ALMA MATER STUDIORUM – UNIVERSITÀ DI BOLOGNA

Dottorato di ricerca in
Biologia Funzionale dei Sistemi Cellulari e Molecolari

XXII CICLO
Settore scientifico disciplinare di afferenza: BIO/11 BIOLOGIA MOLECOLARE

Transcriptional functions of DNA Topoisomerases at a genome-wide scale and a single gene levels

Presentata da: Laura Baranello

Coordinatore Dottorato:

Vincenzo Scarlato

Relatore:

Giovanni Capranico

Index

Abstract

The DNA topology is an important modifier of DNA functions. Torsional stress is generated when right handed DNA is either over- or underwound, producing structural deformations which drive or are driven by processes such as replication, transcription, recombination and repair. DNA topoisomerases are molecular machines that regulate the topological state of the DNA in the cell. These enzymes accomplish this task by either passing one strand of the DNA through a break in the opposing strand or by passing a region of the duplex from the same or a different molecule through a double-stranded cut generated in the DNA. Because of their ability to cut one or two strands of DNA they are also target for some of the most successful anticancer drugs used in standard combination therapies of human cancers. An effective anticancer drug is Camptothecin (CPT) that specifically targets DNA topoisomerase 1 (TOP 1).

The research project of the present thesis has been focused on the role of human TOP 1 during transcription and on the transcriptional consequences associated with TOP 1 inhibition by CPT in human cell lines. Previous findings demonstrate that TOP 1 inhibition by CPT perturbs RNA polymerase (RNAP II) density at promoters and along transcribed genes suggesting an involvement of TOP 1 in RNAP II promoter proximal pausing site. Within the transcription cycle, promoter pausing is a fundamental step the importance of which has been well established as a means of coupling elongation to RNA maturation. By measuring nascent RNA transcripts bound to chromatin, we demonstrated that TOP 1 inhibition by CPT can enhance RNAP II escape from promoter proximal pausing site of the human Hypoxia Inducible Factor 1α (HIF-1α) and c-MYC genes in a dose dependent manner. This effect is dependent from Cdk7/Cdk9 activities since it can be reversed by the kinases inhibitor DRB. Since CPT affects RNAP II by promoting the hyperphosphorylation of its Rpb1 subunit the findings suggest that TOP 1inhibition by CPT may increase the activity of Cdks which in turn phosphorylate the Rpb1 subunit of RNAP II enhancing its escape from pausing.

Interestingly, the transcriptional consequences of CPT induced topological stress are wider than expected. CPT increased co-transcriptional splicing of exon1 and

2 and markedly affected alternative splicing at exon 11. Surprisingly despite its well-established transcription inhibitory activity, CPT can trigger the production of a novel long RNA (5'aHIF-1α) antisense to the human HIF-1α mRNA and a known antisense RNA at the 3' end of the gene, while decreasing mRNA levels. The effects require TOP 1 and are independent from CPT induced DNA damage. Thus, when the supercoiling imbalance promoted by CPT occurs at promoter, it may trigger deregulation of the RNAP II pausing, increased chromatin accessibility and activation/derepression of antisense transcripts in a Cdks dependent manner. A changed balance of antisense transcripts and mRNAs may regulate the activity of HIF-1α and contribute to the control of tumor progression

After focusing our TOP 1 investigations at a single gene level, we have extended the study to the whole genome by developing the "Topo-Seq" approach which generates a map of genome-wide distribution of sites of TOP 1 activity sites in human cells. The preliminary data revealed that TOP 1 preferentially localizes at intragenic regions and in particular at 5' and 3' ends of genes. Surprisingly upon TOP 1 downregulation, which impairs protein expression by 80%, TOP 1 molecules are mostly localized around 3' ends of genes, thus suggesting that its activity is essential at these regions and can be compensate at 5' ends. The developed procedure is a pioneer tool for the detection of TOP 1 cleavage sites across the genome and can open the way to further investigations of the enzyme roles in different nuclear processes.

Chapter 1

Introduction

The classical double-helix structure of DNA, as deduced by Watson and Crick, is one of the most recognizable icons of modern scientific endeavor [1, 2]. Although some details are still debated (the exact number of base pairs per turn for example), it is almost universally accepted that DNA in its lowest energy state exists in the form known as B-DNA. The biological activity of DNA is profoundly influenced by its antiparallel double-helical structure in terms of benefits for stably storing genetic information.

Owing to its structure, the double helix protects information-rich nucleotides from solvent avoiding chemical damage, while complementarity between the sister strands allows mistakes and lesions in the genome to be recognized and repaired. Because the two strands of the B-DNA are wound one around the another, processes that require strand separation – for example, replication and transcription – lead to DNA overwinding and strand entanglement [3]. These topological problems may impair some fundamental cellular processes such as gene expression, DNA duplication, and chromosome segregation. At the same time, cells depend on DNA supercoiling as a means of compacting the genome and promoting DNA unwinding at replication origins and promoter regions [4]. Thus, while on one hand the cell try to preserve the information content of the genome; on the other hand it also invests enough energy to maintain the topology of its chromosomes in an appropriate topological state [2]. Therefore, the access to the genetic information is tightly regulated and the modulation of helical winding provides mechanisms for regulating this access. In particular, these include underwinding, which facilitates strand separation, and overwinding, which inhibits strand separation.

A full understanding of the biological functions of DNA requires a detailed knowledge of its topological properties. The linking number parameter (Lk) refers to the number of times the two helical strands are interwound. The Lk for a relaxed molecule, termed Lk0, is equal to the number of base pairs divided by the period of the DNA helix. Most theoretical and computational modeling of DNA assumes that the double helix behaves as an isotropic elastic rod. Thus, overwinding or underwinding

of the helix changes the twist (Tw), a parameter describing the number of times the individual strands coil around the helical axis. If the DNA behaves as an isotropic elastic rod, then as the value of Tw increases, the associated torque should increase linearly. When Tw reaches a critical density, the molecule bends to form plectonemic structures in which the double helix coils about itself, a property known as writhe (Wr). The coiling of the double helix about itself is more commonly known as supercoiling [2]. The sum of Tw and Wr is equal to the linking number such that:

$$Lk = Tw + Wr$$

Any change in the linking number must result in a change in the twist and/or writhe such that:

$$\Delta Lk = \Delta Tw + \Delta Wr$$

As the writhe of a relaxed molecule is equal to zero, hence:

$$\Delta Wr = Wr - 0 = Wr$$

It is often useful to model DNA supercoiling by using a piece of rubber tubing. DNA is a right-handed helix, i.e. the helix spirals in a clockwise direction. Therefore, to simulate the effects of overwinding (positive supercoiling) one can introduce clockwise twist into the tubing. When sufficient twist is added, the tubing coils about itself analogously to DNA writhe. Positively supercoiled DNA coils about itself in a left-handed direction. In contrast negatively supercoiled DNA assumes a right-handed superhelical structure. Mathematically, this means that the value of Lk is lower than that of Lk0. The negative linking difference, ΔLk, is defined by:

$$\Delta Lk = Lk - Lk0$$

1.1 Biological importance of being underwound.

The DNA of most of the species, from bacteria to humans, is maintained in a homeostatically underwound, i.e. negatively supercoiled, state [2]. Negative supercoiling of the DNA facilitates localized melting of the helix, in order to facilitate the processes that require access to the DNA bases like replication and transcription. In both cases the B-DNA has to be denatured, which can only happen when it is slightly underwound [2, 5].

The regulation of genomic DNA topology is better understood in bacteria than in mammalian cells. The bacterial topoisomerase IV, is approximately 20-fold more efficient at removing positive than negative supercoils [6], and this has a biological

meaning because in bacterial cells negative supercoiling is essential and is very tightly regulated [7]. Topoisomerase IV is relatively inefficient in relaxing negative supercoils, thus, evolution appears to have provided a solution to the need of removing potentially problematic positive supercoils that would prevent polymerase progression while leaving the beneficial negative supercoils unaltered.

On the other side some examples in nature highlight the biological importance of being overwound under peculiar environmental conditions. Reverse gyrase is an enzyme that introduces positive supercoils into DNA [8]. It is found uniquely in hyperthermophilic organisms, which grow in extreme conditions where temperatures can exceed 80°C. Proteins from many hyperthermophilic Archaea wrap DNA in nucleosome-like structures that constrain DNA in a positively supercoiled conformation [9]. The existence of positively supercoiled DNA in hyperthermophilic Archaea is an elegant solution to the problem of preventing thermal denaturation in these organisms.

In eukaryotes, regulation of DNA topology overall and at specific genomic regions is much less clear. Topological problems may arise from different DNA transactions that occur in the cell because of the large size of eukaryotic DNA molecules. Thus, during DNA replication, the two strands of the DNA must become completely unlinked, and during transcription, the translocating RNA polymerase generates supercoiling tension in the DNA that must be relaxed [10, 11]. The association of DNA with histones and other proteins introduces supercoiling that requires relaxation. In all cells, completely replicated chromosomes must be untangled before partitioning and cell division can occur. These examples illustrate that not only DNA structure leads to topological problems that must be solved, but that the topological state of the DNA itself must be fine-controlled to optimize DNA function [8]. In an *in vitro* system reconstituted DNA replication machinery is only active on negatively supercoiled DNA. The DNA polymerase stalls before replication is complete, and this is attributed to the build-up of positive supercoiling ahead of the replication machinery. The addition of enzymes, as DNA topoisomerases (see below) that relax positive supercoiling, restores replication [12].

To solve unfavorable DNA supercoils, knots, and interwound duplexes one or both strands of the duplex have to be cleaved, let it rotated with respect to a second (uncleaved) segment, and then religated. To accomplish these tasks without damaging

the genome, organisms have evolved a class of proteins known as DNA topoisomerases. These enzymes are able to create transient DNA breaks through the formation of a covalent link between the enzyme and the duplex. Once the break occurs the topological stress associated with DNA is released and this could alter either the number of times the two strands of one duplex wind one around the other or the number and type of intertwinings between two DNA duplexes. Thus, the fundamental need of topoisomerases is deeply linked with the double helical structure of DNA.

1.2 DNA topoisomerases.

DNA topoisomerases are ubiquitous essential enzymes that unknot, decatenate, supercoil and relax DNA (Figure 1) [8, 13, 14].

DNA topoisomerases are molecular machines that regulate the topological state of the DNA in the cell. The immense interest in DNA topoisomerases in recent years derives not only from the recognition of their crucial role in managing DNA topology, but also from two other major reasons. First, a wide variety of topoisomerase-targeted drugs have been identified, many of which generate cytotoxic lesions by trapping the enzymes in covalent complexes on the DNA. These topoisomerase poisons include both antimicrobials and anticancer chemotherapeutics, some of which are approved by US Food and Drug Administration to treat human cancers or infectious diseases. Second, the crystal structures of numerous topoisomerase fragments provide key insights into how these machines work.

These enzymes accomplish this task by either passing one strand of the DNA through a break in the opposing strand or by passing a region of the duplex from the same or a different molecule through a double-stranded cut generated in the DNA. Because of their ability to cut one or two strands they are able to change the Lk in steps of one or two, respectively. DNA cleavage by all DNA topoisomerases is accompanied by the formation of a transient phosphodiester bond between a tyrosine residue in the active site of the enzyme and one of the two ends of the broken strand. DNA topology can be modified during the lifetime of the covalent complex, and the enzyme is released as the strand is religated.

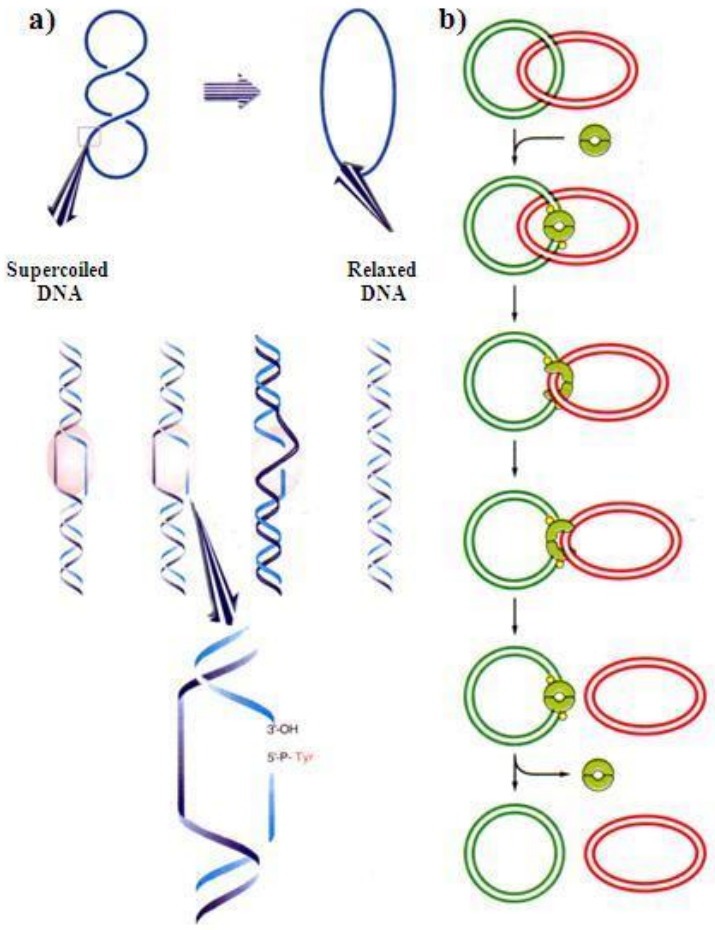

Figure 1. Mechanisms of action of eukaryotic DNA topoisomerases I and II. a) Mechanism of DNA relaxation by DNA topoisomerase I. b) Mechanism of decatenation by DNA topoisomerase II (Lewin B., 1999; and Alberts B., 1998).

