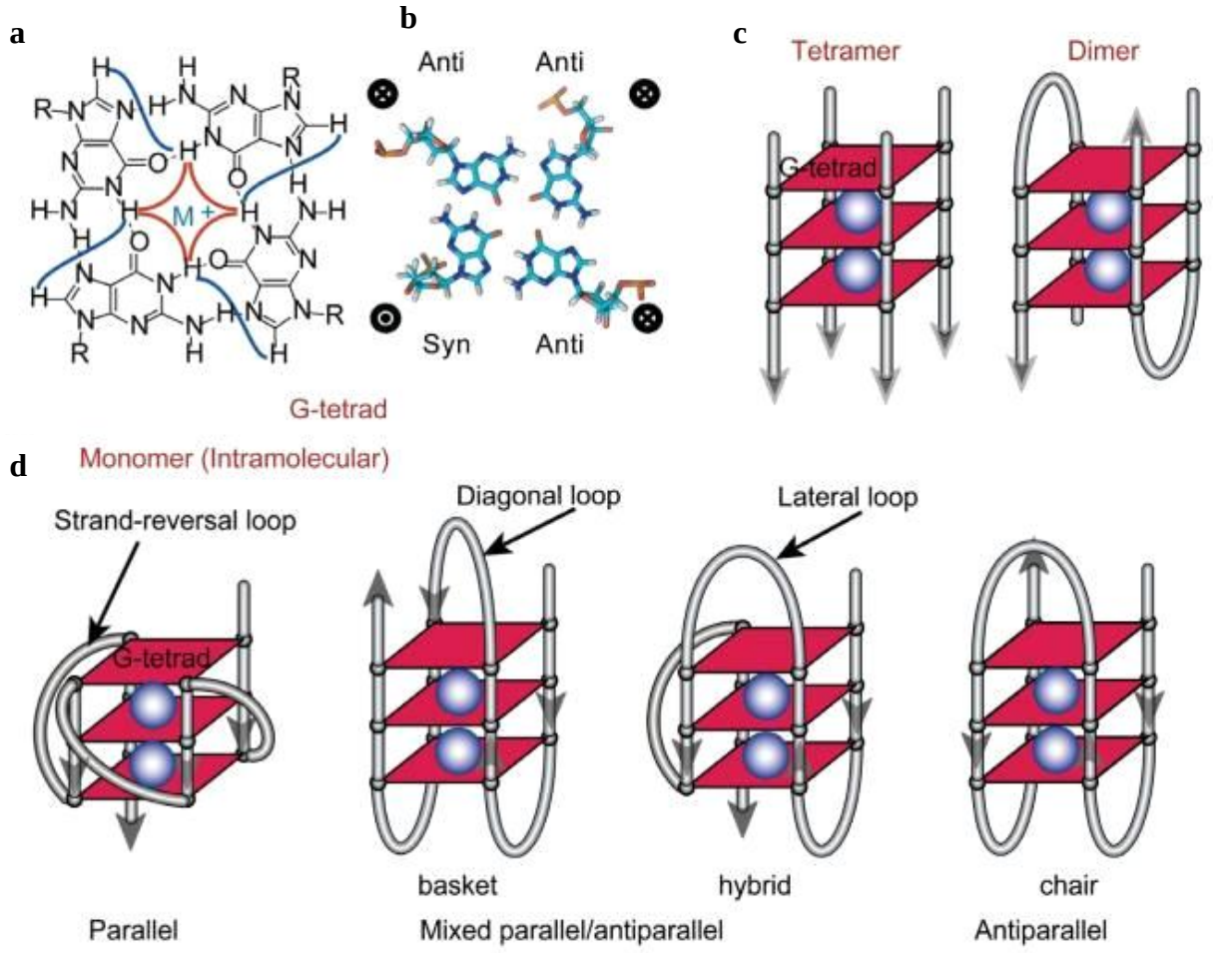


Figure 2.2. Schematic illustration of a G-quadruplex structures. a) Four guanine bases arranged in a square plane with Hoogsteen hydrogen bonding. The H1–H1 and H1–H8 connectivity are shown in red and blue respectively. b) A G-tetrad structure. Guanines in a G-tetrad may adopt either syn or anti glycosidic conformation; the guanines from parallel G-strands adopt the same glycosidic conformation

and the guanines from antiparallel G-strands adopt opposing glycosidic conformations. c) The tetrameric and dimeric G-quadruplexes with three G-tetrads. d) Types of monomeric (intramolecular) G-quadruplexes with different folding structures and loop conformations. Monovalent cations (K+ or Na+, shown as blue spheres), are required to stabilize G-quadruplexes by coordinating with the O6 atoms of the adjacent G-tetrad planes. (Adapted from Onel Buket1 Et Al 2014)

3.2.2 Effects of G-quadruplex on DNA replication:

During DNA replication, the DNA duplex is unwound by the aid of replicative helicase where one strand serves as the template for leading strand synthesis and the other for lagging strand synthesis. While the leading strand DNA replication is continuous, the lagging strand is replicated discontinuously. Due to this irregularity in synthesis a transient single-stranded state may occur in lagging strand; providing an opportunity for G-quadruplex structure formation (**Figure 2.3**). Besides, some G-quadruplexes could be present on the template prior to its replication because they have roles in transcriptional regulation (Huppert, J. L et al 2008). For completion of DNA replication the presence of G-quadruplex structures in its path may act as barriers. It is therefore important that these secondary structures are resolved efficiently, irrespective of whether they are pre-existing or formed during DNA replication. Likewise, the sequence forming G-quadruplexes also cannot serve as a template for replication, thus helicases play a critical role in avoiding such problems.

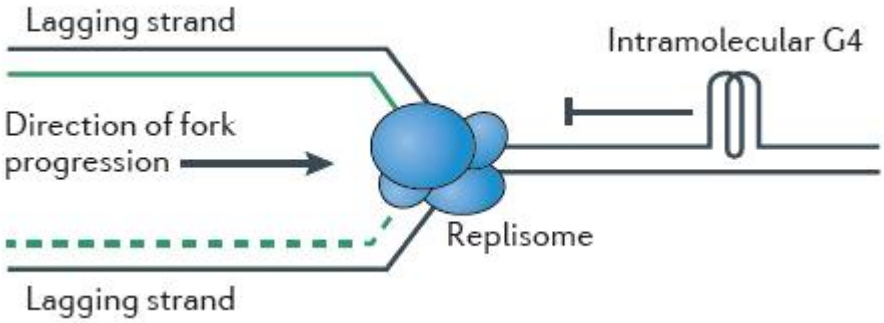


Figure 2.3. G-quadruplex structures during DNA replication. Putative hindrance for the approaching replication fork by the formation of G-quadruplex. Replisome in blue (adapted from Matthew L. Bochman et al 2012)

3.2.3 Role of G4 unwinding DNA helicases:

In human, mutations in helicases that unwind G4 structures *in vitro* (London, T. B. et al 2008) (Sanders, C. M.et al 2010) are associated with diseases that cause genomic instability. These include the RecQ helicase WRN that is associated with premature ageing, BLM & FANCJ that

are associated with increased cancer risk and lastly PIF1 mutations which also increase cancer risk. The most evident link of human disease and G-quadruplex unwinding came from the findings from human patient cell lines with Fanconi anemia carrying FANCJ mutations. It was found that the deletions due FANCJ mutations overlapped with G-ich regions that potentially could form G-quadruplexes (London, T. B. et al 2008). Likewise, treatment with telomestatin; a chemical ligand that stabilizes G-quadruplexes *in vitro* (Gomez, D. et al 2004) (Neidle, S.et al 2009) led to impaired proliferation, increased apoptosis and DNA damage in FANCJ-deficient cells (Wu, Y.et al 2008). Thus, with the association of these helicases with genome instability highlights the possibility that G-quadruplexes unwinding might suppress both premature ageing and cancer.

Similarly, some enzymes are far more active on G-quadruplexes than others. The *S. cerevisiae* Pif1 helicase (Paeschke, K.et al 2011) and members of the Pif1 DNA helicase family are particularly efficient *in vitro* unwinding of parallel intramolecular G-quartet substrates (Ribeyre, C. et al 2009). Pif1 is a class of DNA helicase that binds >1,000 sites in the genome of mitotic cells, of which ~10% overlap G-quadruplex motifs i.e. ~25% of the G-quadruplex motifs within the organism. However, this estimate excludes the G-quartet motifs in ribosomal and telomeric DNA, both of which are strong Pif1 binding sites (Paeschke, K.et al 2011). Several genetic assays show that in the absence of Pif1, DNA replication slows and DSBs occur at many of the G-quadruplex motifs that are normally bound by Pif1. The G-quadruplex motifs also show a high mutation rate in Pif1-deficient cells and the inability of the motif to form a G-quadruplex, without reducing the high GC content of the motif. When these mutated motifs are put back in the genome, they no longer show Pif1 binding or slow DNA replication and cause DSBs. Thus, making a strong argument that G-quadruplexes are formed *in vivo* and their resolution by Pif1 suppresses genome instability (Paeschke, K.et al 2011). There are also references that human PIF1 acts at G-quadruplex motifs. In one study the chromatin immunoprecipitation followed by sequencing (ChIP–seq) in combination with *in vivo* labeling with pyridostatin, a G-quadruplex binding molecule was conducted (Rodriguez, R. et al 2012). A genome-wide, pyridostatin bound preferentially to G-quadruplex motifs, where replication and transcription-dependent damage caused was detected by its high γH2Ax content. Since many of the γH2Ax foci overlapped with GFP–PIF1 foci in the pyridostatin-treated human cells, it was hypothesized that G-quadruplex formation or stabilization blocks transcription and/or replication, resulting in DNA damage.

3.2.4. CEB25 a G-quadruplex forming human minisatellite:

CEB25 is a human minisatellite locus, composed of slightly polymorphic 52-nucleotide (nt) tandem repeats. Genetically, most if not all individuals of the human population are heterozygous, carrying alleles ranging from 0.5 to 20 kb, maintained by mendelian inheritance but also subject to germline instability. G-rich human CEB25 minisatellite is a remarkable G-quadruplex forming motif which contains successive guanines prone to G-quadruplex formation, in particular, a segment, AAGGGTGGGTGTAAGTGTGGGTGGGT, composed of four GGG tracts (underlined) separated by linkers of 1, 9, and 1 nucleotides, respectively. This 25-nt G-rich

sequence is responsible for harboring four G-tracts taken from the 52-nt full repeating unit of the CEB25 minisatellite (**Figure 2.4.b**). In addition, this sequence forms a propeller-type parallel-stranded G-quadruplex in presence of K+ ion involving a 9-nucleotide (nt) central double-chain-reversal loop (**Figure 2.4.b**). This long loop is anchored to the 5′ end of the sequence by an AT Watson-Crick base pair and a potential GA non-canonical base pair, contributing to the stability of the overall G-quadruplex structure by increasing the enthalpy by 17 kcal/mol or 6 °C in melting temperature (Samir Amrane et al 2012) .

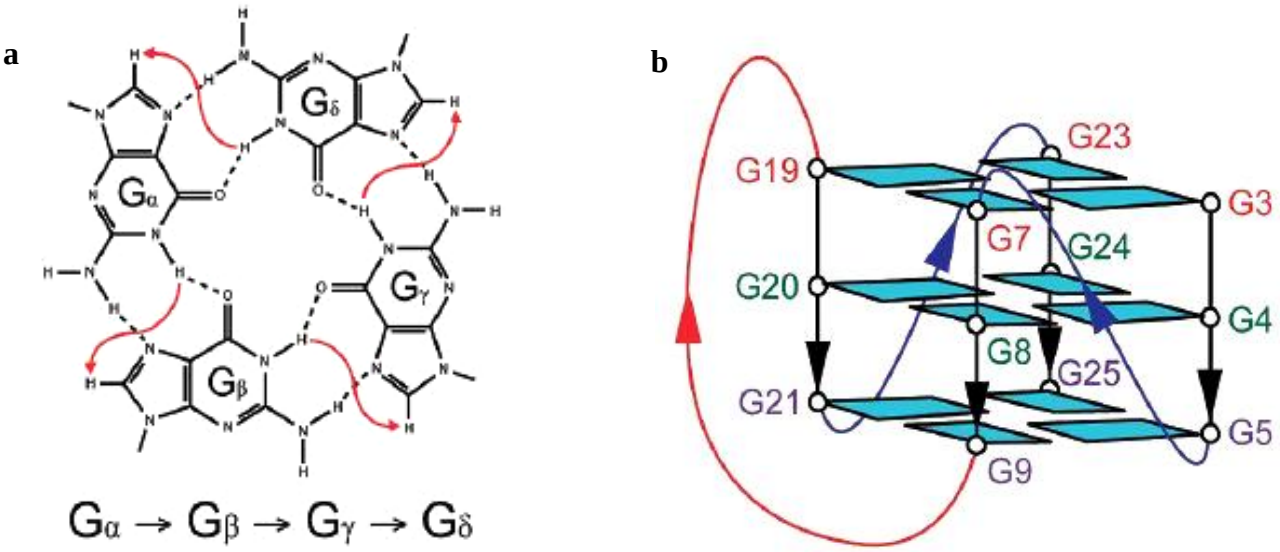


Figure 2.4. The CEB25 G-quadruplex structure depicted by two-dimensional nuclear magnetic resonance spectroscopy (2D NMR). a) Four guanine bases Gα·Gβ·Gγ·Gδ arranged in a square plane with Hoogsteen hydrogen bonds showing the proximity of imino (H1) and H8 protons (in red). b) Schematic structure of the 25nt CEB25 G-quadruplex where guanines in the G-tetrad core are numbered. The backbones of the core, 1-nt loops, and 9-nt loop are colored black, blue, and red, respectively (adapted from Samir Amrane et al 2012)

3.2.5 C-MYC; three-tetrad G-quadruplexes in gene promoters:

The formation of three-tetrad G-quadruplex structures near promoter regions suggests a potential function of G-quadruplex structures in gene regulation. This holds significance as one or more G-quadruplex motifs are found within 1,000 nucleotides (nt) upstream of the TSS of 50% of human genes (Huppert, J. L et al 2007). Interestingly, bioinformatics has also shown that the promoters of human oncogenes and regulatory genes like transcription factors are likely to contain G-quadruplex motifs than an average gene (Eddy, J.et al 2006). Likewise, during transcription negative supercoiling is accumulated, which favors G-quadruplex structure to form as a result of supercoiling-induced stress during transcription (Sun, D. et al 2009). *In vitro* studies also show that the formation of G-quadruplex structures can compensate for the negative supercoiling so

accumulated during transcription (Sun, D. et al 2009) (Brooks, T. A.et al 2010). These findings suggest that G-quadruplex structures in or near promoter regions may influence transcription machinery.

One such best-studied system for role of G-quadruplex structures in transcription involves the mammalian *C-MYC* locus. *C-MYC* is an important oncoprotein and transcription factor that plays an important role in cell proliferation and induction of apoptosis (K.B. Marcu.et al 1992). Over-expression of *C-MYC* is associated with a significant number of human malignancies e.g. breast, colon, cervix, and small-cell lung cancers, to list a few (S. Pelengaris.et al 2003). *C-MYC* transcription is under the complex control of multiple promoters where the nuclease hypersensitivity element III_1 (NHE III_1) is in the proximal region of the *C-MYC* promoter (-142 to -115 base pairs). It also controls about 80-90% of the total transcriptional activity of this gene (O. Sakatsume.et al 1996). The *C-MYC* G-quadruplex consist of a 37 nt long tandem repeat which has a stretch of 24 nt G-rich region forming G-quadruplex. The 24nt sequence is: 5′ -GGGGAGGGTGGGGAGGGTGGGGCC-3′ and consist of four GGG tracts (underlined) (**Figure 2.5**). This G-rich strand can readily form a G-quadruplex structure under physiological conditions and its formation is further stabilized by the presence of K+ ion (T. Simonsson.et al 1998). The *C-MYC* G-quadruplex also has a melting point in excess of 85^{o}C. Additionally, it has been also postulated that the C-rich strand of the *C-MYC* NHE III_1 might adopt an 'i-motif' structure too.

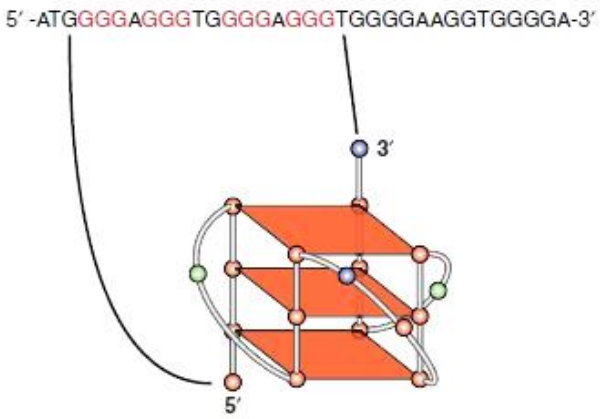


Figure 2.5. Schematic representation of NMR structure of *C-MYC* minisatellite forming G-quadruplex. (Adapted from Tracy A. Brooks et al 2010)

With this knowledge on how, when and where these human minisatellites form G-quadruplex, little is known till date about how there formation affects the fork stability. Thus, part of my work focused on gaining insight into the impact of these highly polymorphic G-rich repeats on replication fork progression by the aid of genetic analysis techniques using yeast as a study model.

OBJECTIVES

The aim of my thesis is to investigate the mechanisms suppressing replication fork instability in context of conflicts of replication machineries with active gene transcription and during replication progression across DNA secondary structure forming G-quadruplexes.

Objectives:

- Characterize role of the THO complex subunit Hpr1 in modulating transcribed chromatin topology and protecting replication fork integrity.

- Study the contribution of the Pif1 helicase essential for replication of different G-quadruplex forming human minisatelite sequences.

MATERIALS AND METHODS

7.1 Strains and plasmids.

All W303 strains used in this study are isogenic derivatives of **W303-1a** *RAD5+* background (Thomas & Rothstein, 1989). The genotype is listed in the table below (**Table 7.1**). Plasmids used by specific strategies are listed in Table 7.2. Deletion strains were constructed using PCR-based gene disruption strategy (Brachmann et al., 1998; Longtine et al., 1998). Protein tagging was performed by introducing the in-frame sequence of the corresponding epitope (FLAG, HA, MYC, PK) at the C-terminal end of the gene of interest.

Table 7.1. Genotype of strains generated and used in this study.

STRAIN	NUMBER	GENOTYPE	REFERENCE
WT	RB718	MATa, ADE2, can1-100, his3-11,15 leu2-3,112 trp1-1, ura3-1 RAD5+	Lab collection
WT (pYES2)	RB1155	MAT a ADE2+ CAN1+, ura3-1, his3-11,15 leu2-3, 112 trp1-1, RAD5+, pYES2-URA3	This study
WT (YCp50)	RB1644	MATa, ADE2, can1-100, his3-11,15 leu2-3,112 trp1-1, ura3-1 RAD5+,c	This study
Hpr1-Flag	RB2100	MAT alpha ,W303, RAD5+, HPR1-FLAG: KanMX	This study
hpr1-Y590A	RB856	MATa ADE2+ CAN1+, ura3-1, his3-11, leu2-3, 112trp1-1, RAD5+, hpr1-Y590A	Lab collection
hpr1-Y590A (pYES2)	RB1163	MATa ADE2+ CAN1+, ura3-1, his3-11, leu2-3, 112trp1-1, RAD5+, hpr1-Y590A, pYES2-URA3	This study
hpr1-Y590A(Ycp50)	RB1805	MATa ADE2+ CAN1+, ura3-1, his3-11, leu2-3, 112trp1-1, RAD5+, hpr1-Y590A	This study
hpr1-Y590A-Flag	RB2088	MAT alpha ,W303, RAD5+, HPR1-Y590A-FLAG: KanMX	This study
top2-1	RB59	MATa ADE2+ CAN1+, ura3-1, his3-11,15 leu2-3, 12 trp1-1, RAD5+, top2-1	Lab collection
top2-1 (pYES2)	RB1158	MATα ADE2+ CAN1+, ura3-1, his3-11,15 leu2-3, 112 trp1-1, RAD5+, top2-1, pYES2-URA3	This study
top2-1 (Ycp50)	RB1641	MATα ADE2+ CAN1+, ura3-1, his3-11,15 leu2-3, 112 trp1-1, RAD5+, top2-1, Ycp50	This study

top2-1 hpr1-Y590A	RB877	MATa, ADE2+ CAN1+, ura3-1, his3-11, leu2-3, 112trp1-1, RAD5+, hpr1-Y590A, top2-1	This study
top2-1 hpr1-Y590A (pYES2)	RB1170	MATa, ADE2+ CAN1+, ura3-1, his3-11, leu2-3, 112trp1-1, RAD5+, hpr1-Y590A, top2-1, pYES2-URA3	This study
spt4Δ (pYES2)	RB1171	MATa, ADE2+ CAN1+, ura3-1, his3-11, leu2-3, 112trp1-1, RAD5+, spt4Δ, pYES2-URA3	This study
top2-1 hpr1-Y590A (Ycp50)	RB2406	MATa ADE2+ CAN1+, ura3-1, his3-11, leu2-3, 112trp1-1, RAD5+, hpr1-Y590A, top2-1, Ycp50	This study
top1 Δ	RB357	MATa ADE2+ CAN1+, ura3-1, his3-11, leu2-3, 112trp1-1, RAD5+, top1::TRP	Lab collection
top1 Δ (Ycp50)	RB1647	MATa ADE2+ CAN1+, ura3-1, his3-11, leu2-3, 112trp1-1, RAD5+, top1::TRP, Ycp50	This study
top1 Δ hpr1-Y590A	RB875	MATa ADE2+ CAN1+, ura3-1, his3-11, leu2-3, 112trp1-1, RAD5+, hpr1-Y590A, top1::TRP	This study
top1 Δ hpr1-Y590A (Ycp50)	RB2403	MATa ADE2+ CAN1+, ura3-1, his3-11, leu2-3, 112trp1-1, RAD5+, hpr1-Y590A, top1::TRP, Ycp50	This study
top1 Δ top2-1	RB866	MATa ADE2+ CAN1+, ura3-1, his3-11, leu2-3, 112trp1-1, RAD5+, top1::TRP, top2-1	This study
top1 Δ top2-1 (Ycp50)	RB1904	MATa, ADE2+ CAN1+, ura3-1, his3-11,15 leu2-3, 112 trp1-1, RAD5+, top2-1 top1::TRP, Ycp50	This study
top1 Δ top2-1 hpr1-Y590A	RB882	MATa, ADE2+ CAN1+, ura3-1, his3-11, leu2-3, 112trp1-1, RAD5+, hpr1-Y590A, top1::TRP, top2-1	This study
Tho2-PK	RB2102	MATa, ADE2+ CAN1+, ura3-1, his3-11, leu2-3, 112trp1-1, RAD5+, THO2-PK :TRP	This study
CEB25-1nt WT	RB1008	Mata, ura3-1, trp1 -1, leu2-3,112, his3-11,15 , ADE2, ARG4, CAN1+ RAD5+, ARS305-CEB25-Loop1-0.4 orientation I	Lab collection (A.Nicolas)
CEB25-1nt pif1Δ	RB1009	Mata, ura3-1, trp1 -1, leu2-3,112, his3-11,15 , ADE2, ARG4, CAN1+ RAD5+, ARS305-CEB25-Loop1-0.4 orientation I, pif1::HIS3	Lab collection (A.Nicolas)
CEB25-9nt WT	RB1006	Mata, ura3-1, trp1 -1, leu2-3,112, his3-11,15 , ADE2, ARG4, CAN1+ RAD5+, ARS305-CEB25-WT-0.7 orientation I	Lab collection (A.Nicolas)
CEB25-9nt pif1Δ	RB1007	Mata, ura3-1, trp1 -1, leu2-3,112, his3-11,15 , ADE2, ARG4, CAN1+ RAD5+, ARS305-CEB25-WT-0.7 orientation I, pif1::HIS3	Lab collection (A.Nicolas)
c-MYC WT	RB1696	Mata, ura3-1, trp1 -1, leu2-3,112, his3-11,15 , ADE2, ARG4, CAN1+ RAD5+, C-MYC CEB1spacer (Orientation I / leading strand) WT (ORT7338-2)	Lab collection (A.Nicolas)
c-MYC pif1Δ	RB1697	Mata, ura3-1, trp1 -1, leu2-3,112, his3-11,15 , ADE2, ARG4, CAN1+ RAD5+, C-MYC CEB1spacer (Orientation I / leading strand) pif1 delta (ORT7345-8)	Lab collection (A.Nicolas)

Table 7.2: Plasmid used in this study.