Each type of topoisomerase has specific features: some are able to relax only negative supercoils, some enzymes relax supercoils of both signs; bacterial DNA gyrase and archea reverse gyrase introduce negative and positive supercoils into the DNA, respectively. Moreover the activity of topoisomerases can promote the catenation and decatenation of circular DNAs or the disentanglement of intertwined linear chromosomes [14].

Based on both amino-acid sequence and reaction mechanism, DNA topoisomerases are divided into different classes and subfamilies (Table 1). Those enzymes that cleave only one strand of the DNA are defined as type I. In particular, if the protein is attached to the 5' phosphate of the DNA, it is a type IA subfamily member. If it is attached to the 3' phosphate, the enzyme is a type IB subfamily member. Moreover, these enzymes also differ for their catalytic mechanisms. The type

IA topoisomerase family includes bacterial topoisomerase I [14] and topoisomerase III, archeal reverse gyrase and eukaryotic topoisomerase III. These enzymes are able to relax highly negatively supercoiled substrates and to efficiently unknot and decatenate DNAs containing single-stranded regions or nicks. A short stretch of double-stranded DNA is first unpaired by the binding of the enzyme, and a transient nick is introduced into the single-stranded region. The less negatively supercoiled the DNA is, the more difficult it is for the enzyme to unpair the dsDNA. This is the basis for the enzyme preference for highly negatively supercoiled DNA. Overwound is refractive to the type IA enzymes unless a pre-existing single-stranded region is present. Reverse gyrase belongs to this family; it is a unique enzyme that is capable of positively supercoiling DNA [15]. The protein shows a homology with other bacterial type IA enzymes, but the region of similarity is limited to the C-terminal half of the protein. The N-terminal half contains an ATP-binding site and motifs characteristic of helicases. The presence of separate topoisomerase and helicase domains within the same protein molecule suggests a model for the mechanism of reverse gyrase in which the helicase domain unwinds the helix as it translocates along the DNA, generating positive supercoils in front and negative supercoils behind. In turn, the topoisomerase domain, restricted to relaxing only negative supercoils, relaxes the DNA behind the translocating helicase to leave the DNA with net positive supercoiling [8].

Type IB topoisomerases include eukaryotic topoisomerase I and poxvirus topoisomerase I, as well as homologous enzymes found in certain bacteria [14]. These enzymes catalyze the relaxation of both negatively and positively supercoiled DNAs [16]. Type IB topoisomerases are believed to use a “controlled rotation” mechanism for relaxing DNA (see below) [8].

Topoisomerases that cleave both strands to generate a double-strand break are grouped together in the type II family of topoisomerases. Type II topoisomerases are subclassified as IIA or IIB on the basis of amino acid sequence and structure. All type II topoisomerases are multisubunit enzymes, and form double strand breaks with enzyme subunits covalently linked to 5’ ends at the DNA cleavage site. The enzyme catalyze topological changes by passing a second DNA duplex through the cleavage of the first DNA segment with a mechanism dependent on ATP and Mg^{2+} [8, 17]. The type IIA family includes: eukaryotic topoisomerase II (TOP 2), viral and bacteriophage topoisomerase II, bacterial DNA gyrase and topoisomerase IV.

Bacterial Topoisomerase IV can relax negative and positive supercoils, although it is particularly effective at catalyzing the latter reaction [18]. Gyrase is also capable of relaxing positive supercoils and is further distinguished by the ability to introduce negative supercoils into DNA [19]. All type IIA topoisomerases are also capable of decatenating and unknotting DNA. Type IIA topoisomerases are widespread in nature, in contrast type IIB enzymes are confined to Archaea, plants, and certain algae. Even if the distribution of topoisomerases among all species is complex, a unifying characteristic is the presence of at least one type IA and one type II enzymes. This defines a minimal set of DNA topoisomerases required for cell life.

Type IA	Type IB	Type IIA	Type II B	Organism	Gene
TOP I				*E.Coli*	*Top A*
TOP III					*Top C*
		DNA Gyrase			*GyrA/GyrB*
		TOP IV			*ParC/ParcE*
TOP III				*S. cerevisiae*	*TOP III*
	TOP I				*TOP 1*
		TOP II			*TOP 2*
Reverse Gyrase				Archaeal	
	TOP V		TOP VI		
TOP IIIα				Mammalian	*TOP 3a*
TOP IIIβ					*TOP 3b*
					TOP1 *TOP1m*
		TOP IIα			*TOP 2a*
		TOP IIβ			*TOP 2b*

Table 1. Subfamilies of DNA topoisomerases.

The DNA topoisomerases evolved to solve many topological problems of DNA. Main and general topological problems can be listed as follows:

- **DNA replication**: e**longation of replicating DNA chains.** As during replication the replication machinery advances on the DNA the unreplicated strands turns one around each other and supercoils generates. The topological consequences of this process and the roles of different DNA topoisomerases, depend on whether the replication machinery is allowed to rotate on the DNA. If the replication

machinery is immobile positive supercoils accumulate ahead of the advancing replication fork (Figure 2a). If the replication machinery is allowed to rotate around the helical axis of the unreplicated DNA, this rotation allows the redistribution of positive supercoils ahead of the advancing fork into the region behind it, and leads to the intertwining of the pair of duplicated double (Figure 2b). The positive supercoils generated can be removed by a type IB or a type II enzyme [8]. If the replication machinery could rotate, a type II enzyme might also act behind the fork to remove the intertwines between the newly synthesized DNA double helices

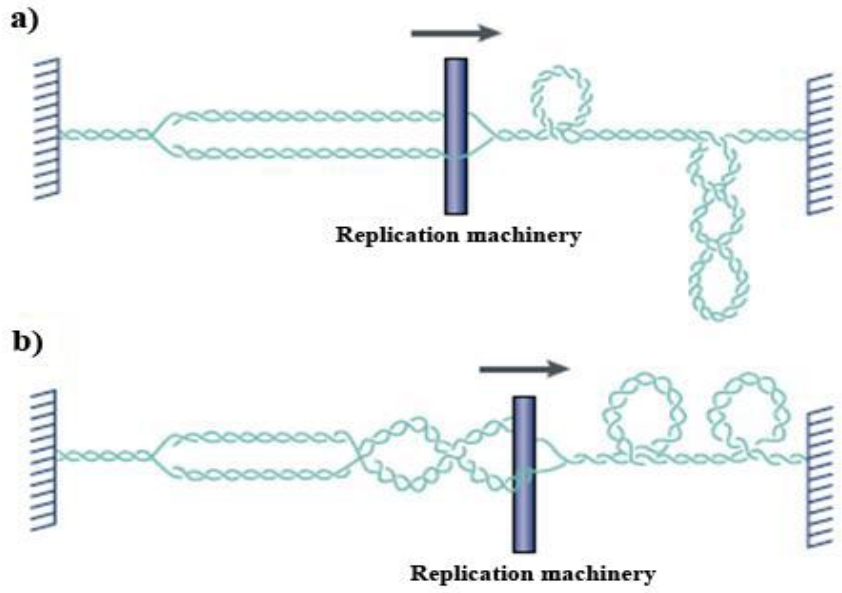

Figure 2. Topological problems associated with an elongating replication fork.
See text for details [13].

- Segregation of newly replicated chromosomes. Distinct topological problems occur when two replication forks converge. When the unreplicated DNA segment becomes very short, the type IB topoisomerase enzyme cannot remove the last few intertwines between the parental strands. These single-stranded intertwines can probably be removed by a type IA enzyme, or be converted to double-stranded intertwines for removal by a type II DNA topoisomerase. Studies of replication *in vivo* in eukaryotes are mainly based on yeast as a model system. Yeast can essentially complete replication with either topoisomerase I or topoisomerase II.

With no topoisomerase II activity *in vivo*, yeast plasmids are fully replicated, but accumulate as catenated dimers. There is good evidence that yeast chromosomal DNA is also intertwined. Holm and colleagues showed that at least some yeast chromosomes are broken at mitosis if replication takes place in the absence of topoisomerase II [20, 21].

- Transcription. During transcription the elongating RNA polymerase (RNAP) translocates along the double helix. This generates oppositely supercoiled domains ahead and behind the transcription machinery. Thus, the requirement for DNA topoisomerases is essential to allow a proper progression of transcription elongation. In yeast deletion of TOP1 has no principal effects on cell growth [13]. This is probably due to the compensation of DNA topoisomerase II that could presumably substitute for any role of DNA topoisomerase I in solving the topological problems of transcription. Indeed yeast *top1 top2* temperature-sensitive double mutants have reduced ribosomal RNA synthesis at non-permissive temperatures demonstrating the fundamental contribution of both the enzyme in the transcriptional process [13].

Moreover transcription of a chromatin template compacted with nucleosomes poses other topological problems. Passage of an advancing RNAP through a nucleosome may form a DNA loop containing both the nucleosome and the polymerase [22]. A DNA topoisomerase might be involved to allow translocation of the polymerase in such a loop.

In the recent years a great interest is focused on the role of topoisomerases in coordinating replication and transcription. By performing a genome wide analysis of TOP 2 binding site Bermejo and co-workers came to hypothesize that TOP 2 is implicated in the formation of architectural domain containing one or more transcription units in which topoisomerase I has a fundamental role in coordinating fork progression and ongoing transcription [23].

- DNA recombination. The presence of at least one type IA DNA topoisomerase in all organisms indicates that these enzymes have a key cellular function and several studies have pointed to their role in recombinational repair and in chromosome segregation. It was shown that inactivation of any one of the three yeast DNA topoisomerases increases genome instability [24]. In an effort to further define the role of *E. coli* type I topoisomerases in DNA metabolism, Zhu *et*

al. have characterized the effect of repressing both type I topoisomerase activities in *E. coli*. Cells lacking both topoisomerase activities possess an abnormal nucleoid structure. This defect can be suppressed by the deletion of the RecA gene, suggesting that these enzymes may be involved in RecA-mediated recombination and may specifically resolve recombination intermediates before partitioning [25]. In the yeasts, inactivation of the single type IA enzyme, DNA topoisomerase III, leads to slow growth and genome instability in *S. cerevisiae.* Genetic screens for suppressors of *S. cerevisiae* top III null mutants led to the identification of mutations that map to the gene SGS1. This gene encodes a helicase that has been implicated DNA recombination and repair. Significantly, the authors proved that the inactivation of SGS1 suppresses the requirement for a type IA enzyme, which suggests that a type IA enzyme might be needed to resolve a structure that is formed during recombination or repair [13, 26].

Moreover some scientific evidences relate meiotic recombination and type IIB topoisomerase. This is the case of the Spo11 in *S. cerevisiae*, a protein structurally homologous to the type IIB archaeal DNA topoisomerase VI. Because Spo11 protein links covalently to the 5′ ends of double strand DNA breaks at the hot-spots of meiotic recombination, this raises the possibility that also type IIB topoisomerase may be a key player in the same process in archetypal organism [13].

- Chromosome condensation. Chromatin compaction, chromosome segregation and DNA topology are interrelated. When DNA undergoes to protein compaction changes in twist and writhe are consistent with a possible requirement of DNA topoisomerase [20]. In bacteria, it seems that DNA supercoiling can also affect chromosome compaction and segregation. The efficiency of *E. Coli* plasmid partition seems to be increased by the presence of a strong gyrase-binding site on a plasmid [13], or by a reduction of the intracellular activity of DNA topoisomerase I. It appears that a higher degree of negative supercoiling makes the plasmid molecules more compact, and hence more easily segregated [13, 27].

- Chromosome structure. Early studies suggested that topoisomerase II plays an important role in chromosome structure. A structure termed ‘chromosome scaffold’ was defined as a set of proteins that remained following removal of DNA and histones from mitotic chromosomes. An examination of the proteins

constituting this fraction identified topoisomerase II as one of the major components. In metaphase chromosomes, topoisomerase II was found to be associated with the axial elements, suggesting that the protein resided at the base of chromosome loops [20, 28]

1.3 Human DNA topoisomerase IB.

The best known type IB enzyme is the human DNA topoisomerase I (TOP 1). Based on sequence comparisons, it is likely that most features described here for the human enzyme apply to the cellular topoisomerases I from other eukaryotic species.