PLASMID	NUMBER	AIM	REFERENCE
Ycp50	BB 70	Chloroquine 2D gel electrophoresis	Lab Collection (J.Roca)
pYES2	BB36	6-Azauracil Experiments	Lab Collection

7.2. Growing media for *Saccharomyces cerevisiae*:

Solid media:

- Complete media YPDA:

Yeast extract	4g
Peptone	8g
D-glucoase 40%	20ml
Agar	8g
H_2O(milli Rho)	380ml

Insoluble Adenine was added at a final concentration of 25µg

- Minimum media:

YNB (w/o aa)	2.8 g
Agar	8 g
D-glucose 40%	20 ml
H2O (milli Rho)	380 ml
Drop-out*	16 ml

Aminoacids used for selection (HIS, TRP, LEU, URA) were added at the final concentration (Cf) of 25 µg/mL

***DROP-OUT**: (Final Volume 2 litres)	
Thr	1.2 g
Phe	1.2 g
Ile	1.2 g
Lys	1.2 g
Arg	1.2 g
Tyr	1.2 g
Ino	1.74 g
Ade	1.2 g
Etoh 100%	120 ml
ddH2O	1800 ml

- YNB (Yeast Nitrogen Base):

YNB (w/o aa)	2.8 g
Agar	8g
D-glucose 40%	20 ml
H2O (milli Rho)	380 ml

- Sporulation medium (VB):

Anhydro CH_3CO_2Na	3.28 g
KCl	0.76 g
NaCl	0.48 g
MgSO4	0.14 g
Agar	6 g
H2O (milli Rho)	400 ml

Liquid medium:

- Complete medium YPD:

Yeast extract	4 g
Peptone	8 g
D-glucose 40%	20 ml
H2O (milli Rho)	380 ml

- Complete medium YPDA:

Yeast extract	4 g
Peptone	8 g
D-glucose 40%	20 ml
H2O (milli Rho)	380 ml

Insoluble Adenine was added at a final concentration of 25µg/ml

- -URA medium:

YNB (w/o aa)	2.8 g
Agar	8 g
D-glucose 40%	20 ml
H2O (milli Rho)	380 ml
Drop-out *	16 ml
HIS, TRP, LEU	Cf 25 µg/ml

Media with drugs:

To minimal or complete (YPDA) liquid/solid the corresponding amount of hydroxyurea (HU) and 6-Azauracil (6AU) were added depending on the desired final concentration.

7.3 List of buffers

Blot#2: 1M $AcNH_4$, 0.02M NaOH

Buffer G2 (digestion buffer)**:** 800mM guanidine HCl, 30mM Tris-HCl pH 8.0, 30mM EDTA pH 8.0, 0.5% Tween-20, 0.5% Triton X-100

Buffer QBT (equilibration buffer)**:** 750mM NaCl, 50mM MOPS pH 7.0, 15% Isopropanol, 0.15% Triton X-100

Buffer QC (wash buffer)**:** 1M NaCl, 50mM MOPS pH 7.0, 15% Isopropanol

Buffer QF (elution buffer)**:** 1.25M NaCl, 50mM Tris-HCl pH 8.5, 15% Isopropanol

Denaturing solution: 0.5M NaOH, 1.5M NaCl

Elution Buffer: 50mM Tris-HCl pH 8.0, 10mM EDTA, 1% SDS

FACS Buffer solution: 200mM Tris-HCl pH 7.4, 200mM NaCl, 80mM $MgCl_2$

Laemmli Buffer 1X: 2% SDS, 10% Glycerol, 5% β-mercaptoethanol, 0.002% Bromophenol blue, 0.125 M Tris-HCl pH 6.8

Lysis buffer: 50mM Hepes-KOH pH 7.5, 140mM NaCl, 1mM EDTA, 1% Triton X-100, 0.1% Na-deoxycholate

NIB buffer (pH 7.2): 17% Glycerol, 50mM MOPS, 150mM K-acetate, 2mM $MgCl_2$, 500mM Spermidine, 150mM Spermine

PBS: 137mM NaCl, 10mM PO4, 2.7mM KCl

Ponceau S: 0.1% Ponceau S, 1% acetic acid, H2O

SSC 20X: 3M NaCl, 0.3M Na citrate ($C_6H_5Na_3O_7$)

Running buffer 1X: 25mM Tris-base, 192mM Glycine, 0.1% SDS

Spheroplast Solution: 1M Sorbitol, 100mM Tris, 20mM EDTA, 0.1% MercaptoEtOH, 1mg/ml Zymolyase

TAE: 0.04M Tris-Acetate, 0.001M EDTA

TBE: 89mM Tris-Borodate, 89mM Boric Acid, 2mM EDTA

TBS: 20mM Tris-HCl pH 7.5, 150mM NaCl

TE: 10mM Tris-HCl pH 7.4, 1mM EDTA

Toluene Solution: 20mM Tris, 95% Ethanol, 3% Toluene, 10mM EDTA

Transfer buffer: Glycine 1%, Tris-base 0.02M, Methanol 20%

Wash buffer: 10mM Tris-HCl pH8.0, 250mM LiCl, 0.5% NP-40, 0.5% Na-deoxycholate, 1mM EDTA

Washing solution I: SSC 2X, 1% SDS

Washing solution II: SSC 0.1X, 0.1% SDS

7.4. PCR:

Polymerase chain reaction (PCR) is a technique used for the *in vitro* amplification of specific DNA sequences, which can further be used to transform yeast cells and produce yeast mutants. A PCR reaction requires two primers (forward & reverse) which are oligonucleotide sequences (17-30 base pairs) flanking the DNA region to synthesize complementary strand to the template. Depending on the specific temperatures required, the process of PCR amplification is divided into three steps namely, Denaturation- annealing- extension. These 3 steps are repeated 20-30 times to obtain a satisfactory amplification of desired sequence:

- **Denaturation**: the DNA double helix is separated into the two single helixes by exposing to high temperature. (T = 94°C).
- **Annealing**: primer recognizes and binds to its complementary sequence in one of the two separated strands of the DNA helix (T = 45-60°C). The primers have a free 3'-end in order to make possible the synthesis on both DNA strands.
- **Extension**: DNA polymerase synthesizes a new complementary strand to the template in 5' to 3' direction, by adding dNTPS to the flanking primers. (T = 72°C).

Depending on the DNA sequence to be amplified (cassette), different PCR reaction mixtures and programmes were used.

- **DELETION** (HIS, TRP, URA) or **MYC-TAG cassette**

Buffer 10X (Biotools 10.002)	100 µl	94°C	3'	
dNTPs (2mM)	100 µl	94°C	30"	
primer forward (250 ng/µL)	20 µl	42°C	30"	8 cycles
primer reverse (250 ng/µL)	20 µl	72°C	1'30"	
specific DNA template (20 ng/µL)	10 µl	94°C	30"	
Dynazyme polymerase	20 µl	58°C	30"	30 cycles
ddH20 sterile	730 µl	72°C	1'30"	
	----------	72°C	7'	
final volume	1000 µl			

- **FLAG TAG cassette**

Buffer 10X (Biotools 10.002)	100 µl	94°C	3'	
dNTPs (2mM)	100 µl	94°C	1'	
primer forward (250 ng/µl)	10 µl	42°C	1'	32 cycles
primer reverse (250 ng/µl)	10 µl	72°C	1'30"	
BB7 (10 ng/µl)	20 µl	94°C	10'	
Dynazyme polymerase	20 µl			
ddH20 sterile	730 µl			
	----------			
final volume	1000 µl			

- **PK-TAG cassette**

Buffer 10X (Biotools 10.002)	100 µl	94°C	5’	
dNTPs (2mM)	100 µl	94°C	15”	
primer forward (250 ng/µl)	10 µl	45°C	15”	32 cycles
primer reverse (250 ng/µl)	10 µl	72°C	2’	
BB6 (10 ng/µl)	20 µl	72°C	7’	
Dynazyme polymerase	20 µl			
ddH20 sterile	730 µl			
final volume	1000 µl			

PCR product purification:

PCR products were analysed by electrophoresis on an agarose gel (0.8% in 1X TAE). 5 volume of PB buffer was added to 1 volume of PCR sample. The sample mixture was then placed in a QIAquick column (in a 2ml collection tube) and centrifuged (to bind the DNA). The flow through was discarded, further to the column 750µl of PE buffer was added and centrifuged. Flow through was discarded and the column was placed in a 1.5ml tube. To elute out the DNA, 30µl of EB (10 mM Tris·Cl, pH 8.5) was added to the centre of column and centrifuged. Concentration of the PCR product was found out using the nano drop or by loading in gel.

Plasmids used to amplify PCR cassette are listed in the table below:

Table 7.3. Plasmids used for PCR.

PLASMID NAME	NUMBER	REFERENCE
pPK9-KanMX6	BB6	Lab collection
pU6H3FLAG	BB7	Lab collection
pFA6a-His3MX6	BB8	Lab collection
pFA6a-TRP1	BB9	Lab collection
pRS406 (URA3)	BB47	Lab collection

7.5 High efficiency LiAc transformation:

To generate strains that expressed a tag version of the protein of interest, high efficiency transformation protocol was used. Strains to be transformed were transferred into 5-10 ml of YEPD liquid media and put in a shaking incubator set at 25^{o}C for 6 to 10 hrs. Cell concentration was then calculated by counting cells in a neubauer chamber under a microscope. Fresh YEPD liquid media was taken into a conical flask (100 to 200 ml depending on the experiment). Using the excel sheet the volume of pre-inoculum (containing cells with known conc.) to be added to the media for overnight growth was calculated such that the final cell concentration reaches 1x10^{7} cells/ml. The volume of pre-inoculum calculated was added to the fresh media and kept on

a shaking incubator set at 25^{o}C overnight. The culture was centrifuged for 3 minutes at 4000 rpm and the pellet was rinsed with 25 ml of sterile water to completely wash away any traces of the medium. The pellet obtained was resuspended in 1 ml di 0.1M LiAc/TE 1X and transferred in a 1.5 ml Eppendorf tube. Cells were centrifuged at maximum speed for 15 seconds and resuspended in 500 µl di 0.1M LiAc/TE 1X. The cell suspension was vortexed and split into individual aliquots of 50 µl, for each transformation. Alongside, salmon sperm DNA (ss-DNA) was boiled for 5 minutes at 95°C, for it to be used as DNA carrier.

The 50µl cell suspension was centrifuged at maximum speed for 15 seconds and the transformation mix was added to the pellet in the following order:

PEG (50% W/v)	240 µl
1M LiAc	36 µl
ss-DNA (9.5 mg/ml)	10.5 µl
DNA (plasmid or PCR product)	1-5 µg (x µl)
Sterile ddH2O	73.5 - x µl

Final volume	360 µl

For cell transformation different quantities (1-5 µg) of DNA were taken and depending on that the corresponding volume of ddH2O was added to reach the final volume of 360 µL. To obtain a homogenous mixture the transformation mix was vortexed vigorously for at least 1 minute followed by incubation for 40 minutes at 42°C. This step, called "heat Shock", allows the cells to incorporate the foreign DNA contained in the mix. Following the heat shock, cells were centrifuged for 15 seconds at 7000 rpm. The supernatant was removed with the vacuum pump without disturbing the pellet, which was resuspended in sterile water and plated in the corresponding selective medium. If the cassette used to transform cells carried an antibiotic resistance marker, for example naturomycin (NAT), kanamycin (KAN) or hygromycin (HPH), cells were allowed to grow for at least 3 hours in 3 ml YPDA before plating them. This was done to allow them to express their resistance gene.

7.6 Growth conditions, cell cycle arrest:

Pre-inoculum cultures of required *S.cerevisiae* strains were transferred into rich YPDA media at required temperature and allowed to grow till they reach a final concentration of $1x10^{7}$ cells/ml. Strains with mating type MAT-***a*** were synchronized in G1 by the addition of synthetic α-factor pheromone at a final concentration of 5 µg/mL(Asynchronous samples were withdrawn before the addition of α-factor). Cells are constantly checked after 45mins for their G1 arrest which is identified by the characteristic 'Shmoo' formation. Once 90% of cells showed the schmoo shape, samples were withdrawn and the remaining culture was centrifuged at 3000rpm. The supernatant was discarded and cell pellet was resuspended in fresh YPDA medium, to remove traces of α-factor.

7.7 Serial dilutions and spot assay:

The required strains were allowed to grow on the YEPD agar media in an incubator at 25^{o}C for 24-28 hrs. Taking culture from solid media the cells were then grown in 200 µl of YPDA at 25°C in a 96-multiwell plate over night to reach stationary phase (plateau). 10 fold serial dilutions were performed. Using a replica plater, the contents of the 96-multiwell plate were plated on YPDA medium, other selective media (YNB-URA, YNB-HIS, YNB-TRP etc) YPDA containing HU at the indicated concentrations and incubated at required temperatures.

7.8 TCA protein extraction:

10 ml of a 1x10^{7} cells/ml culture was collected in 15 ml falcon tubes at various time points, centrifuged for 3 minutes at 4000 rpm, resuspended in 2 ml of 20% TCA (Trichloroacetic Acid). The contents were transferred into a 2 ml eppendorf tube and centrifuged for 1 minute at maximum speed. The pellet was resuspended in 100 µl of 20% TCA and glass beads were added leaving a layer of liquid over the beads. The tubes were vortexed for 3 minutes to allow the cells to break open followed by addition of 200 µl of 5% TCA to the mixture (to have final volume of 300 µl 10% TCA). The liquid phase was transferred to a new 1.5 ml eppendorf tube and centrifuged at 3000 rpm for10mins. The supernatant was discarded and the pellet was allowed to completely dry off. The pellet was resuspended in 100 µl Laemli Buffer 1X plus 50 µl of 1M Tris Base to neutralize the acidic pH. The contents were resuspended by vortexing, boiled at 95°C for 3 minutes followed by centrifugation at 3000 rpm for 10 minutes. The supernatant containing the protein was transferred into a new 1.5 ml eppendorf tube and stored at -20°C.