1.3.1 Structure.

Based on conservation of sequence, sensitivity to limited proteolysis, hydrodynamic properties, and fragment reconstitution experiments, the 91-kDa human topoisomerase I protein has been subdivided into four distinct domains [29] (Figure 3). The N-terminal 214 amino acids of the human enzyme are dispensable for relaxation activity in vitro and constitute a hydrophilic, non-conserved and unstructured domain of the protein [29] (Figure 3). The N-terminal domain contains four nuclear localization signals and sites for interaction with other cellular proteins, including such proteins as nucleolin, SV40 T-antigen, certain transcription factors, p53, and the WRN protein [8] The N-terminal domain is followed by a highly conserved, core domain that contains all of the catalytic residues except the active site tyrosine [8]. The linker domain comprising 77 amino acids connects the core domain to the C-terminal domain. The active site Tyr723 is found within the C-terminal domain. The crystal structures of several forms of the human enzyme with both noncovalently and covalently bound DNA have been determined [30-32].

These co-crystal structures represent the only examples to date of a topoisomerase bound to DNA. Two views of the structure of human TOP 1 noncovalently complexed with a 22 base pair DNA are shown in Figure 4 with the entire N-terminal domain missing from the structure (Figure 3). The bi-lobed protein clamps tightly around the DNA with contacts between the protein and the DNA phosphate backbone. Most of the contacts are grouped around the five base pairs upstream (-5 to -1) of the cleavage site, which occurs between base residues -1 and +1. The core domain is subdivided into core subdomains I, II, and III, as shown by the

color scheme in Figure 4. One of the lobes of the protein comprises core subdomains I and II (yellow and blue) and forms what has been referred to as the "cap" of the protein [32]. The other lobe forms a base around the DNA and consists of core subdomain III (red) and the C-terminal domain (green). This second lobe is connected to the cap by a long α-helix labeled the "connector" in Figure 4a. On the side opposite to the connector are pairs of opposing loops called the "lips" that allow non-covalent interactions of the cap with the base of the protein. The break of this interaction and lifting of the cap results in the opening and closing of the protein clamp during the catalytic cycle. The dispensable 77 amino acid linker domain (orange) is a coiled-coil structure that protrudes from the base of the protein [8].

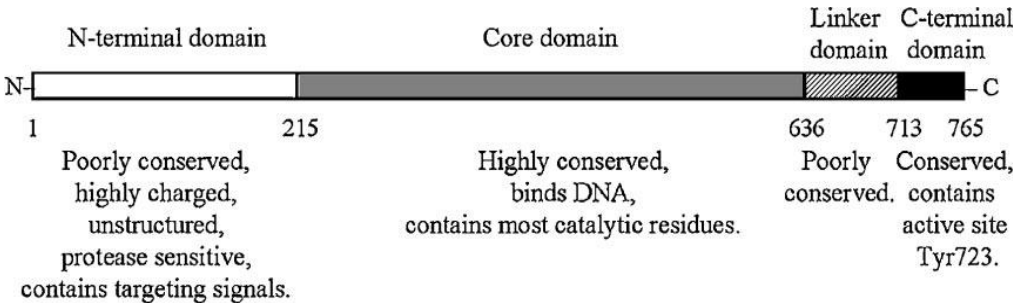

Figure 3. Domain structure of the human DNA topoisomerase I. Human TOP 1 comprises an N-terminal domain (open box), a core domain (gray box), a linker domain (diagonally striped box), and a C-terminal domain (black box). Each domain has specific properties indicated in the lower portion of the figure [8].

Despite very little sequence homology, most of core subdomain III of human TOP 1 share some structural features with the core region of a family of tyrosine recombinases that includes bacteriophages HP1 and¸ integrases, and *E. coli* XerD and bacteriophage P1 Cre recombinases. The structural similarity reflects many biochemical properties shared by the two classes of enzymes. The tyrosine recombinases and type IB topoisomerases use the same chemistry to carry out cleavage and religation. The two classes of enzymes differ by what occurs during the cleaved intermediate. The topoisomerases allow topological changes in the DNA and restore the original DNA linkage, whereas the recombinases promote strand exchange and join the DNA ends to new partners [30].

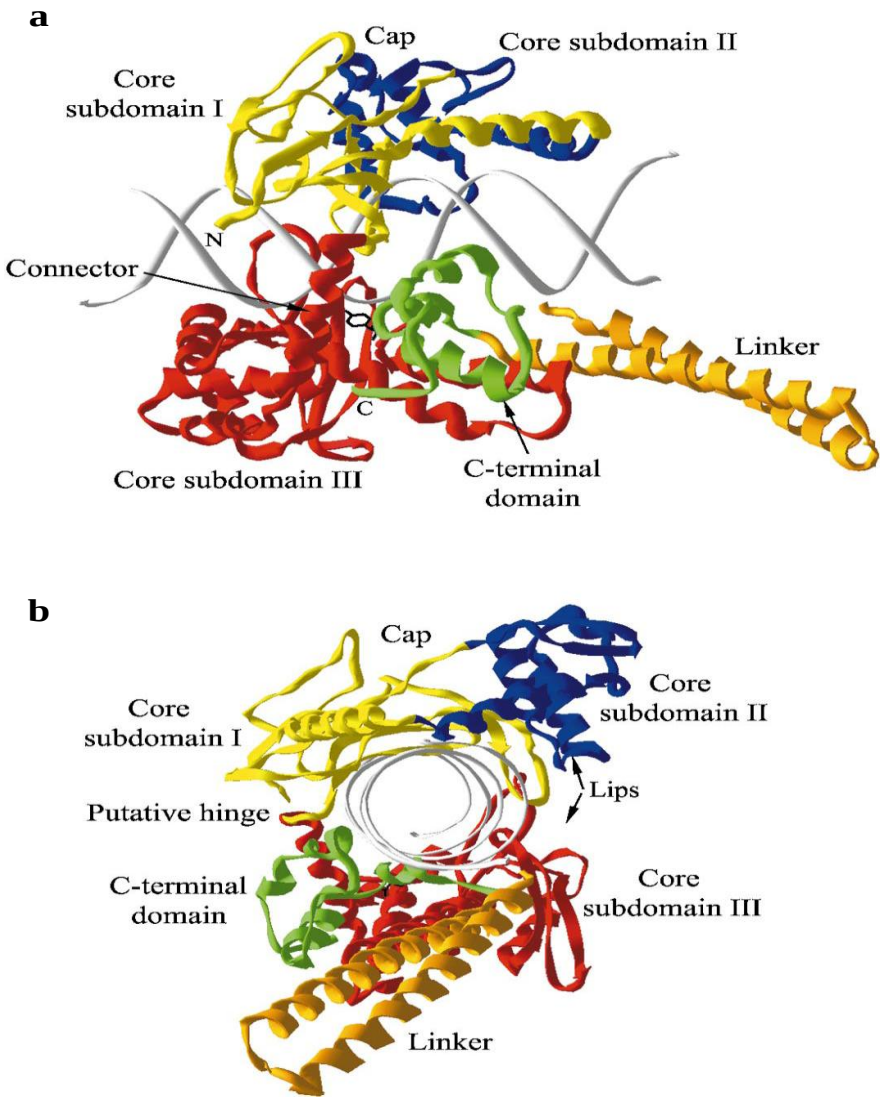

Figure 4. Two views of the structure of human DNA topoisomerase I bound to DNA. a) Side view of TOP 1 with the DNA axis horizontally oriented. b) View of the enzyme looking down the axis of the DNA. Core subdomains I (residues 215–232 and 320–433), II (residues 233–319) and III (residues 434–633) are colored yellow, blue, and red, respectively. The linker (residues 641–712) and C-terminal domain (residues 713– 765) are shown in orange and green, respectively. In panel a) the active site tyrosine is shown in black ball and stick and the amino terminus (N) and carboxyl terminus (C) of the protein are indicated. The cap and the base are connected by an α-helix of the core labeled "Connector" In panel b) the arrows indicate the "Lips" that is a region where the protein opens during DNA binding and unbinding. The hinge region at the top of the connector helix is labeled "Putative hinge" [8].

The substrate specificity of eukaryotic TOP 1 has been characterized at both the nucleotide sequence level and at the level of DNA tertiary structure. The preferred nucleotides in the scissile strand are 5'-(A/T)(G/C)(A/T)T-3' with the enzyme covalently attached to the -1 T residue. Occasionally a C residue is found at the -1 position. In the crystal structure, it is possible to recognize only one specific base

contact between the protein and the DNA, which involves a hydrogen-bond between Lys532 and the O^{-2} atom of the thymine base located at the -1 position. It appears that further protein-DNA interactions play an important role in cleavage site selection.

A number of studies have indicated that eukaryotic TOP 1 binds preferentially to supercoiled rather than relaxed plasmid DNA. The use of a mutant form of the protein with phenylalanine in place of tyrosine at the active site (Y723F) showed that the enzyme has a strong preference for binding to supercoiled over relaxed DNA [33]. Since the enzyme preferentially binds to supercoils of either sign, it seems likely that the structural feature recognized in the DNA is a bent DNA segment. Alternatively, the enzymes may recognize the node where two duplexes cross, this is the case of replication.

1.3.2 Mechanisms of Catalysis and Relaxation.

The mechanism of catalysis (Figure 5) involves a nucleophilic attack by the O^{-4} oxygen of Tyr723 on the scissile phosphate of the DNA strand. This generates a phosphodiester link between the tyrosine and the 3' phosphate, releasing a 5' hydroxyl [30]. In the crystal structure, Tyr723 is perfectly aligned for a nucleophilic attack and other basic amino acids are positioned to stabilize the complex through interactions with the non-bridging oxygens of the scissile phosphate. When the reaction starts the enzyme is in an open conformation, with the cap and catalytic base spread apart. Upon binding to DNA, these domains close to form a clamp around the duplex. Then the nucleophilic attack on the scissile phosphate of the DNA links covalently the enzyme to a strand end and permits the tethered duplex segments to rotate with respect to one another in order to relax positive or negative supercoils.

Rotation is slowed by contacts with the inner cavity of the enzyme. The linker found in eukaryotic TOP 1 provides additional interactions between the protein and DNA, further regulating the cleavage/religation equilibrium and adding an additional layer of control over the rate of rotation. Thus, the term "controlled rotation" was coined to indicate that these structural domains of the protein likely hinder or impede the rotation reaction [14, 31]. Religation requires the 5′-hydroxyl-group at the DNA end to be aligned with the tyrosine-DNA phosphodiester bond. Under normal conditions, the cleavage intermediates are transient and religation is favored.

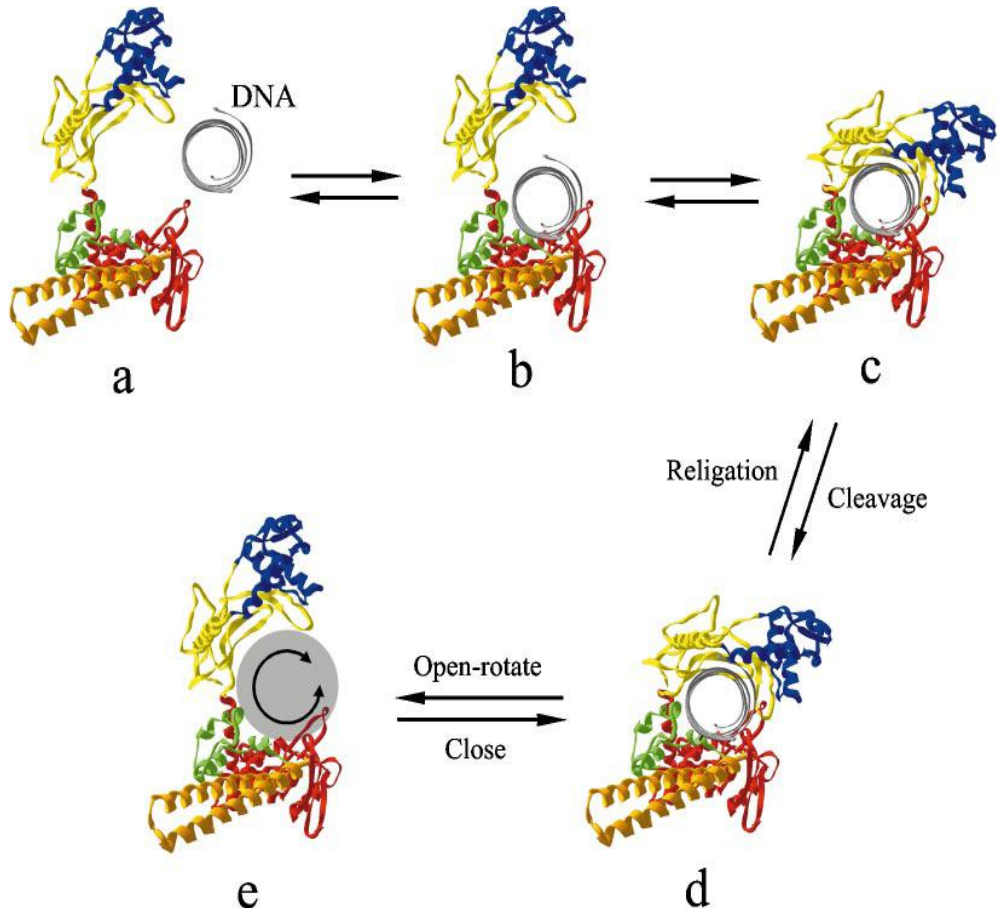

Figure 5. Proposed mechanism of DNA relaxation by human DNA Topoisomerase I. a) To bind DNA the enzyme assumes an open conformation with the "cap" and the catalytic base spread apart. b) Upon binding to DNA, the enzyme catalyzes the cleavage of the scissile strand (c and d). The cleaved DNA rotates around the intact strand (e). Multiple rotations are possible before the re-ligation step. Then, DNA is either cleaved again or released [8].