7.9 SDS-PAGE and Western Blot analysis:

It is a technique used for separating proteins based on the difference in their molecular weight. The separation process was performed in denaturing conditions on a polyacrilamide matrix. Depending on the size of the protein being analyzed specific percentages of acrylamide and bisacrylamide were used. The relation between the two parameters being; bigger (heavier) the protein, lower the percentage of acrylamide/bisacrylamide in the gel. The gel apparatus was placed in SDS-PAGE running buffer and in the presence of electric field the proteins migrate through the gel according to their molecular weights; heavier proteins travelling less compared to the lighter ones. The proteins in the gel were transferred onto a porous nitrocellulose membrane by electric transfer in Transfer buffer. The quality of the transfer was checked by Ponceau-S staining. The stained membrane was washed with 1% Tween-TBS 1X (T-TBS) and highly reactive protein epitopes were blocked with 4% milk solution in TBS 1X-0.2% TritonX-100 for 1hour at RT. After blocking, the membrane was incubated with a milk solution containing the specific primary antibody (12CA5 for HA epitope, V5 for PK epitope and FLAG antibodies for the corresponding tags) at the appropriate dilution for 2 hours. After incubation with the primary antibody, the membrane was rinsed and washed twice with T-TBS 1X for 10 minutes and further incubated with the secondary antibody which is conjugated to alkaline peroxidise, for 1 hour at RT (anti-mouse-IgG unless differently stated). Post incubation with the secondary

antibody the membrane was washed twice in T-TBS 1X (10 mins each). Following the washes the membrane was incubated in a solution containing the substrate for 1 min (the chemoluminiscent reaction). (Amersham™ ECL™ Western Blotting Detection Reagents by GE Healthcare). The membrane was then exposed to photographic films and developed.

7.10 FACS analysis:

Fluorescence Activated Cell Sorting (FACS) technique was used to determine the DNA content of cells by analyzing the fluorescence signal of propidium iodide intercalated with the DNA. For this type of analysis, 2 ml was withdrawn at various time points from exponentially growing cell culture (Cell conc. =$1x10^7$ cellules/ml) and centrifuged for 1 minute at 13,300 rpm. Samples were then resuspended in 1 ml of cold 70% EtOH-250mM Tris-HCl pH 7.6 and incubated for 1 hour at room temperature. Samples were centrifuged at 13,300 rpm for 1 minute, resuspended in a solution containing 450 µl of 50mM Tris-HCl pH 7.5 plus and 50µl RNAse 10 mg/ml and further incubated at 37°C for 1 hour. Post treatment with RNAse, cells were centrifuged and resuspended in 500 µl of FACS buffer with 50 µl IPr (propidium iodide 0.5 mg/ml). Propidium iodide is a DNA intercalating agent which fluoresces when activated by the laser of the flow cytometer where, the intensity of fluorescence produced is proportional to the DNA content. 200 µl of cell samples treated with IPr solution were transferred to specific tubes for FACS analysis containing 1 ml of 50mM Tris-HCl pH 7.6. Each sample was sonicated for 15 seconds at 40% of power to separate cells and then read by the flowcytometer.

7.11 Co-immunoprecipitation assay:

This technique was used to study protein-protein interactions from a whole cell extract.

1st day

• Antibody-bound magnetic beads preparation:

20 µL of magnetic beads with protein G (Dynabeads® by Life Technologies) were used per 100 mL of culture. The beads were washed twice with 1 mL of PBS 1X/BSA 5mg/mL and resuspended in the appropriate final volume of the same solution (PBS 1X/BSA 5mg/mL). We added 7.5 µg of the specific antibody each 20 µL of beads and incubated O/N in pre-lubricated Costar tubes on a wheel at 4°C.

• Cell lysis:

We started with 100 mL of a culture at a concentration of $1x10^7$ cells/mL. Cells were centrifuged 5 minutes at 5000 rpm and resuspended in 500 µL of Lysis Buffer supplemented with 2X Protease Inhibitor (IP) and 10mM PMSF. The resuspension was then split into 4 O-ring tubes and 500 µL of glass beads were added. Cells were broken using the fast-prep machine by alternating 5 times 30 seconds of breakage at 4.5 power and 1 minute on ice. After the breakage it was checked that at least 90% of cells was lysed; the supernatant was collected into new 1.5 mL Eppendorf tubes and centrifuged twice 5 minute at maximum speed at 4°C to clarify the cell

extracts. Protein concentration of the samples was measured by spectrophotometer at 595 nm. Cell extracts were stored at -80°C or immediately used for immunoprecipitation.

2nd day

We used 1 mg of protein for each CoIP in a final volume of 300 µL of Lysis Buffer supplemented with IP 2X and 10mM PMSF. Before adding the beads previously washed twice with 1 mL of PBS 1X/BSA 5mg/mL, 3.75 µL of extract were taken as WCE (Whole Cell Extract) sample. After adding 6.7 µL of Laemly 3X and 9.55 µL of H2O, the WCE sample was stored at 4°C untill loading. The extract was incubated from 4 hours to O/N at 4°C on a steering wheel. The time of incubation for the Co-immunoprecipotations performed is 4 hours unless differently noted. After incubation, we took the SUP (SUPernatant) sample (3.75 µL of the extract + 6.7 µL of Laemly 3X + 9.55 µL of H2O) and the beads were washed with increasing concentration of salt to reduce the unspecific binding to the antibody: twice with 1 mL of Lysis Buffer, twice with 1 mL of Lysis Buffer supplemented with 72 µL/mL 5M NaCl, twice with 1 mL of Wash Buffer and one last wash with 1 mL of TE 1X. All the washes were performed in a magnetic greed and the liquid is taken away with the vacuum pump except for the 1 mL of TE 1X which was eliminated with the pipette. The tubes were centrifuged 3 minutes at 3000 rpm and placed back into the magnet in order to eliminate all the TE 1X. The beads were finally resuspended in 20 µL of Laemly Buffer 1X, boiled with WCE and SUP samples 5 minutes at 95°C and loaded on acrylamide gel.

7.12 Neutral/Neutral 2D gel electrophoresis analysis:

Two-dimensional electrophoresis is an important technique used to analyse the replication intermediates (RI) of a specific DNA fragment. The arrays of signals detected by neutral/neutral bi-dimensional gel electrophoresis (**Figure 7.1**) are described as followed:

- **Monomer spot:** DNA fragments that have not been replicated: these are the smallest fragments with the simplest structure and thus traverse less distance in the gel.

- **Bubble arc:** These signify fragments from the replication bubble which are formed by the firing of an active origin in the region being analyzed. With the progression of replication, the bi-directional movement of the fork produces structures with increased mass and complexity.

- **Y's arc:** When the replication bubble formed is not perfectly in the centre of the fragment being analyzed or the replication fork progresses asymmetrically, one of the fork formed from the bubble exits the restriction fragment before the other. This generates a structure resembling the letter 'Y'. A Y-shaped RI can also be generated in a passively replicated region. Depending on the dimensions of the Y-shaped molecules, migration along the big Y's or small Y's arc takes place.

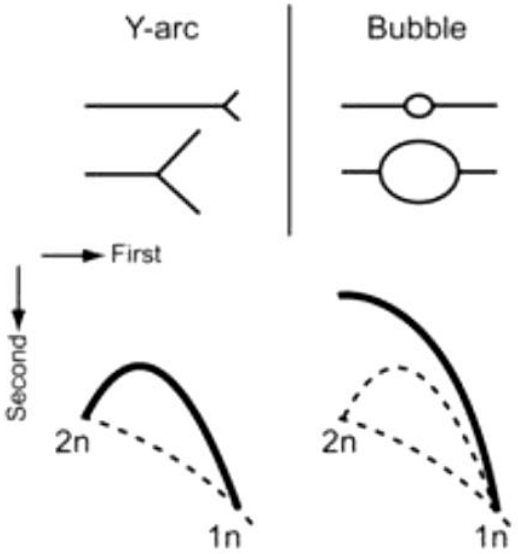


Figure 7.1 Schematic representation of the replication intermediates detectable by 2D gel analysis. Replication intermediates (RI) detected using the 2D gel technique. See text for details.

Procedure for DNA extraction and psoralen crosslink:

Required strains were grown in suitable culture media to a concentration of $1x10^7$cells/ml, samples for DNA extraction were withdrawn from these cultures. Withdrawn culture containing cells were treated with 2 ml of 10% Na-azide and kept on ice (post this step all the steps were performed keeping samples on ice). Samples were centrifuged at 5000 rpm (JA-14 Beckman tubes) for 5 minutes at 4°C. The cell pellet formed was washed with 20 ml of cold water and centrifuged again to obtain the cell pellet. The pellet was dried and resuspended in 5 ml of ice-cold water and processed for psoralen crosslinking. Most commonly used psoralen for *in vivo* crosslinking of DNA is Trimethylpsoralen (TMP). Psoralen intercalates with DNA and when irradiated with UV light (366 nm) it forms covalent crosslinks between the pyrimidines of opposite strands. Psoralen derivatives easily penetrate the membrane of living cells. The cell suspension obtained previously was transferred into a 6-well plate to which 300 µl of 0.2 mg/mL psoralen solution was added and the plate was kept in dark for 5 minutes. Cells were then irradiated with UV radiation in a stratalinker at 365 nm for 10 minutes, at a distance of 2-3 centimetres from the light source. Adding of psoralen and irradiation with UV light was repeated 3 times with 5-minute intervals in the dark for 1-hour total time. Cells suspension was transferred to a 50 ml falcon tube and the 6-well plate was washed with 5ml of cold water twice to collect the remaining cells. The cell suspension in the falcon was centrifuged at 4000 rpm for 3 mins, supernatant was discarded and the pellet was dried and stored in at -20°C (later to be used for DNA extraction). To extract DNA, the cell pellet stored in at -20^{o}C was thawed, followed by addition of 5 ml NIB Buffer and an equal amount of autoclaved glass beads. Each tube was vortexed for 30 seconds at maximum speed and then put on ice for 30 seconds; this intermittent vortexing followed by incubation on ice was repeated 15 times. Using *pasteur* pipettes the cell suspension was collected (leaving the glass beads in the falcon) and transferred to a fresh falcon. The glass beads were washed with 5ml of NIB Buffer twice and using the *pasteur* pipettes the solution was transferred into the fresh falcon from the previous step. The cell suspension was centrifuged for 10 minutes at 8000 rpm at 4°C, the pellet obtained was resuspended in 5 ml of Buffer G2 (Quiagen Genomic DNA extraction kit) containing 100 µl of RNAse A 10 mg/ml and

left for at least 30 minutes at 37°C. Post incubation with RNAse, 100 µl of Proteinase K was added and further incubated for 1 hour at 37°C. The lysates were centrifuged at 5000 rpm, 4°C for 5 mins and the supernatant was diluted in 5 ml of QBT Buffer (Quiagen Genomic DNA extraction kit). Quiagen tip 100G anion exchange columns were set up and pre-equilibrated with 4 ml of QBT Buffer. The samples containing DNA were loaded onto the pre-equilibrated columns and after the sample solution had passed through the membrane of the columns were washed twice with 7.5 ml of QC Buffer (Quiagen Genomic DNA extraction kit). The DNA was then eluted into corex glass tubes with 5 ml of QF Buffer (Quiagen Genomic DNA extraction kit), pre-heated at 50°C. DNA was precipitated by adding 3.5 ml of isopropanol , centrifuged at 8000 rpm(4°C) for 25 mins. The supernatant was transferred into a new corex tube and put at -20°C over night (O/N) and processed the next day as just mentioned for residual DNA content. The dry pellet was resuspended in 150 µl of sterile TE 1X, and left O/N in agitation at room temperature. The 300µl (150 µl +150 µl) of DNA extracted was stored at 4°C.

DNA digestion and precipitation:

Depending on the experiment 10 - 5 µg of DNA was added into an eppendorf along with BSA 1X, the other appropriate enzymes, buffers and Mili-Q water to make up the voulume to a minimum of 150µl. The mixture containing DNA was digested overnight. For our analysis the DNA was digested with HindIII and PstI. The digested DNA was precipitated by adding 1/8 of the volume of 2.5M potassium acetate (CH_3CO_2K) pH6.0 and 1 volume of isopropanol. The contents were mixed by inverting the tubes and then centrifuged at maximum speed for 10 minutes. The pellet obtained was then washed with 500 µl of 75% EtOH centrifuged and supernatant was discarded. Once dry, the pellet was resuspended in 20 µl of sterile TE 1X and kept on shaker for at least 1 hour. Before loading into the gel 5 µl of loading dye 20X was added to the sample.

DNA electrophoresis:

The first dimension electrophoresis was carried out in a gel with agarose conc. of 0.35% in TBE 1X without ethidium bromide placed in buffer TBE 1X. The gel was solidified at 4°C 30 mins before the samples were to be loaded. Before loading the DNA samples into the designated wells for them, 20 µl of loading dye 1X was loaded to check if the wells were intact (between two samples space of atleast one well was kept). The dye in the wells was cleaned using pasteur pipette and the samples were loaded. The first dimension was run at room temperature at 75 Volts constant for 20 hours. This first run separates the fragments according to the mass. The gel was then stained with 0.3 µg/ml ethidium bromide for 30 minutes. Once stained, the gel containing linear and the replicated fragments were cut into individual strips for each sample. The gel strips were then rotated at 90° with respect to the direction of the first dimension and then placed in the second dimension **(Figure 7.2)**. The second dimension run was carried out in conditions that maximize the resolution of shape complexity. The parameters which favour this are; high agarose concentration (0.9%), high voltage (250 Volts) and ethidium bromide (0.3 µg/ml). Gel for the second dimension was poured around the strips of first dimension gel at RT and run at 4°C in TBE 1X buffer containing ethidium bromide (0.3 µg/ml). The gel was constantly checked on the UV illuminator to see how far the DNA had run in the second

dimension. The run was stopped when the DNA line was 1 cm distance from the edge of the gel (approximately 4 hours and 30 minutes of running). The second dimension gel was then cut into two rectangles with two DNA samples in each, which were further used for Southern blot.

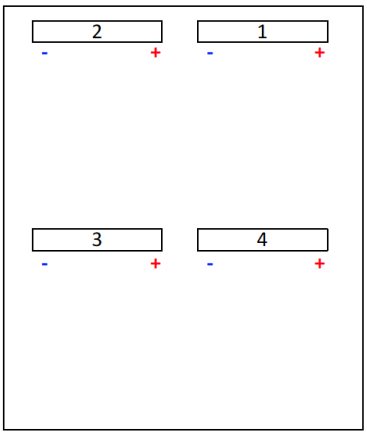


Figure 7.2. Schematic representation of the second dimension preparation. See text for details.
The – and + represents the direction in which the first dimension was run, where the smaller fragments are towards the + sign and the larger fragments towards the – sign.

7.13 Southern blot and hybridization procedure:

For efficient transfer of DNA from the gel to the membrane the cross linking achieved by psoralen needs to be reversed. This was achieved by irradiating the gel with 265 nM UV lamps for 10 mins. Post the irradiation washing steps were carried out to ensure efficient DNA transfer. The gel was first incubated with 0.25 N HCl for 5 mins on a shaker, followed by a quick wash with miliQ. The gel was then incubated with Denaturing solution for 30 mins and then with Blot#2 for 30 mins. The Gene Screen neutral transfer membrane was equilibrated in SSC 10X. The southern blot apparatus was set up and transference of DNA from the gel to the membrane was allowed to take place O/N. Post the transference membranes were auto crosslinked using stratalinker to fix DNA onto the membrane by exposing to 265 nM UV light. The membranes were then incubated in Pre hybridizing solution 1X (PerfectHybTM Plus Hybridisation Buffer by Sigma) until the specific radio labelled probes were prepared for hybridization step.
50 ng of the specific purified DNA was labelled with 50 μCi of P^{32} dCTP. The probe is usually synthesized in 1 hour at RT using the Kleenow polymerase (Prime-a-GeneR Labeling System by Promega). The synthesized probe was purified by passing through Sephadex DNA grade resin columns (illustra™ MicroSpin™ G-50 Columns by GE Healthcare) to remove the non-incorporated nucleotides. Before the membranes were hybridized with the probe, the tube containing the membranes in pre hybridization solution was incubated for at least 1 hr at 65^{o}C. The membranes were then hybridized with the specific radio labelled probe O/N in the hybridization tube. Post hybridization the membranes were incubated for 15mins with 450 ml of Washing solution-I preheated at 65^{o}C in a shaking tray followed by incubation for 15 mins twice

with 500 ml of Washing solution-II, preheated at 42^{o}C. The membranes were then transferred into the Phosphorimager screen cassette and kept for 48 hrs exposure.
The IR signals on the screen were analyzed by FujiFilm Image Analyzer FLA-3000 series.

7.14 Chloroquine two-dimensional gel electrophoresis technique:

This electrophoresis technique is basically used to understand the topological state of a DNA molecule so as to differentiate between supercoiled and relaxed DNA states. To separate highly supercoiled DNA topoisomers, agarose gel electrophoresis is performed in the presence of an intercalator i.e. Chloroquine. By unwinding the double helix, chloroquine alters the supercoiled state of the topoisomers, allowing their resolution in a gel. In the first-dimensional electrophoresis using chloroquine a ladder of DNA bands is observed separating highly supercoiled from relaxed DNA topoisomers in an agarose gel. However, the resolution is not sufficient to separate topoisomers with high supercoiling density which co-migrate in the first dimension i.e. both positively and negatively supercoiled topoisomers. By running a second-dimensional agarose gel electrophoresis with higher chloroquine concentration a difference between the positive & negative supercoiled DNA molecules can be identified. The second dimension is run perpendicular to the first one. Since chloroquine unwinds DNA, negatively supercoiled topoisomers become less supercoiled and migrate more slowly, while positively supercoiled ones gain extra supercoils and migrate more rapidly. Consequently, mobilities of previously co-migrated topoisomers now separate differentially to form an arch-like shape **(Figure 7.3)** where the right arm represents positively supercoiled topoisomers and the left arm corresponds to negatively supercoiled topoisomers.

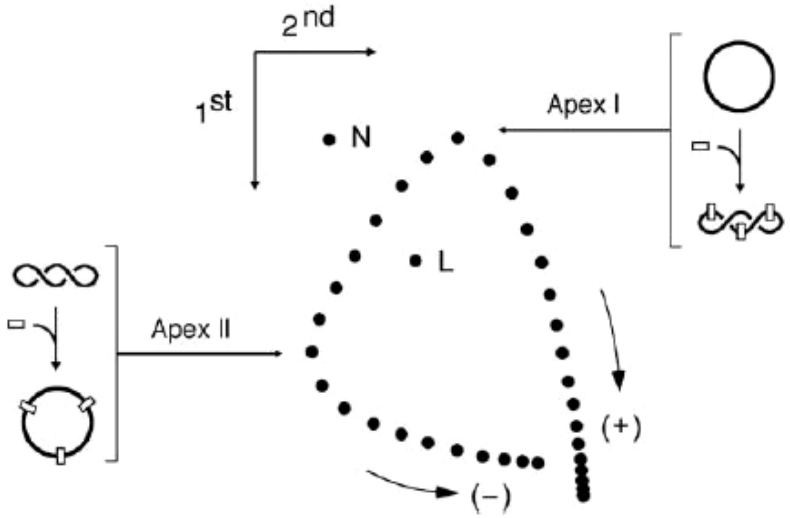


Figure 7.3. Schematic representation of the migration of topoisomers in two dimension chloroquine gels. The dots show the accumulation of (+) and (-) supercoiled DNA. The Apex I indicates the relaxed circle DNA, which on binding with the intercalator (rectangular box) started to get more supercoiled and

thus migrate faster in 2nd dimension. Apex II shows the previously (-) supercoiled DNA which became more relaxed by binding of intercalator and thus migrated slower in the 2nd dimension.

Minichromosome extraction from yeast cells using spheroplast method:

Inoculum of yeast strains was prepared in 20ml with 40µl tetracycline (5mg/ml) and was allowed to grow overnight. 10ml of culture was collected with approx $1x10^7$ cells/ml and 10ml of Toluene solution (prepared prior and kept at -20°C) was added to the collected cell culture. The cell suspension along with toluene was centrifuged at 4000 rpm for 5 mins at 4°C. The cell pellet was washed with 10 ml of cold TE 1X solution (to remove residual toluene) and centrifuged at 4000 rpm for 5 mins at 4°C. The cell pellet was resuspended in 1ml of spheroplast solution (prepared just prior to the experiment) and was transferred to a 1.5 ml microfuge tube.

The cell suspension was incubated at 37°C for 15 mins, followed by centrifugation at 2000 rpm for 5 mins at RT (room temperature). Supernatant was discarded and the spheroplast was resuspended in 300µl of TE 1X. To lyse spheroblasts 30µl of 10% SDS were added to the suspension, mixed gently and left undisturbed for 5 mins at RT. Post incubation, 200µl of 5M KAc was added, mixed well and the mixture was centrifuged at 13,300 rpm for 5 mins. The supernatant was transferred to a new 2ml microfuge tube where 1.2ml of cold 100% EtOH is added. Let it stand for 10 mins at -20°C. Post incubation centrifuge at 13,300 rpm for 10 mins at 4°C. A white pellet should be visible at this point. The supernatant was discarded carefully and the pellet was washed with 1ml of cold 70% EtOH, followed by centrifugation for 5 mins at 13,300 rpm at RT. Supernatant was discarded, pellet was dried, resuspended in 100µl of TE 1X + 1mg/ml RNaseA (10mg/ml) and incubated for 20-30 mins at RT. Post incubation add 60.6µl 5M KAc + 363.6µl of cold 100% EtOH and centrifuge at 13,300 rpm for 5 mins at 4°C. Supernatant was discarded and pellet washed with 1ml 70% EtOH and centrifuged at 13,300 rpm for 5 mins at 4°C. The supernatant was discarded and the pellet was dried and dissolved in 25µl TE 1X with 6.25ul of Chloroquine Loading Buffer (5X) and directly loaded in the 1st dimension agarose gel.

Chloroquine Electrophoresis:

The first dimension gel of the required consistency (i.e 0.6% agarose) was prepared, by weighing 3g and adding in 500 ml of TBE 1X buffer. The suspension was heated till the agarose was completely dissolved and placed in a water bath for 10 mins set at 65°C. The casting tray was prepared by sealing the sides with tape and the combs required for optimum resolution were cleaned thoroughly. Prior to casting, chloroquine from stock solution (1 mg/ml) was added to the gel solution at a final concentration of 0.6 µg/ml (i.e 300 µl of stock chloroquine for 500ml). The prepared gel solution was poured into the casting tray avoiding bubble formation, especially near the comb. Once the gel solution had been poured and left to solidify, the casting tray was

covered to avoid any exposure to light. The gel was allowed to solidify for at least an hour before being transferred into the electrophoresis tank filled with TBE 1X (3 litres). To the TBE 1X buffer in the tank, stock chloroquine was added such that the final concentration was the same as that in the gel (0.6 µg/ml). 31.25µl of DNA sample prepared before was added into the designated well (4 samples were run in one gel). Once the samples had been loaded the apparatus was covered to avoid exposure to light. The gel was run at 50 V for 20 hrs at RT (depending on the size of the minichromsome of study). After the first dimension run, the gel was trimmed by cutting the side above the wells. The gel was then stained in TBE 1X containing 3 µg/ml chloroquine for 1 hour to equilibrate the gel before running the second dimension in higher chloroquine concentration (while equilibration the gel was covered to avoid exposure from light). The orientation of the gel was changed by rotating it to the right by 90^{o}. Fresh TBE 1X buffer (3litres) was poured into the tank with chloroquine concentration of 3µg/ml. Once the gel had been set up, the apparatus was covered to avoid exposure from light. The samples were run in second dimension at 80V for 4 hour at RT.

After the second dimension the Southern Blot and hybridization techniques are performed similar to those explained above in section 7.13.