1.3.3 TOP 1 and gene expression.

Mammalian TOP 1 is enriched in transcribed genomic regions as established by TOP 1 DNA cleavage sites [8, 13, 34] and chromatin immunoprecipitation (ChIP) [35]. The localization of TOP 1 in nuclear chromatin is likely guided by its N-terminal domain [36]. Early studies in the yeast *S. cerevisiae* indicate that neither TOP 1 nor TOP 2 are essential for transcription by RNAP II [8, 13]. However, plasmids carrying transcriptionally active genes are found to be extremely negatively supercoiled when isolated from mutants lacking both TOP 1 and TOP 2, and slightly negatively supercoiled in mutants lacking TOP 1 only [37]. Thus, a main molecular function of TOP 1 is generally considered the relaxation of transcription-dependent DNA supercoils.

It is well established that waves of positive and negative supercoils are generated ahead and behind the elongating RNA polymerase if the translocating transcriptional apparatus cannot turn around the DNA template (the "twin supercoiled-domain model") (Fig 6) [8, 13]. The twin supercoiling model has been supported by several findings [11]. Within a topological domain, that may have diverse consequences. First, a high degree of template supercoiling can inhibit further transcription, particularly when the transcription rate is high [38]. As a consequence, the TOP 1 function may be the relaxation of torsional stress to allow a normal rate of RNA synthesis. Secondly, it is interesting to point out that localized changes in supercoiling of DNA can be exploited to regulate other nuclear processes [39].

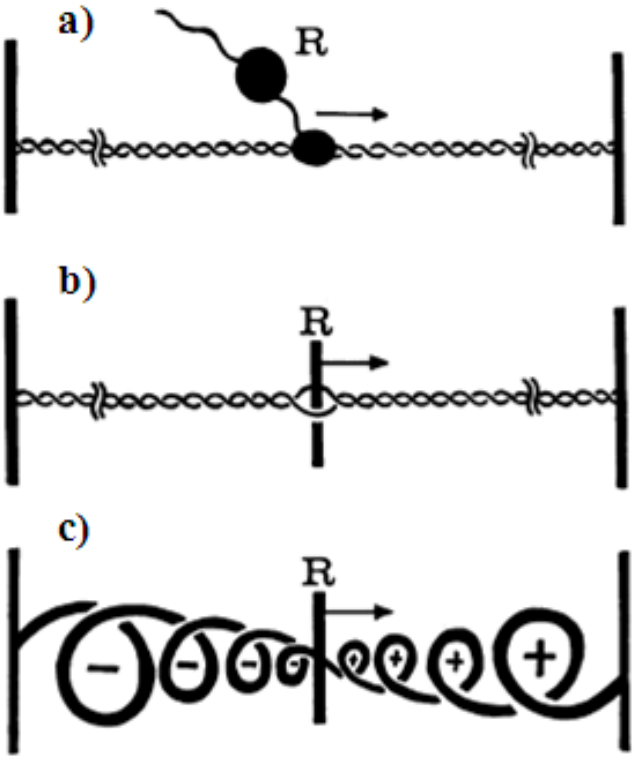

Figure 6. Topological problems associated with transcription. a) The transcription machinery indicated as an R is moving in the direction of the arrow along a DNA segment. The ends of the DNA are attached to a hypothetical immobile structure, which could be the nuclear or cell membrane. b) and c) When R cannot rotate around the helical axis of the DNA template, overwinding or positive supercoiling of the DNA template ahead of R is accompanied by underwinding or negative supercoiling of the DNA template behind R [11].

In an attempt to understand how TOP 2 might solve topological constraints arising when replication forks encounter transcription Bermejo and co-workers

performed a genome wide analysis of TOP 2 binding site in S phase *S. cervisiae*. The authors came to demonstrate that TOP 2 preferentially binds to promoter and transcription termination regions, but the enzyme is not needed to licence the transcriptional program as shown by TOP 2 mutants [23]. This is probably due to the contribution of TOP 1 in transcription. TOP 1 might coordinate fork progression and ongoing transcription while TOP 2 in might be involved in formation of architectural domain containing one or more transcription units [23].

Interestingly, TOP 1 has also been shown to activate gene transcription, and to bind to general transcription factors at promoters [40, 41]. TOP 1 roles in transcription initiation were examined in a highly defined *in vitro* transcription system containing RNAP II and purified factors [42]. TOP 1 was able to either repress basal transcription as well as to co-activate transcription. Activation was first observed to be dependent on TFIID-TFIIA, and maximal activation by TOP 1 required the concomitant presence of an inducible activator. In this contest a possible target of TOP 1 could be RNAP II. Indeed it was shown that TOP 1 could bind the C-terminal domain of Rpb1, the largest subunit of RNAP II [43].

Interestingly, a catalytically inactive TOP 1 mutant and the WT enzyme were similarly able to stimulate transcription and formation of TFIID–TFIIA-DNA complexes. Taken together, the results showed that TOP 1 can likely function by enhancing the formation of an active TFIID-TFIIA complex at the studied promoter [42]. To explain the activation by the inactive mutant enzyme, the authors proposed that the loading of TOP 1 onto the polymerase at initiation can be necessary for the subsequent elongation phase to relieve transcription-generated supercoils. If the enzyme is not loaded, then transcription is not activated since it would be less efficient during elongation [41].

However, it must be pointed out that negative supercoils may enhance transcription initiation at eukaryotic promoters, and TOP 1 could thus modulate such a process at least in some cases [44]. At the c-MYC promoter, a DNA loop has been proposed to form by the interactions of the FBP-FIR protein system with both the upstream FUSE element and TFIIH at the P2 site where an RNAPII is paused [45, 46]. Therefore, TOP 1 might modulate the negative superhelicity of a DNA loop at transcribed gene promoters, consequently affecting transcription initiation in human cells [47].

The biological functions of human TOP 1 have been difficult to study because knocking out TOP 1 gene is not viable for the cell [48]. Through the generation of stable TOP 1 small interfering RNA (siRNA) cell lines Miao and co-workers [49] recently showed the pleiotropic nature of human TOP 1 activities. Indeed, TOP 1 plays critical non-classic roles in genomic stability, gene-specific transcription, and response to various anticancer agents. Interestingly in these cells topoisomerase 2α compensates for TOP 1 deficiency. This was shown also in mutant yeast without TOP 1 and reflects the redundancy of their role in the nuclear context [13, 49]. From a broader point of view, Miao findings are significant because they indicate that TOP 1 may have novel unexpected functions that might be considered as independent of its classic DNA nicking-closing functions during transcription.

1.3.4 CPT: a specific Topoisomerase I B inhibitor.

In order to carry out their critical physiological functions, topoisomerases generates transient breaks in DNA. Consequently, while essential for many vital functions of DNA during normal cell growth, these enzymes are also highly vulnerable under various physiological and nonphysiological stresses because of their capability to cut the DNA. These stresses (e.g. exposure to topoisomerase poisons, acidic pH, and oxidative stresses) can convert DNA topoisomerases into DNA-breaking nucleases, resulting in cell death and/or genomic instability.

The DNA cleavage/ligation reaction of topoisomerases is the target for some of the most successful anticancer drugs currently in clinical use [34, 50]. DNA topoisomerases are particularly vulnerable to inhibitors during their cleavage intermediate step (referred to as the cleavage complex- TOP 1cc). Despite their high frequency throughout the genome, TOP 1 cleavage complexes are normally so transient that they are not detectable. Campthotecin (CPT) and close derivatives (Fig 7) are able to specifically and reversibly trap TOP 1ccs. Because the religation of TOP 1ccs requires nucleophilic attack of the tyrosyl-DNA-phosphodiester bond by the free DNA end (the 5'-hydroxyl end), it is crucial for the 5'-hydroxyl-DNA end to be perfectly aligned with the tyrosyl-phosphodiester bond. It is thought that CPT intercalating into DNA base pairs at the break site, where the enzyme is linked to the double helix, promotes a misalignment of the 5'-hydroxyl end of the DNA increasing the half life of the complex [30, 34]. The rapid resealing of TOP 1cc is also inhibited

by common DNA base alterations such as base mismatches, base oxidation, abasic sites, carcinogenic adducts and pre-existing DNA breaks [51, 52] as these lesions may likely cause a misalignment.

Camptothecin was originally identified as the antitumor component in the extract of the plant *Camptotheca acuminate*. Then, it was tested clinically and showed effective anticancer activity in human patients [34]. In 1985 human TOP 1 was identified as a molecular target for CPT. Subsequent studies in yeast using a *ΔTOP 1* mutant strain demonstrated that the yeast cells become totally resistant to CPT when TOP 1 is removed. This findings firmly established that human TOP 1 is the sole antitumor target of camptothecin [53, 54].

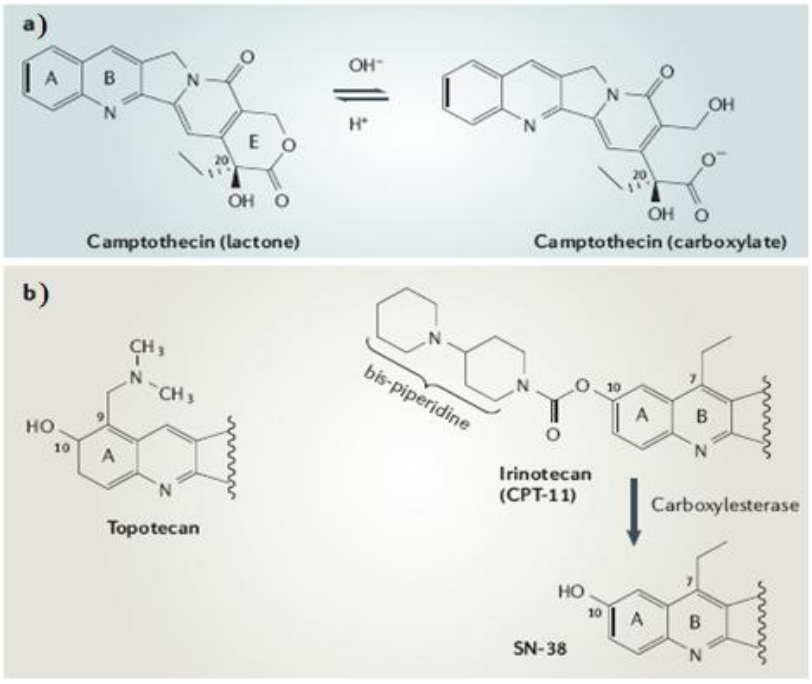

Figure 7. The chemical structure of camptothecin and its derivatives. a) Camptothecin is a 5-ring heterocyclic alkaloid that contains an hydroxylactone within its E-ring that is unstable at physiological pH. The picture shows the equilibrium between the active lactone form and its inactive carboxylate derivative. For all camptothecin derivatives, the carboxylate form is inactive as a TOP 1 inhibitor. b) Topotecan (Hycamtin®) and irinotecan are routinely used for IV infusion in cancer treatment [34].

CPT has a potent antitumor activity against a broad spectrum of solid tumors and has been approved by FDA for the treatments of human cancers, as ovary, colon and lung tumors. Several CPT derivatives or other TOP 1 poisons are currently in clinical trials for solid tumors and leaukemia [34, 55-57].

Camptothecin is a non-competitive inhibitor of TOP 1. The drug reversibly binds to the TOP 1-DNA complex, however it does not bind to each component alone [58, 59]. The reversible TOP 1-camptothecin-DNA ternary complex is apparently non productive in catalysis [58]. Camptothecins penetrate mammalian cells readily and target TOP 1 within minutes of exposure. Because the cleavage complexes reverse within minutes after camptothecin removal from the medium, camptothecin is an incisive molecular tool, as the drug exposure and trapping of TOP 1ccs can be precisely controlled.