BIBLOGRAPHY

- **A, Azvolinsky.; P,G, Giresi.; J,D, Lieb.; V,A, Zakian. Highly transcribed RNApolymerase II genes are impediments to replication fork progression in Saccharomyces cerevisiae, *Mol. Cell*. 2009; 34, 722–734.**

- **Aguilera, A. Cotranscriptional mRNP assembly: from the DNA to the nuclear pore. *Curr Opin Cell Biol.* 2005; 17, 242-250.**

- **Aguilera, A.; Gomez-Gonzalez, B. Genome instability: a mechanistic view of its causes and consequences. *Nat. Rev. Genet.* 2008; *9*, 204-217.**

- **Aurèle Piazza et al. Short loop length and high thermal stability determine genomic instability induced by G-quadruplex-forming minisatellites. *EMBO J*, 2015; 34, 1718-1734.**

- **Aygün, O.; Svejstrup, J. Q. RECQL5 helicase: Connections to DNA recombination and RNA polymerase II transcription. *DNA Repair*, 2010; *9*, 345-353.**

- **Azvolinsky, A.; Dunaway, S.; Torres, J.Z.; Bessler, J.B.; Zakian, V.A. The *S. cerevisiae* Rrm3p DNA helicase moves with the replication fork and affects replication of all yeast chromosomes. *Genes Dev.* 2006; *20*, 3104-3116.**

- **Azvolinsky, A.; Giresi, P.G.; Lieb, J.D.; Zakian, V.A. Highly transcribed RNA polymerase II genes are impediments to replication fork progression in *Saccharomyces cerevisiae*. *Mol. Cell* 2009; 34, 722–734.**

- **Balasubramanian, S.; Hurley, L.H.; Neidle, S. Targeting G-quadruplexes in gene promoters: a novel anticancer strategy. *Nat Rev Drug Discov*. 2011; 10, 261–275.**

- **Baldacci, G.; Chérif-Zahar, B.; Bernardi, G. The initiation of DNA replication inthe mitochondrial genome of yeast. *EMBO J.* 1984; 3, 2115–2120.**

- **Bang, I. Untersuchungen über die guanylsäure. *Biochemische Zeitschrif*, 1910; 26, 293–311**

- **Bates. A.; Maxwell. A.; DNA Topology. 2005 Oxford University Press.**

- **Bermejo, R. et al. The replication checkpoint protects fork stability by releasing transcribed genes from nuclear pores. *Cell* 2011; 146, 233–246.**

- **Bermejo, R.; Lai, M.S.; Foiani, M. Preventing replication stress to maintain genome stability: resolving conflicts between replication and transcription. *Mol. Cell* 2012; 45, 710–718.**

- **Bernstein, B.E.; Birney, E.; Dunham, I.; Green, E.D.; Gunter, C.; Snyder, M.; Consortium, E.P. An integrated encyclopedia of DNA elements in the human genome. *Nature*, 2012; 489, 57–74.**

- **Besnard, E.; Babled, A.; Lapasset, L.; Milhavet, O.; Parrinello, H.; Dantec, C.; Marin, J.M.; Lemaitre, J.M. Unraveling cell type-specific and reprogrammable human replication origin signatures associated with G-quadruplex consensus motifs. *Nat Struct Mol Biol.* 2012; 19, 837– 844.**

- **Biffi, G.; Tannahill, D.; McCafferty, J.; Balasubramanian, S. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat Chem.* 2013; 5, 182–186.**

- **Brewer, B. J. When polymerases collide: replication and the transcriptional organization of the E. coli chromosome. *Cell* 1988; 53, 679–686.**

- **Brewer, B. J.; W. L. Fangman. A replication fork barrier at the 3' end of yeast ribosomal RNA genes. *Cell* 1988; 55, 637–643.**

- **Brooks, T. A.; Kendrick, S.; Hurley, L. Making sense of G-quadruplex and i-motif functions in oncogene promoters. *FEBS J.* 2010; 277, 3459–3469.**

- **Bugaut, A.; Balasubramanian, S. 5'-UTR RNA G-quadruplexes: translation regulation and targeting. *Nucleic Acids Res.* 2012; 40, 4727–4741.**

- **Bzymek, M.; Thayer, NH.; Oh, SD.; Kleckner, N.; Hunter, N. Double Holliday junctions are intermediates of DNA break repair. *Nature*, 2007; 464, 937–941**

- **Calladine CR.; Drew HR. Understanding DNA: The Molecule and How it Works. 1997. 2nd edn. Sandiego, USA: Academic Press.**

- **Champoux, J. J. DNA topoisomerases: structure, function, and mechanism. *Annu. Rev. Biochem.* 2001; 70, 369–413.**
- **Chávez ,S.; Beilharz, T.; Rondón, A.G.; Erdjument-Bromage, H.; Tempst, P.; Svejstrup, J.Q et al. A protein complex containing Tho2, Hpr1, Mft1 and a novel protein, Thp2, connects transcription elongation with mitotic recombination in Saccharomyces cerevisiae. *EMBO J.* 2000; 19, 5824–34.**

- **Chávez, S. A protein complex containing Tho2, Hpr1, Mft1 and a novel protein, Thp2, connects transcription elongation with mitotic recombination in Saccharomyces cerevisiae. *EMBO J.* 2000;19, 5824-34.**

- **Chen, Y.; Yang, D.Z. Sequence, stability, and structure of G-quadruplexes and their interactions with drugs. *Curr Protoc Nucleic Acid Chem*. 2012; 50, 17.15.11–17.15.17.**

- **Christiansen, J.; Kofod, M.; Nielsen, F.C. A guanosine quadruplex and two stable hairpins flank a major cleavage site in insulin-like growth factor II MRNA. *Nucleic Acids Res.* 1994; 22, 5709–5716.**

- **Costa, A.; Ilves, I.; Tamberg, N.; Petojevic, T.; Nogales, E.; Botchan, M.R.; Berger, J.M. The structural basis for MCM2-7 helicase activation by GINS and Cdc45. *Nat. Struct. Mol. Biol.* 2010; 18, 471–477**

- **Darzacq, X. *et al. In vivo* dynamics of RNA polymerase II transcription. *Nat. Struct. Mol. Biol.* 2007; 14, 796–806.**

- **De Zamaroczy, M.; Faugeron-Fonty, G.; Baldacci, G.; Goursot, R.; Bernardi. G. The ori sequences of the mitochondrial genome of a wild-type yeast strain: number, location, orientation and structure. *Gene*. 1984. Dec; 32, 439-57.**

- **Drolet, M. Growth inhibition mediated by excess negative supercoiling: the interplay between transcription elongation, R-loop formation and DNA topology. *Mol. Microbiol.* 2006; 59, 723-730.**

- **Drygin, D.; Siddiqui-Jain, A.; O'Brien, S.; Schwaebe, M.; Lin, A.; Bliesath, J.; Ho, C.B.; Proffitt, C.; Trent, K.; Whitten, J.P.; Lim, J.K.C.; von Hoff, D.; Anderes, K.; Rice, W.G. Anticancer activity of CX-3543: a direct inhibitor of rRNA biogenesis. *Cancer Res.* 2009; 69, 7653–7661.**

- **Eddy, J.; Maizels, N. Gene function correlates with potential for G4 DNA formation in the human genome. *Nucleic Acids Res.* 2006; 34, 3887–3896.**
- **El Hage, A.; French, S.L.; Beyer, A.L.; Tollervey, D. Loss of Topoisomerase I leads to R-loop–mediated transcriptional blocks during ribosomal RNA synthesis. *Genes Dev.* 2010; 24, 1546–1558.**

- **Evrin, C.; Clarke, P.; Zech, J.; Lurz, R.; Sun, J.; Uhle, S.; Li, H.; Stillman, B.; Speck, C. A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. *Proc. Natl. Acad.* Sci. USA 2009; 106, 20240–20245.**

- **Fry, M. Tetraplex DNA and its interacting proteins. *Front Biosci.* 2007; 12, 4336–4351.**

- **Gan, W. *et al.* R-loop–mediated genomic instability is caused by impairment of replication fork progression. *Genes Dev.* 2011; 25, 2041–2056.**

- **García-Rubio, ML.; Aguilera, A.; Topological constraints impair RNA polymerase II transcription and causes instability of plasmid-borne convergent genes. *Nucleic Acids Res*. 2012; 40, 1050-64.**

- **Gellert, M.; Lipsett, M.N.; Davies, D.R.; Helix formation by guanylic acid. *Proc Natl Acad Sci USA* 1962; 48, 2013–2018.**

- **Giaever, GN.; Wang, JC. Supercoiling of intracellular DNA can occur in eukaryotic cells. *Cell*. 1988; 55,849-56.**

- **Gilbert, D. E.; Feigon, J. Multistranded DNA structures. *Curr. Opin. Struct. Biol*. 1999; 9, 305–314.**

- **Gomez, D. *et al.* Interaction of telomestatin with the telomeric single-strand overhang. *J. Biol. Chem.* 2004; 279, 41487–41494.**

- **Gozuacik, D.; Chami, M.; Lagorce, D.; Faivre, J.; Murakami, Y.; Poch, O.; Biermann, E.; Knippers, R.; Brechot, C.; Paterlini-Brechot, P. Identification and functional characterization of a new member of the human Mcm protein family: hMcm8. *Nucleic Acids Res.* 2003; 31, 570–579.**

- **Greenfeder, S. A.; Newlon, C. S. Replication forks pause at yeast centromeres. *Mol. Cell. Biol*. 1992; 12, 4056–4066.**
- **Heller, RC.; Marians, KJ. Replication fork reactivation downstream of a blocked nascent leading strand. *Nature*, 2006; 439, 557–562.**

- **Helmrich, A.; Ballarino, M.; Tora, L. Collisions between replication and transcription complexes cause common fragile site instability at the longest human genes. *Mol. Cell* 2011; 44, 966–977.**

- **Hiasa, H.; Marians, K.J.. Topoisomerase III, but not topoisomerase I, can support nascent chain elongation during theta-type DNA replication. *J. Biol. Chem.*1994; 269, 32655–32659.**

- **Hiratani, I. *et al.* Global reorganization of replication domains during embryonic stem cell differentiation. *PLoS Biol.* 2008; 6, 245.**

- **Hobeika, M.; Brockmann, C.; Gruessing, F.; Neuhaus, D.; Divita, G.; Stewart, M.; Dargemont, C. Structural requirements for the ubiquitin-associated domain of the mRNA export factor Mex67 to bind its specific targets, the transcription elongation THO complex component Hpr1 and nucleoporin FXFG repeats. *J Biol. Chem* 2009; 284, 17575-17583.**

- **Houseley, J.; LaCava, J.; Tollervey, D. RNA-quality control by the exosome. *Nat Rev Mol Cell Biol.* 2006; 7, 529–39.**

- **Huberman, J.A.; Riggs, A.D. Autoradiography of chromosomal DNA fibers from Chinese hamster cells. *Proc. Natl. Acad. Sci. USA* 1966; 55, 599–606.**

- **Huertas, P.; Aguilera, A. Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. *Mol. Cell* 2003; 12, 711-721.**

- **Huppert, J. L.; Balasubramanian, S. G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res.* 2007; 35, 406–413.**

- **Huppert, J. L.; Bugaut, A.; Kumari, S.; Balasubramanian, S. G-quadruplexes: the beginning and end of UTRs. *Nucleic Acids Res.* 2008; 36, 6260–6268.**

- **Itoh, T.; Tomizawa, J. Formation of an RNA primer for initiation of replication ofColE1 DNA by ribonuclease H. *Proc. Nat. Acad. Sci. U.S.A.* 1980; 77, 2450–2454.**

- **Ivessa, A. S., J. Q. Zhou, and V. A. Zakian. The Saccharomyces Pif1p DNA helicase and the highly related Rrm3p have opposite effects on replication fork progression in ribosomal DNA. *Cell* 2000; 100, 479–489.**

- **Ivessa, A. S.; Lenzmeier, B. A.; Bessler, J. B.; Goudsouzian, L. K.; Schnakenberg, S.L.; Zakian, V. A. The Saccharomyces cerevisiae helicase Rrm3p facilitates replication past nonhistone protein-DNA complexes. *Mol. Cell* .2003; 12, 1525–1536.**

- **Ivessa, A.S. *et al.* The *Saccharomyces cerevisiae* helicase Rrm3p facilitates replication past nonhistone protein-DNA complexes. *Mol. Cell* 2003; 12, 1525–1536.**

- **Jiang, X.; Klimovich, V.; Arunkumar, A.I.; Hysinger, E.B.; Wang, Y.; Ott, R.D.; Guler, G.D.; Weiner, B.; Chazin, W.J.; Fanning, E. Structural mechanism of RPA loading on DNA during activation of a simple pre-replication complex. *EMBO J.* 2006; 25, 5516–5526.**

- **Joaquim Roca et al. The torsional state of DNA within the chromosome. *Chromosoma* 2011; 120, 323–334.**

- **John, L, Nitiss. Targeting DNA topoisomerase II in cancer chemotherapy. *Nat Rev Cancer*. 2009; 9, 338–350.**

- **Johzuka, K.; Horiuchi, T. Replication fork block protein, Fob1, acts as an rDNA region specific recombinator in *S. cerevisiae*. *Genes Cells* 2002; *7*, 99-113.**

- **Judith Lopes et al. G-quadruplex-induced instability during leading-strand replication. *EMBO J*. 2011 Aug 26; 30, 4033-46.**

- **Kireeva, M. L. et al. Nucleosome remodeling induced by RNA polymerase II: loss of the H2A/H2B dimer during transcription. *Mol. Cell* 2002; 9, 541–552.**

- **Kobayashi, T. The replication fork barrier site forms a unique structure with Fob1p and inhibits the replication fork. *Mol. Cell. Biol.* 2003; *23*, 9178-9188.**

- **Kobayashi, T.; Horiuchi, T. A yeast gene product, Fob1 protein, required for both replication fork blocking and recombinational hotspot activities. *Genes Cells* 1996; *1*, 465-474.**

- **Kopytova, D.V.; Orlova, A.V.; Krasnov, A.N.; Gurskiy, D.Y.; Nikolenko, J.V.; Nabirochkina, E.N.; Shidlovskii, Y.V.; Georgieva, S.G. Multifunctional factor ENY2 is associated with the THO complex and promotes its recruitment onto nascent mRNA. *Genes Dev*. 2010; 24, 86-96.**

- **Kornberg, A.; Baker, T. DNA replication, 2nd ed. W. H. Freeman and Co., New York, NY 1992.**

- **Krasilnikova, M. M.; Mirkin, S. M. Replication stalling at Friedreich's ataxia (GAA)n repeats in vivo. *Mol. Cell. Biol.* 2004; 24, 2286–2295.**