Because TOP 1 inhibitors bind reversibly to TOP 1ccs, and TOP 1 readily religates the cleaved DNA after drug removal, TOP 1 inhibitors do not directly damage DNA. It is TOP 1 itself that damages DNA in connection with some DNA processes, primarily replication and transcription. Camptothecin is known to kill S-phase cells selectively [60]. It is now well established that camptothecin kills S-phase cells by a mechanism involving collision between advancing replication fork and TOP 1cc (Figure 8). This collision triggers both S-phase-specific cell death and cell cycle arrest at the G2 phase of the cell cycle [60]. Moreover it leads to three main events: formation of a double-strand break at the fork, irreversible arrest of fork movement, and formation of an irreversible TOP 1-DNA covalent adduct [34].

Replication double-strand breaks produce a pleiotropic response that starts with activation of phosphorylation cascades (Figure 9). These cascades primarily involve checkpoint kinases ATM (ataxia telangiectasia mutated) ATR (ataxia telangiectasia and RAD3 related) and DNAPK (DNA-dependent protein kinase), and their target CHK2 and CHK1 [34]. One molecular target of the kinase activity is the histone variant H2AX [61]. The phosphorylation of H2AX (γ-H2AX) is a ubiquitous response to DNA double-strand breaks that occur in non-replicating or replicating cells. As it happens within minutes after the formation of CPT induced double-strand breaks it is a useful biomarker to monitor DNA breaks and tumour response to TOP 1 inhibitors.

The phosphorylation cascades lead to several downstream events such as CDC25 phosphatase inhibition, MDM2 inactivation and the activation of p53, which in turn induce a cell cycle arrest and eventually repair of the damage [34].

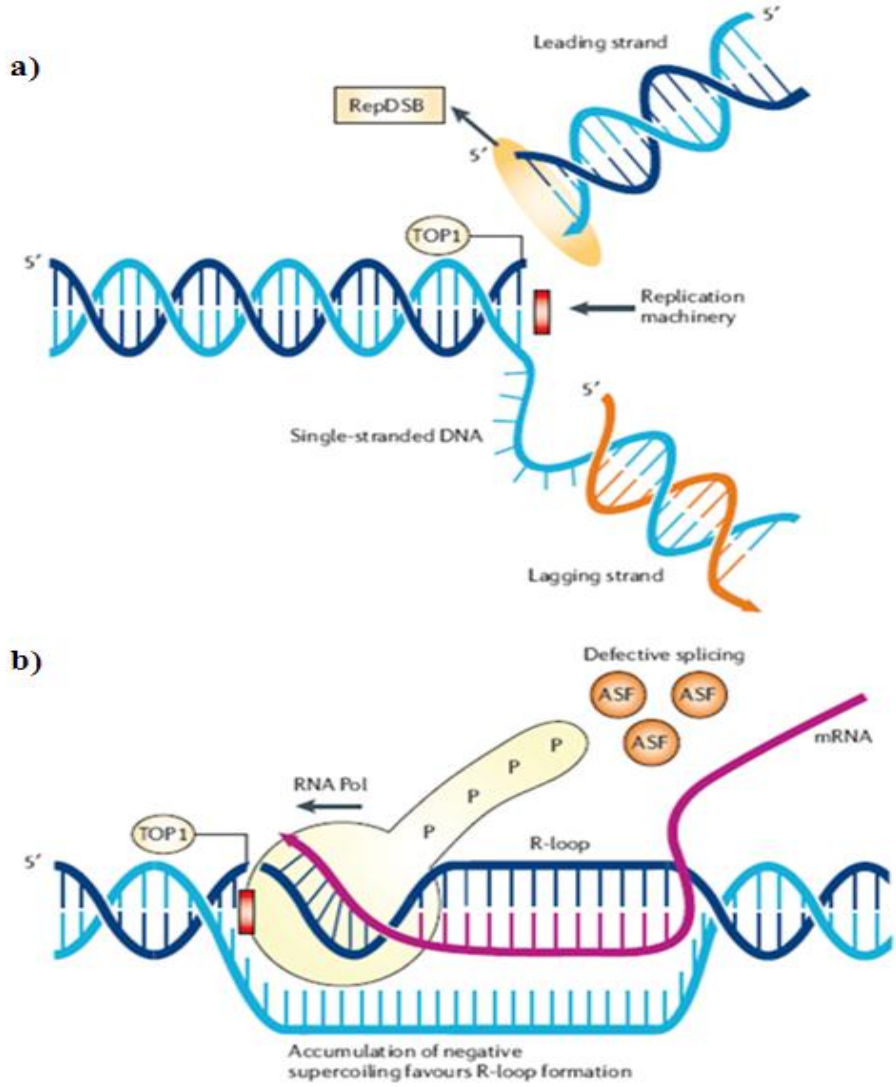

Figure 8. Conversion of TOP 1 cleavage complexes into DNA damage by replication-fork collision and transcription. a) When the elongating replication forks collides with a CPT trapped TOP 1cc a replication double-strand break generates. Moreover the collision induces arrest of the fork, generation of an irreversible TOP 1 covalent complex and cell death. b) When a CPT trapped TOP 1cc is on the transcribed strand (dark blue), the RNA polymerase can crash the cleavage complex and arrest RNA elongation. The RNA–DNA duplex prevents the religation of the TOP 1cc generating an irreversible single strand break. TOP 1 inhibition eventually leads to an accumulation of negative supercoiling that could promote the formation of an R-loop [34].

Early studies demonstrated that camptothecin inhibits both DNA and RNA synthesis [62]. A broad and general inhibition of transcription elongation is an immediate effect of CPT in cultured cell, likely due to the stalling of elongating RNAP II by TOP 1ccs and/or by persistent transcription generated DNA supercoilos [34]. Studies in human HeLa cells showed that camptothecin treatments result in rapid transcription arrest and redistribution of the RNA polymerase elongation complexes

[62, 63]. The RNA polymerase elongation complexes accumulate to a higher level at the 5' end of the gene and gradually decrease in density toward the 3' end. This result suggests that camptothecin specifically inhibits transcription elongation. The molecular mechanism of transcription-elongation inhibition by camptothecin has been suggested by *in vitro* studies. Moreover some *in vivo cell* studies indicate that CPT treatment promotes hyperphosphorylation of the Rpb1 large subunit of RNA polymerase II likely through Cdk7 and/or Cdk9 kinases activity [35, 64]. Since recent studies showed that UV-induced hyperphosphorylation of RNAP II may cause lower rate in transcription elongation [65], CPT could also impair RNAP II elongation through a similar mechanism. However, the mechanism has not been demonstrated *in vivo*.

Transcription complexes, similar to replication complexes, have been proposed to convert reversible TOP 1ccs to irreversible TOP 1 covalent complexes [66, 67] (Figure 8). However, in highly proliferative cancer cells, transcription seems to contribute much less than replication to the anticancer activity of camptothecins [68]. The nature of transcription-mediated DNA damage and its cellular responses is less characterized than that of replication-induced damage. Interestingly, Cockayne syndrome cells which are deficient in transcription-coupled repair, show strong hypersensitivity when exposed to CPT [69]. This indicates a possible link with transcription-associated DNA repair.

Besides forming irreversible TOP 1 covalent complexes, transcription collisions can induce the proteasomal degradation of both TOP 1 and RNAP II [70]. In a paper published by Liu and colleagues in 2001, it was shown that non-transformed cells are much more proficient in CPT-induced TOP1 down-regulation than their transformed counterparts. TOP 1 degradation in normal tissue might function as a protective mechanism by which normal cells adapt to camptothecin. In cancer cells this mechanism is impaired. Based on these findings it appears that CPT sensitivity is related to the ability of the cell to degrade TOP 1 [71]. Another consequence of trapped TOP 1ccs during transcription might be the accumulation of negative supercoiling upstream of the transcription block, which could favour the formation of R-loop structures (DNA:RNA hybrids; Figure 8). R-loops result from the extended pairing of nascent mRNA with the corresponding unwound DNA template behind the elongation machinery [72]. These structures are known to induce DSBs and genome

instability [73]. TOP 1 inhibition by camptothecin has also been reported to block the SR-kinase activity of TOP 1. In which case, splicing might be inhibited as pre-mRNA processing and splicing becomes functionally defective (Figure 8) [74].

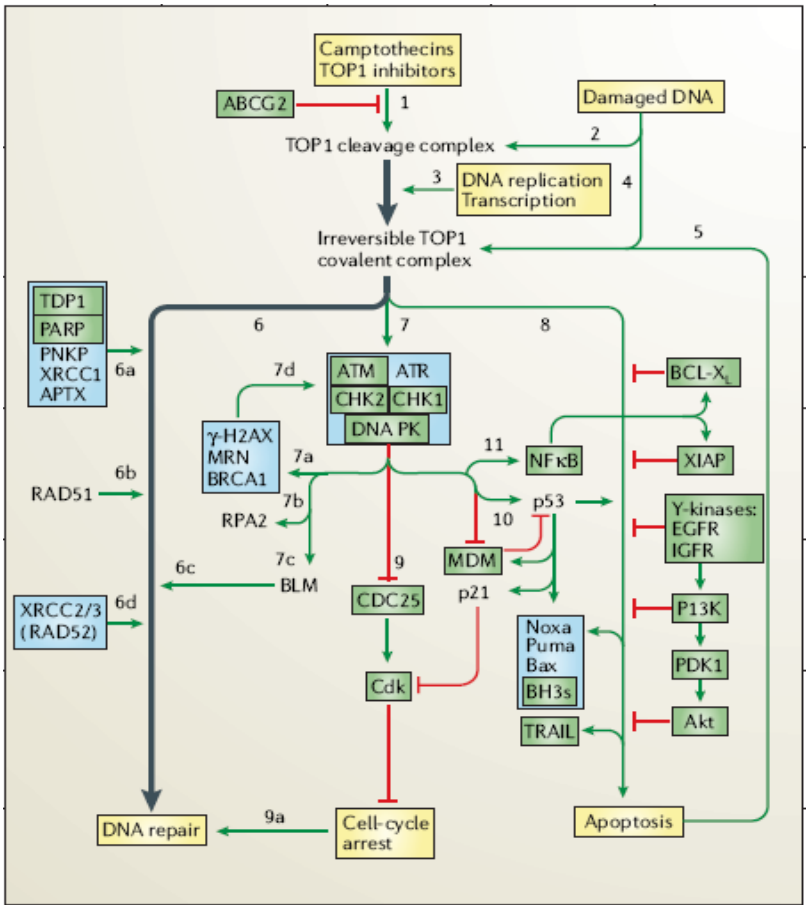

Fig 9 Molecular pathways involved in cellular responses to TOP 1 cleavage complexes [34].

Although the established cellular effects of camptothecin are specific of DNA damage responses, TOP 1cc occurs primarily in actively transcribed regions. Nevertheless what exactly happens when a TOP 1 trapped CPT encounters a transcription apparatus *in vivo* is still unclear and subject of investigation.

Our group is particularly interested in this topic since it could shed light on novel functions of TOP 1 during transcription. To this purpose, we use TOP 1 poisoning by CPT to study molecular dynamics of RNAP II and other components of the transcription machinery at actively transcribed loci.

Interestingly, several reports published by us and other groups have revealed different roles for TOP 1 activity at actively transcribed regions in living cells. First, CPT-induced TOP 1ccs have immediate and specific effects on RNAP II. The drug triggers the phosphorylation of the largest subunit (Rpb1) of RNAP II [35, 70], showing an effect on a critical step of transcription regulation (see below). Hyperphosphorylation occurs selectively on Ser-5 residues of the conserved heptapeptide repeats of the carboxy-terminal domain (CTD) possibly mediated by Cdk7, component of TFIIH [75]. Interestingly, a recent report showed that camptothecin can disrupt the large inactive P-TEFb complex (see below), thus releasing a free active P-TEFb complex (containing the Cdk9 kinase activity), which may then contribute to camptothecin-increased phosphorylation of RNAP II [76]. A second immediate effect of camptothecin on RNAP II reported by us [35] can be correlated to the hyperphosphorylation of Rpb1. Short cell treatments with CPT induced a redistribution of chromatin-bound RNAP II along transcribed genes in human cancer cells (Figure 10a). Interestingly, RNAP II reduction at promoter pause sites occurred within 5-10 min of camptothecin treatment, and was not a response to replication-dependent DNA breaks. ChIP analyses of RNAP II along transcribed genes indicated that RNAP II levels were transiently increased at internal exons, and that camptothecin effects could be fully reversed by DRB, a cdk inhibitor (Figure 10b). The data let us to hypothesize that camptothecin-induced alterations of RNAP II distribution along transcribed genes were due to an enhanced polymerase escape form promoter-proximal pause sites. In an effort to interpret all these different findings we elaborated a model suggesting a novel TOP 1 role at the level of promoter pausing reassumed in Figure 10c [35].

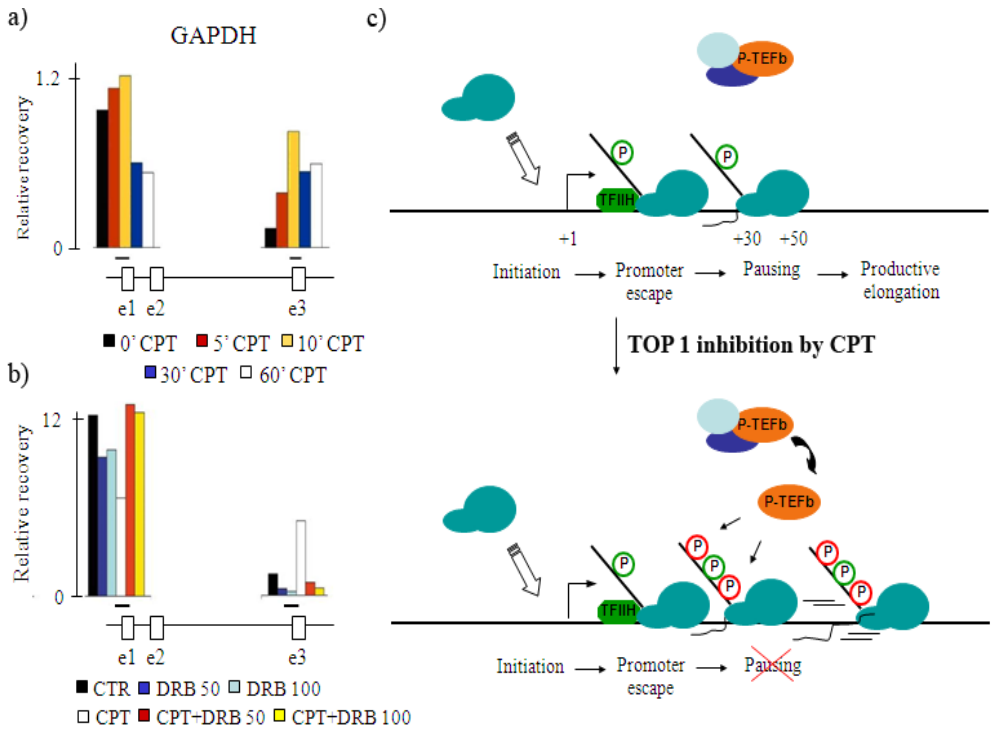

Figure 10. RNAP II dynamics after TOP 1 inhibition by CPT. a) Cells were treated with 10 µM CPT for 5, 10, 30 and 60 minutes. RNAP II binding to GAPD gene fragments was assessed through Chromatin immunoprecipitation (ChIP) and quantitative real time quantitative PCR (qPCR). DNA recovery is normalized relative to maximal recovery of untreated cells. b) Cells were treated with CPT 10 µM or co-treated with CPT and DRB for 1 hour. RNAP II levels on GAPD gene fragments were analyzed by ChIP. c) Model of the regulation of RNAP II pausing by TOP 1. At the promoter level the first step of transcription is the pre-initiation complex (PIC) assembly. After the DNA is unwound and an open complex is formed, RNAP II synthesizes a short RNA and paused partially at 20-50 bases from transcription start site. The polymerase pausing is facilitated by TOP 1 activity. The inhibition of TOP 1 by CPT activates PTEF-b which in turns phosphorylates the carboxy-terminal-domain of RNAP II favoring pausing escape. By reversibly interferes with the TOP 1 activity CPT lead to an increased escape of polymerases from the pausing site.

1.4 Chromatin structure and transcriptional regulation.

Precise temporal and spatial regulation of the transcription of protein-encoding genes by RNAP II is essential to the execution of complex gene expression programs in mammalian cells in response to growth, cell cycle progression, developmental and homeostatic signals. Eukaryotic transcription is a highly regulated process, and chromatin structure, the way in which DNA is packed in the eukaryotic cell together with nucleoprotein complexes, is known to have a major impact on transcription. At the first level of organization, eukaryotic DNA typically exists *in vivo* as a repeating array of nucleosome core particle, in which nearly two tight super helical turns of

DNA (146 bp) are wound around a histone octamer, consisting of two of each histone proteins H2A, H2B, H3, and H4 [77, 78].

Long arrays of nucleosomes, connected by linker DNA of variable length, are further compacted in multiple levels of complex architecture. The linker histone H1 and different non-histone proteins are critically involved in this process (Figure 11). Since accessibility of the DNA that is "sequestered" by nucleosomes differs dramatically from that of linear protein-free DNA, all these aspects of chromatin organization and compaction have fundamental implications not only for transcription but for all the biological processes that use DNA as a substrate, as well as replication, DNA repair, and recombination.

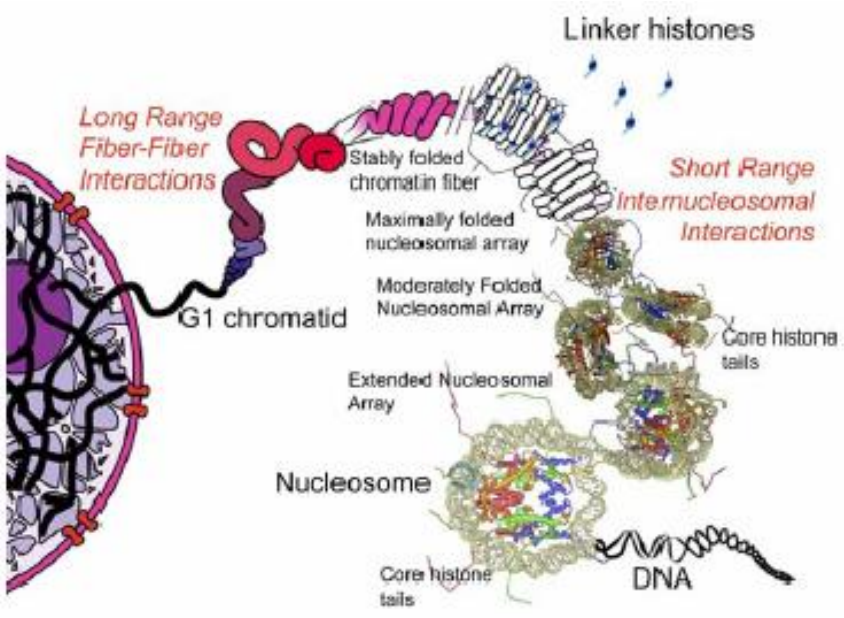

Figure 11. Modular organization of eukaryotic DNA. Multiple levels of DNA compaction are depicted in a schematic manner. Long arrays of nucleosomes are compacted via short-range and long-range interactions into fibers of unknown architecture.

It is well known that eukaryotic chromosomes are divided into domains with distinct structural features. Heterochromatic domains are required for chromosome segregation and telomere maintenance as well as for suppressing recombination between repetitive elements. These domains encompass chromosomal regions that have few genes and are assembled into a highly compacted and condensed chromatin structure that limits the access of transcription factors.

Interspersed within these transcriptionally repressive domains are euchromatic regions that are gene rich and organized in a chromatin packaging structure more accessible to transcription factors binding.

Nucleosomes compact the genome but also restrict the access of DNA-binding transcription factors, so there is a balance between effective genome packaging and accessibility. This regulation involves dynamic competition between nucleosomes and transcription factors for important key *cis*-regulatory sequences in gene promoters.

Several enzymes and protein complexes are known to change the state of chromatin by disrupting nucleosome structure and altering the DNA packaging (chromatin remodelling) in an ATP-dependent manner. As a result, DNA becomes more accessible and both specific and general transcription factors (GTFs) can bind to their sites. Human ATP-dependent chromatin remodelling complexes include the SWI2/SNF2 and ISWI-based ATP-dependent chromatin remodelling complexes that can alter nucleosome structure either individually or in combination with different co-activators [79, 80].

Moreover, it has to be considered that another striking feature of core histones is that they are subjected to a large number of covalent modifications and a direct connection between histones modifications and transcriptional activation/repression has been increasingly demonstrated in recent years. The core histones do not share any primary sequence similarity but they have a common domain structure. The central portion of each histone consists of a globular core domain that folds into the characteristic histone fold. These histone fold-containing regions associate with each other to constitute the bulk of the histone octamer around which DNA wraps. The N-terminal tails of the core histones contain a number of sites that can be subjected to post-translational modifications. These modifications can include phosphorylation, ubiquitination, ADP-ribosylation, but the best-characterized are acetylation, catalyzed by histone acetyltransferase (HATs), de-acetylation catalysed by histone deacetylase (HDACs) and methylation, catalysed by histone methyltransferase (HMTs) [81]. Some modifications, such as acetylation and phosphorylation, can alter the net charge of the tails and, therefore, have the potential to influence chromatin structure through electrostatic mechanisms, for instance acetylation is associated to a more accessible chromatin structure.

Moreover, giving the remarkable diversity and biological specificity associated with distinct patterns of covalent histone modification, the accepted model is that a histone "language" may be encoded on histones tail domains and multiple histone modifications could specify unique downstream functions [82, 83]. Nevertheless some recent works from Keji Zhao group are challenging this classical view. In a genome-wide mapping of HATs and HDACs binding on chromatin it was pointed out that both the enzymes are found at active genes with acetylated histones. Moreover HATs and HDACs are both targeted to transcribed regions by RNAP II. Thus, it appears that the major roles of HDACs is to reset chromatin by removing acetylation at active genes and to remove acetyl groups added by HATs at inactive gene promoters to prevent RNAP II from binding [84]. These novel data debate the classical "histone code" model and open a new scenario about the role of histone acetylation in modulating chromatin structure and function.

A number of studies have demonstrated that nucleosomal DNA is generally repressive to transcription and a recent genome wide study in yeast clearly demonstrated that nucleosome occupancy of both promoter and transcribed regions inversely correlate with gene activity, with nucleosome occupancy being maximally reduced at the promoters of active genes. In contrast, the promoters of transcriptionally inactive genes are more densely populated of nucleosomes. [85] Analyzing the distribution profile of nucleosomes and nucleosome variants at a genome scale in yeast organism Bradley R. Cairns in a recent paper classifies promoters into two contrasting architectural categories, 'open' and 'covered', based on two broad types of gene, constitutive and highly regulated, respectively [86]. Constitutive genes typically have a large (~150-base-pair (bp)) nucleosome-depleted region directly upstream of the transcription start site (TSS), within which key *cis*-regulatory sequences reside.

In contrast, at regulated genes in their repressed state, nucleosomes often cover the TSS, the regions flanking the TSS, and most of the binding sites for transcriptional activators. At covered promoters, nucleosomes compete effectively with transcription factors for occupancy of key *cis*-regulatory binding sites, rendering covered promoters more reliant than open promoters on chromatin remodeling. However, at least one binding site is typically exposed in the linker DNA between nucleosomes. This exposed site allows a transcription factor access to the promoter, but chromatin

modification and remodelling are probably required to expose the additional sites under nucleosomes [86].

1.5 The transcription cycle is a multistep process.

Gene regulation mostly happens at the level of transcription of DNA template to produce RNA. Numerous factors regulate transcription by controlling the ability of RNAP II to access, bind and transcribe specific genes in response to specific signals. Modern biochemical and molecular methods coupled with genetics and genomics approaches have identified thousands of factors that participate in regulated transcription. Most of these factors are proteins, but a growing number of them are RNAs. They enable RNAP II to gain access to the gene's promoter, to initiate RNA synthesis at the transcription start site (TSS) of the gene and to generate a productively elongating transcription complex that produces a full-length RNA transcript [87]. The generation of a mature mRNA molecule by RNAP II involves multiple processes, some of which occur sequentially and others in parallel. The primary phases of transcript generation form a so-called "transcription cycle", and include pre-initiation, initiation, promoter clearance, elongation, and termination. Some of these steps are rate limiting and activators could potentially act to increase the rate of transcription (Figure 12).

The transcription cycle starts with pre-initiation complex (PIC) assembly at the promoter (Figure 12 - Step 2). The PIC includes the general transcription factors (GTFs) IID, IIB, IIE, IIF, and IIH and RNAP II, as well as several additional cofactors [88]. TFIID is one of the general initiation factors containing components capable of sequence specific binding to eukaryotic promoter DNA. These components include TBP, the main functional subunit of TFIID, which directly recognizes the TATA element, and certain TBP-associated factors (TAFs), which have been implicated directly or indirectly in sequence specific DNA binding. After PIC assembling, the DNA is unwound (Figure 12 – Step 3), and RNAP II initiates transcription. Formation of an open complex between RNAP II and the DNA template is a prerequisite for transcription initiation. Melting of the double-stranded DNA into a single-stranded bubble is an ATP-dependent process and requires the action of two GTFs, IIE and IIH [88].