- **Krogh, BO.; Shuman, S. Catalytic mechanism of DNA topoisomerase IB. *Mol Cell*. 2000; 5, 1035-41.**

- **Kuryan, B. G. *et al.* Histone density is maintained during transcription mediated by the chromatin remodeler RSC and histone chaperone NAP1 *in vitro. Proc. Natl Acad. Sci. USA* 2012; 109, 1931–1936.**

- **Levine, M. Transcriptional enhancers in animal development and evolution. *Current Biology* 2010; 20, R754–63**

- **Li, M.; Xu, X.; Liu, Y. The Set2-RPB1 Interaction Domain of Human RECQ5 Is Important for Transcription-Associated Genome Stability. *Mol. Cell. Biol.* 2011; *31*, 2090-2099.**

- **Li, Y.; Lin, AW.; Zhang, X.; Wang, Y.; Wang, X.; Goodrich, DW. Cancer cells and normal cells differ in their requirements for Thoc1. *Cancer Res*. 2007; 67, 6657-64.**

- **Lilley, D. M. J. Structural isomerization in DNA: the formation of cruciform structures in supercoiled DNA molecules. *Chem. Soc. Rev.* 1989; 18, 53–83.**

- **Lin, Y.; Dent, S.Y.; Wilson, J.H.; Wells, R.D.; Napierala, M. R loops stimulate genetic instability of CTG.CAG repeats. *Proc. Natl. Acad. Sci. USA* 2010; 107, 692–697.**

- **Linskens, M.H.; J. A. Huberman. Organization of replication of ribosomal DNA in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 1988; 8, 4927–4935.**

- **London, T. B. *et al.* FANCJ is a structure-specific DNA helicase associated with the maintenance of genomic G/C tracts. *J. Biol. Chem.* 2008; 283, 36132–36139.**

- **MacNeill, S.A. Structure and function of the GINS complex, a key component of the eukaryotic replisome. *Biochem. J.* 2010; 425, 489–500.**

- **Maiorano, D.; Cuvier, O.; Danis, E.; Mechali, M. MCM8 is an MCM2-7-related protein that functions as a DNA helicase during replication elongation and not initiation. *Cell* 2005; 120, 315–328**

- **Marcu, K.B.; Bossone, S.A.; Patel, A.J. Myc function and regulation. *Annu. Rev. Biochem.* 1992; 61, 809-812.**

- **Marıa, L.; Garcıa Rubio.; Andres Aguilera. Topological constraints impair RNA polymerase II transcription and causes instability of plasmid-borne convergent *genes*. 2012; 18, 990-921.**

- **Mariappan, S. V.; Silks, L. A. R.; Chen, X.; Springer, P. A.; Wu, R.; Moyzis, R. K.; Bradbury, E. M.; Garcia, A. E.; Gupta, G. Solution structures of the Huntington's disease DNA triplets, (CAG)n. *Biomol. Struct. Dyn.* 1998; 15, 723–744.**

- **Masuda, S.; Das, R.; Cheng, H.; Hurt, E.; Dorman, N.; Reed, R. Recruitment of the human TREX complex to mRNA during splicing. *Genes Dev.* 2005; 19, 1512-1517.**

- **Mechali, M. DNA replication origins: From sequence specificity to epigenetics. *Nat. Rev. Genet.* 2001; 2, 640–645.**

- **Miguel, A.; External conditions inversely change the RNA polymerase II elongation rate and density in yeast. *Biochim Biophys Acta*. 2013;1829, 1248-55.**

- **Mirkin, S. M.; Malvy, A.; Harel-Bellan.; Pritchard, L. L. Structure and biology of H DNA Triple helix forming oligonucleotides. *Kluwer Academic Publishers, Boston,* MA. 1999; 193–222.**

- **Mischo, H. E.; Gomez-Gonzalez, B.; Grzechnik, P.; Rondon, A. G.; Wei, W.; Steinmetz, L.; Aguilera, A.; Proudfoot, N. J. Yeast Sen1 helicase protects the genome from transcription-associated instability. *Mol. Cell* 2011; 41, 21-32.**

- **Mohanty, B.K.; Bairwa, N.K.; Bastia, D. The Tof1p-Csm3p protein complex counteracts the Rrm3p helicase to control replication termination of *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* 2006; *103*, 897-902.**

- **Neidle, S. The structures of quadruplex nucleic acids and their drug complexes. *Curr. Opin. Struct. Biol.* 2009; 19, 239–250.**

- **Nishimura, K.; Ishiai, M.; Horikawa, K.; Fukagawa, T.; Takata, M.; Takisawa, H.; Kanemaki, M.T. Mcm8 and Mcm9 form a complex that functions in homologous recombination repair induced by DNA interstrand crosslinks. *Mol. Cell* 2012; 47, 511–522.**

- **Nishiyama, A.; Frappier, L.; Mechali, M. MCM-BP regulates unloading of the MCM2-7 helicase in late S phase. *Genes Dev*. 2011; 25, 165–175.**

- **Nomura, M.; Morgan, E. A.. Genetics of bacterial ribosomes. *Annu. Rev. Genet.* 1977; 11, 297–347.**

- **Paeschke, K.; Capra, J.A.; Zakian, V.A. DNA replication through Gquadruplex motifs is promoted by the Saccharomyces cerevisiae Pif1 DNA helicase. *Cell* 2011; 145, 678–691.**

- **Paques F.; Haber J,E. Multiple pathways of recombination induced by double-strand breaks in Saccharomyces cerevisiae. *Microbiol Mol Biol Rev*. 1999, 63, 349-404.**

- **Pavri, R. *et al.* Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II. *Cell* 2006*;* 125, 703–717.**

- **Pearson, C. E.; Edamura, K. N.; Cleary, J. D. Repeat instability: mechanisms of dynamic mutations. *Nat. Rev. Genet.* 2005; 6, 729–742.**

- **Pelengaris, S.; Khan, M. The many faces of c-MYC. *Arch. Biochem. Biophys.* 2003; 416, 3450-3462.**

- **Pelletier, R.; Krasilnikova, M. M.; Samadashwily, G. M.; Lahue, R. S.; Mirkin, S. M. Replication and expansion of trinucleotide repeats in yeast. *Mol. Cell. Biol.* 2003; 23, 1349–1357.**

- **Prado, F.; Aguilera, A. Impairment of replication fork progression mediates RNA polII transcription-associated recombination. *EMBO J.* 2005; 24, 1267–1276.**

- **Prunell, A.; A topological approach to nucleosome structure and dynamics: the linking number paradox and other issues. *Biophys J*. 1998; 74, 2531-44.**

- **R, E, Wellinger.; F, Prado.; A, Aguilera. Replication fork progression is impairedby transcription in hyperrecombinant yeast cells lacking a functional THO complex, *Mol. Cell. Biol*. 2006; 26, 3327–3334.**

- **Randell, J.C.; Fan, A.; Chan, C.; Francis, L.I.; Heller, R.C.; Galani, K.; Bell, S.P. Mec1 is one of multiple kinases that prime the Mcm2-7 helicase for phosphorylation by Cdc7. *Mol. Cell* 2010; 40, 353–363.**

- **Reaban, M.E.; Lebowitz, J.; Griffin, J.A. Transcription induces the formation of a stable RNA.DNA hybrid in the immunoglobulin α switch region. *J. Biol. Chem.* 1994; 269, 21850–21857.**

- **Recolin, B.; van der Laan S.; Tsanov, N.; Maiorano, D. Molecular Mechanisms of DNA Replication Checkpoint Activation. *Genes*. 2014; 5, 147-175.**

- **Ribeyre, C. *et al.* The yeast Pif1 helicase prevents genomic instability caused by G-quadruplex-forming CEB1 sequences *in vivo. PLoS Genet.* 2009; 5, 475.**

- **Rich, A.; Nordheim, A.; Wang, A. H.J. The chemistry and biology of left-handed Z-DNA. *Annu. Rev. Biochem.* 1984; 53, 791–846.**

- **Rodriguez, R. *et al.* Small-molecule-induced DNA damage identifies alternative DNA structures in human genes. *Nature Chem. Biol.* 2012; 8, 301–310.**

- **Rougemaille, M.; Dieppois, G.; Kisseleva-Romanova, E.; Gudipati, R.K.; Lemoine, S.; Blugeon, C et al. THO/Sub2p functions to coordinate 3′-end processing with gene-nuclear pore association. *Cell* 2008; 135, 308–21.**

- **S, Tuduri.; L, Crabbé.; C, Conti.; H, Tourrière.; H, Holtgreve-Grez.; A, Jauch.; V, Pantesco.; J, De Vos.; A, Thomas.; C, Theillet. Topoisomerase I, suppresses genomic instability by preventing interference between replication and tran-scription, *Nat. Cell Biol*. 2009; 11, 1315–1324**

- **Sakatsume, O.; Tsutsui, H.; Wang, Y.; Gao, H.; Tang, X.; Yamauchi, T.; Murata, T.; Itakura, K.; Yokoyama, K.K. Binding of THZif-1, a MAZ-like zinc finger protein to the nuclease-hypersensitive element in the promoter region of the c-MYC protooncogene. *J. Biol. Chem.* 1996; 271 , 313-320.**

- **Salceda, J.; Fernández, X.; Roca, J. Topoisomerase II, not topoisomerase I, is the proficient relaxase of nucleosomal DNA. *EMBO J.* 2006; 25, 2575-83.**

- **Samadashwily, G. M.; Dayn, A.; Mirkin, S. M. Suicidal nucleotide sequences for DNA polymerization. *EMBO J.* 1993; 12, 4975–498**

- **Samadashwily, GM.; Raca, G.; Mirkin, SM. Trinucleotide repeats affect DNA replication in vivo. *Nat Genet,* 1997; 17 298–304.**

- **Sanders, C.M. Human Pif1 helicase is a G-quadruplex DNA binding protein with G-quadruplex DNA unwinding activity. *Biochem. J.* 2010*;* 430,119–128.**

- **Schmidt, K.H.; Derry, K.L.; Kolodner, R.D. *Saccharomyces cerevisiae* RRM3, a 5' to 3' DNA helicase, physically interacts with proliferating cell nuclear antigen. *J. Biol. Chem.* 2002; *277*,45331-45337.**

- **Schramm, L.; Hernandez, N. Recruitment of RNA polymerase III to its target promoters. *Genes Dev.* 2002*;* 16, 2593–2620.**

- **Schramm, L.; Pendergrast, P. S.; Sun, Y.; Hernandez, N. Different human TFIIIB activities direct RNA polymerase III transcription from TATA-containing and TATA-less promoters. *Genes Dev.* 2000; 14, 2650–2663.**

- **Sen, D.; Gilbert, W. A sodium-potassium switch in the formation of four-stranded G4-DNA. *Nature* 1990; 344, 410–414.**

- **Sergei M Mirkin. DNA Topology: Fundamentals. University of Illinois at Chicago, Illinois, USA 2001 Nature Publishing Group.**

- **Sheu, Y.J.; Stillman, B. The Dbf4-Cdc7 kinase promotes S phase by alleviating an inhibitory activity in Mcm4. *Nature* 2010; 463, 113–117.**

- **Siddiqui-Jain, A.; Hurley, L.H. DNA structure: visualizing the quadruplex. *Nat Chem.* 2013; 5, 153–155.**

- **Simonsson, T.; Pecinka, P.; Kubista, M. DNA tetraplex formation in the control region of c-myc. *Nucleic Acids Res.* 1998; 26 , 116-122.**

- **Sinden, R. R. DNA structure and function. Academic Press, Inc., San Diego, CA. 1994.**
- **Singh, J.; Padgett, R.A. Rates of in situ transcription and splicing in large human genes. *Nat. Struct. Mol. Biol.* 2009; 16, 1128–1133.**

- **Skourti-Stathaki, K.; Proudfoot, N. J.; Gromak, N. Human Senataxin Resolves RNA/DNA Hybrids Formed at Transcriptional Pause Sites to Promote Xrn2-Dependent Termination. *Mol. Cell* 2011; 42, 794-805.**

- **Stephen P, Bell.; Anindya Dutta. DNA Replication In Eukaryotic Cells. *Annu. Rev. Biochem*. 2002; 71, 333–74**

- **Stirling, P.C. et al. R-loop–mediated genome instability in mRNA cleavage and polyadenylation mutants. *Genes Dev*. 2012; 26, 163–175.**

- **Strässer, K.; Masuda, S.; Mason, P.; Pfannstiel, J.; Oppizzi, M.; Rodriguez-Navarro, S et al. TREX is a conserved complex coupling transcription with messenger RNA export. *Nature* 2002; 417,304–8.**

- **Sun, D.; Hurley, L. H. The importance of negative superhelicity in inducing the formation of G-quadruplex and i-motif structures in the c-Myc promoter: implications for drug targeting and control of gene expression. *J. Med. Chem.* 2009; 52, 2863–2874.**

- **Sundquist, W.I.; Klug, A. Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops. *Nature* 1989; 342, 825–829.**