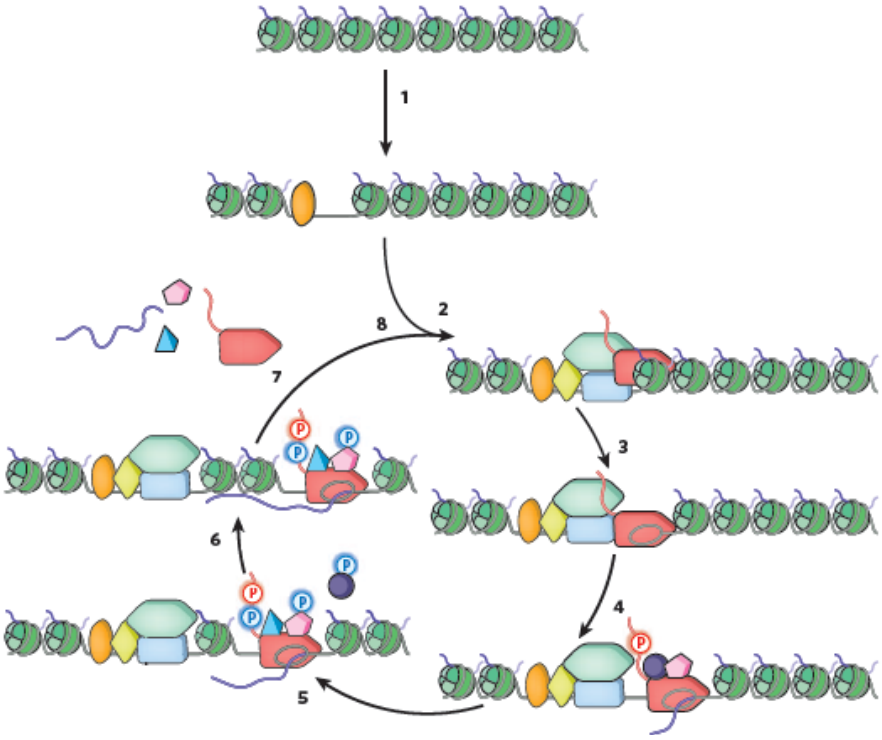

Figure 12. The transcription cycle. Step 1: chromatin opening. Step 2: PIC formation. Step 3: initiation. DNA is unwound. Step 4: promoter escape/clearance. Step 5: escape from pausing. Step 6: productive elongation. Step 7: termination. Step 8: recycling. See text for details [87].

Before RNAP II becomes engaged into productive transcript elongation, it must pass through a stage known as promoter clearance (Figure 12 –Step 4). During this stage, the PIC is partially disassembled: a subset of GTFs remains at the promoter, serving as a scaffold for the formation of the next transcription initiation complex [88]. The GTF TFIIH plays a role in facilitating promoter clearance, in part by preventing premature arrest [88]. Once the promoter is cleared, the next round of transcription can be reinitiated. Reinitiation of transcription has the potential to be a much faster process relative to the initial round, and is responsible for the bulk of transcription in the cell [89]. Promoter clearance coincides with the beginning of another cycling event within the transcription cycle: phosphorylation of the C-terminal domain (CTD) of Rpb1, the largest subunit of RNAP II (Fig. 13). The RNAP II CTD contains multiple repeats of the heptapeptide sequence YSPTSPS. The number of these repeats increases with genomic complexity: 26 in yeast and 52 in mammals [88]. During the transcription cycle the CTD of RNAP II undergoes to a cycle of phosphoryulation and dephosphorylation on Ser 2 and Ser 5 of the heptapeptide sequence. The kinases and

the phosphatases responsible for these modifications are respectively Cdk7 (TFIIH kinase activity), Cdk9 (P-TEFb kinase activity) Ssu72 and FCP1. Different states of phosphorylation associate with different phases of transcription. Moreover modification of the CTD affects its conformation and ability to interact with other factors involved in RNA elongation, maturation or termination [90]. These processes are now known to occur co-transcriptionally, and many events during the synthesis of a mature mRNA are co-regulated [91].

Interestingly, chromatin immunoprecipitation (ChIP) analyses in yeast and Drosophila revealed that RNAP II phosphorylated at Ser 5 associates with promoter-proximal regions of transcribed genes and remains constant or decreases towards the 3'-end region [92]. By contrast phosphorylation at Ser 2 predominates in the body of the gene toward the 3'-end. Thus, it appears that CTD phosphorylation at Ser 5 correlates with transcription initiation and early elongation (promoter clearance), whereas Ser 2 phosphorylation is associated with RNAP II further away from the promoter. The phosphastase activity defines the last part of the process in order to guarantee a correct termination and the turnover of the RNAP II [88, 92] (Figure 13).

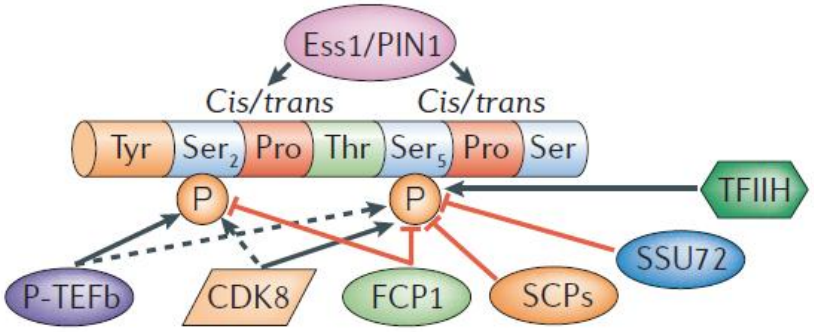

Fig 13. The C-terminal domain (CTD) of the largest subunit of RNA polymerase II Rpb1 and its modification [90].

Once promoter clearance has occurred RNAP II transcribes 20-50 bases downstream of the TSS, produces an RNA and temporary pause at promoter proximal regions. This step functions as checkpoint before committing to productive elongation and assures efficient transcript maturation. While in this "paused state" the C-terminal domain of Rpb1 acts as a docking site for capping and splicing factors in order to facilitate RNA processing [88].

The first well characterized paused polymerase was detected at the *Drosophila melanogaster* heat-shock gene, HPS70 [93]. Under non-inducing conditions, RNAP II fully occupies the HSP70 promoter and transcribes to between +20 and +50 where it pauses. The enzyme resumes transcription only under inducing conditions, such as heat shock. This facilitates rapid induction of the gene's expression. Similar mechanisms have been described also for viral and mammalian genes like c-MYC and FOS [90, 94]. Transcriptional pausing by RNAP II is self-reversible and is thought to be a natural way of transcript regulation. It is probably caused by a structural rearrangement within the enzyme and DNA sequences, which results in the formation of an "inactivated" intermediate [88].

There are many factors that modulate transcriptional pause, and thus, the rate of transcriptional elongation. For different identified model genes, promoter-proximal pausing is mediated by the action of pause factors [95]. These include DRB sensitivity-inducing factor (DSIF) and negative elongation factor (NELF). Both factors were identified by their ability to increase sensitivity to elongation inhibition by DRB, a kinase inhibitor. The factors that alleviate promoter-proximal pausing by DSIF and NELF include TFIIS and the positive transcription elongation factor-b complex (P-TEFb) [95]. In mammalian cells P-TEFb is found in two forms referred to a large and a free small complex. The active small complex contains the Cdk9 kinase which is inhibited by DRB, and cyclin T. In the large inactive complex P-TEFb is associated with the 7SK snRNP [96]. The two complexes are in dynamic equilibrium and perturbation of cellular homeostasis influence the relative presence of the two forms. TFIIS stimulates the intrinsic RNA-cleavage activity of RNAP II to create a new RNA 3'-OH in the active site that enables transcription elongation [90].

The pausing process start when DSIF, an heterodimeric complex composed of the human homologs of *Saccharomyces cerevisiae* Spt4 and Spt5, interacts with RNAP II shortly after initiation (Figure 14). Successively NELF recognizes and binds the RNAPII–DSIF complex thus blocking elongation. The RNAPII pause allows the recruitment of the capping enzyme by the CTD and DSIF (Spt5 subunit), which adds a 5'-cap to the nascent transcript. The activity of the capping enzyme is stimulated by both the CTD and the Spt5 subunit of DSIF. NELF is released by the concerted action of P-TEFb phosphorylation of Spt5 and the CTD on Ser 2 (Figure 14). Further characterization of this process still need to be done.

Moreover, because transcription *in vivo* occurs on nucleoprotein templates, the packaging of DNA into chromatin has profound effects on all processes that require DNA access, including transcription. Chromatin remodeling factor, histone chaperones and histone modifiers are fundamental for efficient transcription. Interestingly, the Facilitates Chromatin Transcription factor (FACT) interacts with PTEF-b in order to alleviate NELF-mediated inhibition of elongation. FACT is a general regulator of chromatin structure which destabilizes nucleosomes selectively removing the H2A/H2B complex allowing the transit of RNAP II along the gene [97]. Besides FACT many proteins control transcription by remodelling chromatin, affecting DNA structure and regulating its topology.

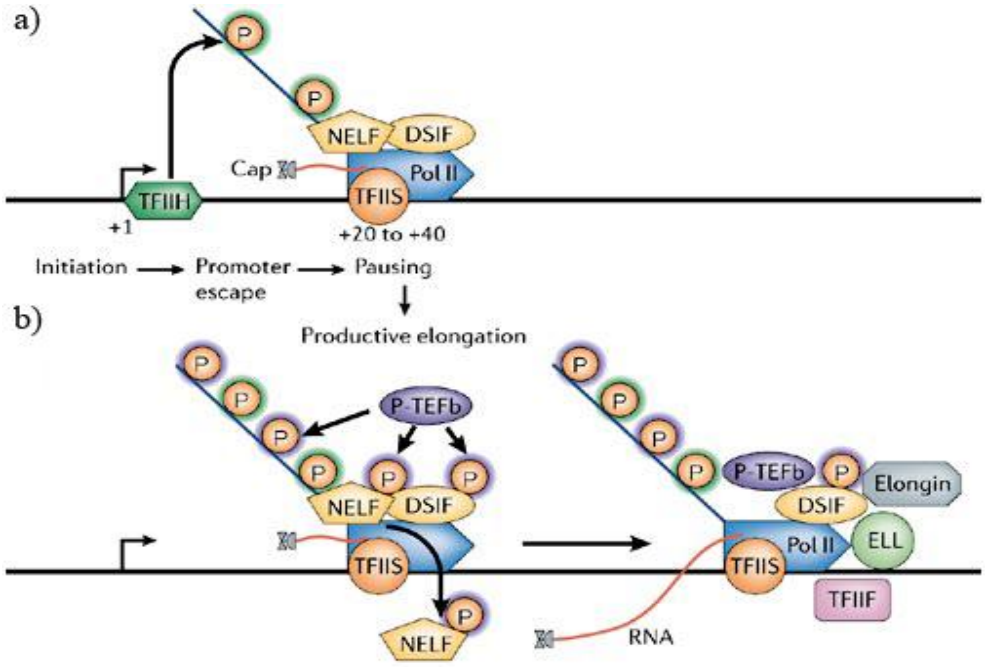

Figure 14. Promoter-proximal pausing and escape to productive elongation. a) The Ser5 of the CTD of RNAP II is phosphorylated by TFIIH during pre-initiation complex formation or before promoter-proximal pausing. DSIF and NELF probably facilitate RNAP II pausing in the promoter-proximal region, and TFIIS also associates with the paused polymerase. TFIIS stimulates the intrinsic RNA-cleavage activity of RNAP II to create a new RNA 3'-OH in the RNAP II active site after backtracking of the polymerase. Capping enzyme associates with the Ser5-phosphorylated CTD and with Spt5 and the nascent RNA becomes capped during this first stage of elongation. b) P-TEFb-mediated phosphorylation of DSIF, NELF and Ser2 of the RNAP II CTD stimulating productive elongation. TFIIS facilitates efficient release of RNAP II from the pause site. Successively NELF dissociates from the transcription complex and DSIF, TFIIS and P-TEFb track with RNAP II along the gene [90].

The final step in the "transcription cycle" is transcript termination. At this stage, the mRNA is cleaved, polyadenylated, and transported to the cytoplasm, where it will be translated [88]. Interestingly, the transcript cleavage-polyadenylation specificity factor (CPSF) can be recruited by TFIID apparently during PIC formation [98]. Thus, transcription initiation and termination are interconnected and might influence each other's efficiency. Moreover elongation arrests also near the 3' flank region to allow co-transcriptional processing by factors recruited to the RNAP II ternary complex [99].

1.6 Hypoxia-Inducible Factor-1alpha: a model to investigate specific gene expression regulation by CPT.

The hypoxia-inducible factor 1 (HIF-1) mediates adaptive responses to changes in tissue oxygenation. This transcription factor is a basic helix-loop-helix heterodimer that consists of a constitutively expressed HIF-1β subunit and a HIF-1α subunit. The latter is highly regulated in its expression at post-translational level. The protein synthesis is regulated via O_2-independent mechanisms, whereas the degradation is regulated primarily via O_2-dependent mechanisms [100]. At high oxygen tensions HIF-1α is subjected to an oxygen mediated hydroxylation of conserved prolyl residues which is required for the interaction with the von Hippel Lindau protein (pHVL). This tumor suppression factor is an E3 ubiquitin-protein ligase that targets HIF-1α for proteasomal degradation [100]. Under hypoxic conditions, HIF-1α protein accumulates and translocates to the nucleus where it forms an active complex with HIF-1β, which activates transcription of target genes important for the adaptation and survival under hypoxia [101].