- **Tanaka, S.; Umemori, T.; Hirai, K.; Muramatsu, S.; Kamimura, Y.; Araki, H. CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast. *Nature* 2007; 445, 328–332.**

- **Tuduri, S. et al. Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription. *Nat. Cell Biol.* 2009; 11, 1315–1324.**

- **Venkatesh, S.; Workman, JL. Histone exchange, chromatin structure and transcription. *Nat Rev Mol Cell Biol*. 2015; 16, 178-189.**

- **Walter, J.; Newport, J. Initiation of eukaryotic DNA replication: Origin unwinding and sequential chromatin association of Cdc45, RPA, and DNA polymerase alpha. *Mol. Cell* 2000; 5, 617–627.**
- **Wells, R. D.; Warren, S. T. Genetic instabilities and hereditary neurological disorders. Academic Press, San Diego, CA. 1998.**

- **Westover, K.D.; Bushnell, D.A.; Kornberg, R.D. Structural basis of transcription: separation of RNA from DNA by RNA polymerase II. *Science* 2004; 303, 1014–1016.**

- **Williamson, J.R.; Raghuraman, M.K.; Cech, T.R. Monovalent cationinduced structure of telomeric DNA: the G-quartet model. *Cell* 1989; 59, 871–880.**

- **Wu, Y.; Shin-ya, K.; Brosh, R. M. Jr. FANCJ helicase defective in Fanconia anemia and breast cancer unwinds G-quadruplex DNA to defend genomic stability. *Mol. Cell. Biol.* 2008; 28, 4116–4128.**

- **Yang, D.; Okamoto, K. Structural insights into G-quadruplexes: towards new anticancer drugs. *Future Med Chem.* 2010; 2, 619–646.**

- **Yasukazu Daigaku et al. A global profile of replicative polymerase usage. *Nature Structural & Molecular Biology.* 2015; 22, 192–198.**

- **Zechiedrich, E.L.; Cozzarelli, N.R. Roles of topoisomerase IV and DNA gyrase in DNA unlinking during replication in *Escherichia coli*. *Genes Dev* 1995; 9, 2859–2869.**

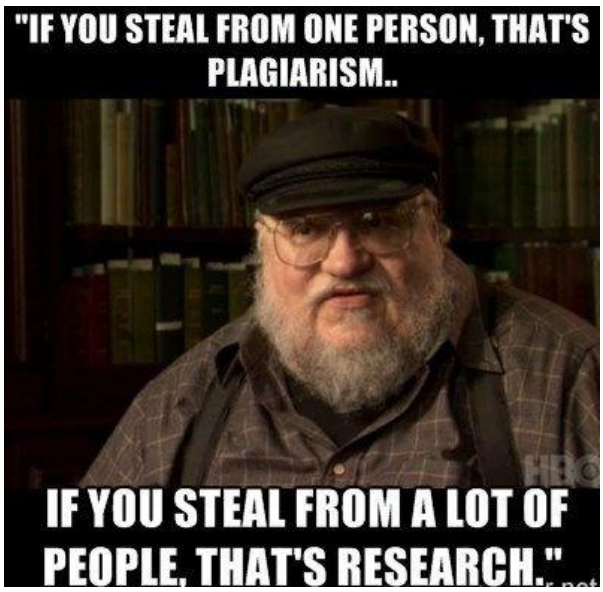
- **Zegerman, P.; Diffley, J.F. Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast. *Nature* 2007; 445, 281–285.**

- **Zenklusen, D.; Vinciguerra, P.; Wyss, JC.; Stutz, F. Stable mRNP formation and export require co-transcriptional recruitment of the mRNA export factors Yra1p and Sub2p by Hpr1p. *Mol Cell Biol*. 2002; 22, 8241-53.**

- **Zheng, M.; X. Huang, G.; Smith, K.; Yang, X.; Gao, X. Genetically unstable CXG repeats are structurally dynamic and have a high propensity for folding. An NMR and UV spectroscopic study. *J. Mol. Biol.* 1996; 264,323–336.**

ACKNOWLEDGEMENT

Well, I think this is the only section in this report where I can present my views in complete "HONESTY"....so to sum it all up:



Disclaimer: The image above states the information which may be perceived in many different forms...no one knows in what context has the author put forward this idea.

Jokes apart I would like to thank a few who all have in good faith tried to make my 4 years a memorable journey in which I achieved my doctorate. As hierarchy and future job references demands, I wish to thank **Dr. Rodrigo Moreno Bermejo** for taking the challenge to prove the world wrong by giving me an opportunity to tax his mind and resources in making research possible. I apologise on personal level for the times I was better at basketball against you and forgot that you were still my boss. But all in all, I did learn a lot from you on both fronts. So, thank you for your utmost patience and guidance in making this happen.

Secondly, I would like to thank my lovely lab mates who gave me an eye opening reality check on how atrociously bad my concepts were and how much hard work was required, which I suppose, did come through in my own Indian way. Firstly, **Dr. Camilla Frattini** who turned out to be not my fellow PhD student but rather my personal "save his thesis" advisor. So thank you for luckily being patient at teaching me and also benefiting to comprehend the Indian version of English speaking. Secondly, **Dr. (very soon to be) Sara Villa** for offering countless times with amazing Spanish food (which I thought was a myth). Your cooking and scientific depth is always what I am a big fan of. Lastly, **Grazia Pellicano and Fabio** from being my first mentor in the

lab to the time of being the only one who matched my level of being "torpe". You guys are a wonderful couple who are blessed with a really cute pseudo baby "Olivia". I wish you both all the success and luck, with my incredible 'Yoga Powers'. Apart these few, I deeply want to thank **Dr. Esther Cabanas** for being my emergency helpline, my lunatic companion, my dance partner, my screwed up Spanish solver, my gossip girl and many more. If it wasn't for you I would have never found a true source of inspiration that I am what it takes to finish this. So thanks a ton. Amongst the other, One boy who stood by me from the very 1st week till now is my brother from another mother **Pablo Tamayo Hermosa.** Man, it's a never ending story if I start on you, but to keep it simple, if it wasn't for you, I would have finished my thesis a year before! So thank you for being the partner in crime in everything and making my 4 years an incredible memory. Last but not the least a big shout out to my Spanish family **Angela Miguel, Sergio & Javier** for all the epic fun and crazy memories we share; which I am sure is still just the beginning. You guys truly speak volumes about what Spanish culture and people is all about, apart from times when Sergio is drunk and talks nothing but how majestic his 'Beep' is.

Finally, I know it's unfair for not mentioning you guys by name but I just can't THANK ALL YOU ERASMUS GUYS enough who all I met in Salamanca from, Belgium, Greece, Japan, Germany, USA, Netherlands, England, France and many more with whom I got shit drunk and shared insane time which is impossible to express in words. P.S. a big shout out to **Raj Koothrappali** from 'Big Bang Theory', because of him I was a great reference point which had countless advantages of being in a foreign land.

With this new paragraph (for better sustenance of my future life) I want to thank the sole person who drove me by my ear; literally and showed me what I can pursue. It's none other than my girlfriend and now WIFE **Dr. Pooja Joshi** who introduced me to the idea of being a Doctorate. She did mind wash all my permutations about future career but I see what she meant then. She cannot be thanked enough for making me also a doctorate, if and when I become one in near future, as she was the sole driving force for me to even 'think' of becoming one. She is a girl with immense talent and one of her biggest strengths is her diligence to know what she wants. Due to which she transforms into Professor X (female version) with unimaginable focus to fathom any herculean task. Though, she is terrible at many more thing but when it comes to science she inspires me and that's why I take this opportunity to thank her with all my heart and love for standing by me through thick and thin and making me what I am. It was not easy for her to do so as I am still distracted and lost on my own but every time she is there to sort my brains out and help me channel my energy where it should be. The best part about all this time we have spent together is, she was there from the beginning I went on trying lab work and failed miserably, but she told that me being less worried can be an amazing attribute to venture in this field of science. It turns out she was right, I'm not really great at Science and neither have a star CV when I finish my PhD. I am still happy and hopeful to see what's next in store for me now. Above all, to have a life partner who shares the same wavelength in personal and professional front while matching the level of stupidity and foolishness, is all I ever wanted to be 'Super

Awesome US'. So thank you Pooja Joshi for being all the above and I hope this makes you happy as I have exceeded a page while thanking you thoroughly.

See it's considered a basic human nature to thank your family and friends but the real story is these people do make a hell load of a difference. Likewise, my childhood idol and mentor who is himself an accomplished doctor with the level of achievements which me and my future child cumulatively won't be able to account for is my father "**Brigadier. Dr. Yashwant Singh Bisht.** He is crazy good in everything I can think of, and always keeps on achieving more and more in life. Figuratively, he is not Bisht, he is a Beast and it's not just me but everyone around him wants to imbibe something or the other to be like him. So, cheers to you baapu you are truly a legend and thank you for everything, hopefully soon your son will also have a pseudo Dr. on his name. Ok now there are some people you can thank in words or gestures or whatever way humanly possible, but then there is this one individual whom none such things work. It is impossible to thank a human being who has not only given birth to you but also made you what you are. It's impossible to pin downs words of appreciation for her, so all I can say is **Mrs. Meena Bisht** thank you for being my mother. You mean so much to us that its stupidity to try and thank you. I love you Maa. Alright too much sentimental values, but I am still not done as I need to thank this guy who has been here with me for the past 6 months, when I was out cold with no one to help and support me or even to talk to, to figure out what is better and what not. He being younger shitty version of me made it perfect to glide through this time in life when I needed him apparently the most so far and thank god he was there. He has been the unpaid maid I so desperately needed, the Fifa buddy to have a kickass stress buster time, the native language speaker when that's all you need to make everything around you fun and much more. He helped literally in this thesis to be compiled in time and even now is concerned whether when will I finish this and sleep. So thank you **Mr. Prashast Bisht** for being there younger brother, you are the best and the worst in right proportions, which will take you to places. Lastly, I am again taking just jotting down names of you all but I am short of time and space so hope you understand. Chronologically thank you **Mr. Ashok Thakur**, you are the best thakur saab one can ever want. You are my uncle but I always found an amazing friend with so much love for me that I cannot express. **Mrs. Vinita Thakur** for being the awesome carefree, jolly hearted, wise lady who has been the example for everything is fine why worry. **Mr. Ajay Kumar Pundir** who is again a 'man of many hats' under whom I learnt many things. You are the beam of our family and we have nothing but respect for you. **Mrs. Gyalmu Pundir** the lady who is our role model of generosity and style. You have always been the one with selfless love for all living things and I thank you for making me inculcate that quality from you. **Maj. Avijit Thakur,** you are that one person with whom I can always blindly trust no matter what may come. You are my brother by pedigree but you are that one factor down the memory lane which has always brought happiness to me time and again. **Mrs. Anisha Mahanta Thakur,** thank you for being a part of my life. I can't thank you enough as the love and happiness you spread around is unaccountable. You are the buttress of our future family tree and your support morally and emotionally I will always cherish in times to come. **Mr. Ravijit Thakur** the actual personified version of what a stud I

always dreamt to be. You are undoubtedly the Starboy who gains accolades in all fronts and I thank you for making me feel proud, popular and young whenever I am with you. **Mr.Shreyyash Pundir**, thank you for being the real gold hearted person you are. I seriously can't believe how you turned out so self-sacrificing with such a bunch of madcap brothers around you. I am proud of you and I love you to core. **Mr. Shreshth Pundir**, you amongst us all is the true lineage of all elder brothers' qualities summarized in one. You are a born talent be in any front, where studies is the least explored territory so far. I thank you for keeping me on my toes always and I hope you beat me in every aspect in life to make me and all of us proud. Above all, Thank you So much **Mr. Ashish Arora, Mr. Shubankar Biswas, Mr. Sanak Mahapatra** and **Mrs. Ashika Arora** without you guys in my life I would have never achieved and will never achieve anything. We are each other's backbone. If I could choose one thing to re-live, it will be the time we all spent together in college. Things change people change but friends like you always makes me live more and more to keep creating memories to fill mine and your children till they die of boredom.

I would like to thank **Mr.Alfredo Ruiz** for being my guardian and a friend during my time and being always there to help and support me. Thank you so much Sir.

In the end I want to thank my elaborated family which has played a very crucial role in making me achieve my goals. Thank you so much **Mr. S.P. Joshi** and **Mrs. Promila Joshi** for welcoming me in your family with open arms and giving me your blessings to achieve things step by step. Finally the two lovely gorgeous hot beautiful girls, **Ms. Sakshi Joshi** and **Ms. Nidhi Joshi** for first bearing with my sad taste in humour and terrible memory. I deeply thank you all for making me a part of your own and showering me with all the undeserved love. I hope and will thrive to make you all equally happy in near future. I earnestly thank god and **Maharaji** for bestowing upon his blessings and almighty wishes due to which I couldn't have even began what I have just finished. So, all in all whole hearted thank you to all of you for helping me accomplish another chapter in life, with all due regards JAI SACCHIDANAND.