HIF 1α is a master regulator of cellular response to oxygen deprivation and a target of antiangiogenic and anticancer drugs. Owing to low levels of intratumoral oxygen and cancer specific genetic alterations, HIF-1α is often overexpressed in human solid cancers. It has indeed been shown that HIF-1α protein level and transcriptional activity were positively modulated by numerous proto-oncogene products such as SRC, RAS and HER-2 [100, 102] and negatively by tumour suppressor proteins such as p53 and von Hippel Lindau gene product (pVHL) during normoxia.

Owing to its critical role in cellular adaptation to hypoxia, the HIF-1α gene locus is tightly regulated and is quite complex. For example, it has been reported the presence of a natural antisense transcript (3'aHIF-1α) that is complementary to the 3'-untranslated region of HIF-1α messenger RNA [103]. Under hypoxic condition the initial rise in HIF-1α protein also increases the natural antisense, which in turn destabilized HIF-1α mRNA through a mechanism that involves exposition of AU-rich elements present in the 3'untranslated region of HIF-1 α mRNA. This leads to the decrease in protein expression. Furthermore, owing to the presence of several putative hypoxia response elements in the antisense transcript promoter, the expression of HIF-1α is subjected to a negative feedback regulation [104].

In the recent years a big interest in the role of HIF-1α has grown since it was established that it regulates transcription of genes that are involved in crucial aspects of cancer biology, including angiogenesis, cell survival, glucose metabolism and invasion. Interestingly, in a recent paper published by Rapisarda *et al.* the camptothecin analogue topotecan (TPT), a TOP 1 poison, was shown to inhibit HIF-1α transcriptional activity and HIF-1 α protein accumulation in hypoxia-treated U251 human glioma cells [101]. The authors demonstrated that TPT does not affect HIF-1α protein half-life or mRNA accumulation but inhibits its translation, in a manner dependent by the presence of TOP 1 in the cells. Experiments performed with aphidicolin, a potent inhibitor of replicative DNA polymerases, indicated that TPT inhibited HIF-1α protein accumulation also in the absence of DNA replication double strand breaks.

1.7 Aim of the Thesis research project.

Previous findings of our lab [35] have led to the hypothesis that TOP 1 is involved in the regulation of RNAP II at the promoters of transcribed genes in human cancer cells (see also Figure 10).

Thus, the aim of the Thesis research project was to test the above hypothesis and to establish whether or not TOP 1 plays a role in RNAP II pausing and elongation. To this end, we used camptothecin as a specific inhibitor of TOP 1. In addition, we intended to establish the molecular mechanism triggered by camptothecin at trascriptional levels. The findings published by Rapidarda and colleagues [101] provided an intriguing challenge to us. If TOP 1 inhibition by camptothecin impairs RNAPII promoter-proximal pausing regulation, that might deregulate transcript maturation and eventually gene expression and gene product activity. Thus, in order to prove this hypothesis we decided to use the human HIF-1α gene as a model. The findings may indeed give remarkable insight into HIF-1α regulation in cancer cells as several aspects of its tight control are still unknown.

Thus, we first aimed at verifying the presence of a RNAP II pausing site at the HIF-1α gene promoter, and the specificity of CPT effects on RNAP II in living cells. We investigated the molecular dynamics of RNAP II at promoter proximal pause site and further downstream regions in CPT-treated cells in different physiological and pathological conditions. To test our model about TOP 1 regulation of promoter pausing site, we developed a specific strategy that enables to catch nascent RNAs bound to chromatin, and we determined the levels of transcription immediately downstream to promoter pausing site in CPT-treated cells. Finally, based on the specific outcome of the study, we investigated nascent transcripts in other regions of HIF-1α gene locus showing an unexpected and intriguing phenomenon. Most of the experiments were done on HCT 116 cells, a colorectal cancer cell line, as camptothecin derivatives are clinically used in this type of solid tumor. However, we have shown that our findings were independent from a specific cell type. Moreover, the findings suggested a new mechanism of HIF-1α regulation that may explain the findings reported by Rapisarda *et al*. previously [101].

After focusing our TOP 1 investigations at a gene specific level, we decided to extend the study to the whole genome. Thus, we next aimed at defining TOP 1 binding site across the genome, since defining a protein's binding site may greatly help in establishing its role. We developed a new mapping approach, taking advantage of the enzyme's mechanism of action. Since the enzyme has to break the double helix to exert its catalytic activity, we used this feature to develop a technique that allows the mapping of TOP 1 cleavage sites. In comparison to a ChIP-Seq approach, which gives information only about protein localization, this novel strategy (which we will name "Topo-Seq") revealed the sites at which the enzyme is active. The knowledge of the genomic sites of TOP 1 activity also provides remarkable insights into the physiological and pathological roles of TOP 1.

Chapter 2

Materials and Methods

2.1 Cell Lines and Reagents.

Human colorectal cancer HCT116 cell lines were cultured in Dulbecco's modified Eagle's medium (Euroclone) containing 10% heat-inactivated fetal calf serum (Invitrogen) and 500 μg/ml hygromycin B. HCT116 (top1siRNA) cells have been described previously [49]. Human MRCV fibroblast cells were routinely cultured in Dulbecco's modified Eagle's medium containing 10% heat-inactivated fetal calf serum. Human Jurkat T cells were maintained in RPMI1640 supplemented with 10% heat-inactivated fetal calf serum and 2 mM glutamine (Invitrogen). Cells were maintained at 37°C in a humidified incubator containing 5% CO_2 in air (referred to as normoxic conditions, 20% O_2). Hypoxia treatments were performed by placing cells in a modular incubator chamber (HERAcell^{R}150, Kendro Laboratory Products), and then flushing with a mixture of 1% O_2, 5% CO_2, and 94% nitrogen for 20 min. The chamber was then placed at 37°C.

All drug treatments were performed on exponentially-growing cells. Cells were exposed to camptothecin, cisplatin or etoposide at 37°C for the indicated time periods and concentrations. In case of co-treatments, cells were first incubated with aphidicolin (3 μM) or 5,6-dichloro-1-β-Image–ribofuranosylbenzimidazole (50 μM) for 15 min, or caffeine (5 mM) for 30 min. Then, camptothecin was added to the medium and cells incubated for the indicated time. All drugs were purchased from Sigma.

2.2 RNA Purification and primer specific cDNA preparation.

After drug treatments, $5x10^7$ cells were washed twice with cold PBS and collected through centrifugation. The pellet was resuspended and well mixed in 3.6 ml AE buffer [50 mM NaOAc (pH 5.2), 10 mM EDTA], 240 μl of SDS 25% and 3.6 ml of acid phenol (pH 4.5). Samples were then incubated for 10 min at 65°C mixing vigorously every minute. After a short incubation on ice, samples were centrifuged for 15 min at 12,000 g. The upper phase was collected; 3.9 ml of chloroform/isoamylic alcohol was added to it, then mixed and centrifuged for 10 min at 1,800 g. The upper

phase was precipitated with isopropanol and NaOAc. The pellet was resuspended in TE, and DNA was digested with DNAse I (Fermentas). Finally, RNAs were purified with phenol and precipitated with ethanol and NaOAc.

RNA integrity was established by running 1% agarose gel electrophoresis. Then, 1 μg of total RNA was used to prepare cDNA using SuperScript III (Invitrogen) with reaction buffers suggested by the manufacturer, for 5 min at 65°C, 5 min at 25°C, and 60 min at 50°C, followed by incubation with RNase H. When a specific primer was used for reverse transcription with SuperScript III, conditions for the reaction were as follows: 5 min/ at 65°C and 50 min at 55°C. As negative controls, we used cDNA prepared with no primer during reverse transcription reactions.

2.3 Northern Blotting.

For Northern blot analysis, 20 μg RNA for each sample were separated by 1.5% agarose-formaldehyde gel electrophoresis in MOPS buffer [20 mM MOPS (pH 7.0), 2 mM NaOAC, 1 mM EDTA (pH 8.0), in DEPC-treated water]. Then, RNA was transferred to a Hybond-N membrane in 20X SSC buffer (3 M NaCl, 0.3 M sodium citrate), and then fixed to the membrane by UV-crosslinking and 1 hour baking at 80°C. Strand-specific probe was designed to be 46 nucleotides in length and was synthesized by Intregrated DNA Technologies (IDT) with the 3' StarFire extension (IDT) (5'- CAG CCC CAA TTC TAA ATA AGC TCT TAG ATT TTC CTC AGC C/NNNNNN/ -3'); The probe was labeled using StarFire kits from IDT using the manufacturer's instructions. Hybridization was performed for 16 h in 7% SDS, 500 mM Sodium Phosphate (pH 7.0), 1 mM EDTA, 25% formamide at 42°C followed by two washes in 2X SSC, 0.5% SDS at room temperature for 20 min.

The HIF-1α mRNA exon 2 (corresponding to the 25-kb amplicon shown in Figure 1-Results) was cloned by PCR fragment into the TOPO-TA vector (Invitrogen). Plasmid DNA was isolated with QIAGEN kit according to manufacturer protocol, and digested with EcoRI (New England Biolabs), and the exon 2 insert was then purified by gel electrophoresis and electroelution. The probe was labeled using Ready-To-Go DNA Labelling Beads (GE Healthcare) using manufacturer's instructions. Hybridization was performed for 16 h in 7% SDS, 500 mM Sodium Phosphate (pH 7.0), 1mM EDTA at 65°C followed by two washes [1% SDS, 40 mM Sodium Phosphate (pH 7.0)], at room temperature for 30 min and then at 65°C for 45 min.

Membranes were then exposed to Phospho Screen (Biosciences Amersham) for 1-2 days and then developed using Storm 860 (Biosciences Amersham).

2.4 Chromatin immunoprecipitation (ChIP).

To prepare chromatin, $1x10^7$cells were fixed with 1% formaldehyde for 15 min. The reaction was stopped with 0.125 M glycin and cells were washed twice with ice-cold PBS followed by 7 ml of TEET [10 mM Tris-HCl (pH 8.0) 10 mM EDTA, 0.5 mM EGTA, 0.25% Triton X-100], 5 ml of TEEN [10 mM Tris-HCl (pH 8.0) 10 mM EDTA, 0.5 mM EGTA, 200 mM NaCl], and resuspended in 0.5 ml of TEE [10 mM Tris-HCl (pH 8.0) 10 mM EDTA, 0.5 mM EGTA]. Protease inhibitors [aprotinin, leupeptin and pepstatin (Sigma) 10 µg/ml] were added to the buffers immediately before use. Chromatin was then sheared by sonication using a Bioruptor (Diagenode) to an average DNA fragment size of 300–400 bp.

Immunoprecipitations were performed at 4°C in RIPA buffer [50 mM Tris-HCl (pH 8.0), 1 mM EDTA, 0.5 mM EGTA, 150 mM NaCl, 1% Triton X-100, 0.1% Na-deoxycholate, 0.1% SDS]. Amounts of chromatin, equivalent to 0.4 OD at 260 nm were taken for each immunoprecipitation. Samples were precleared for 2 hour with 4 µg of non-immune rabbit IgG and 20 µl of 50% suspension of a 1:1 mix of Protein A- and Protein G-Sepharose beads. Then, chromatin was separated by beads and recovered by centrifugation for 3 min at 1,000g. One tenth of the supernatants were saved as "10% input." Supernatants were incubated overnight with 4 µg of specific antibody or nonimmune rabbit IgG (to measure background recovery). ChIP-grade Abs against anti-acetylated K9 and K14 of H3 histone or anti-acetylated K5, K8, K12, K16 of H4 histone were purchased from Upstate (Lake Placid, NY), and Abs against anti dimeK9 of H3 histone or anti H4 histone were purchased from Abcam. The H-224 Ab Antibody against the N-terminal of RNAPII large subunit and the SI-1 against TBP were from Santa Cruz Biotechnology (Santa Cruz, CA). Non-immune rabbit IgG were from Cedarlane (Hornby, Canada). Immunocomplexes were recovered by addition of 40 µl of Protein A-/Protein G-Sepharose beads blocked with DNase-free BSA (9.95 mg/ml) and salmon testes DNA (10.5 mg/ml). Then, the beads were washed four times with RIPA buffer; once with RIPA buffer containing 0.5 M NaCl; once with Li250 buffer [10 mM Tris-HCl (pH 8.0), 1 mM EDTA, 0.25 M LiCl, 0.5% Na-deoxycholate, and 0.5% NP40]; twice with TE [10 mM Tris-HCl, 1 mM EDTA

(pH 8.0)]; and finally resuspended in TE. Each wash was performed for 10 min by rocking at 20 rpm followed by 3 min of centrifugation at 1,000g.

The pellets were then adjusted to 0.5% SDS and incubated overnight at 65°C to reverse cross-links. Samples were then digested with proteinase K (500 µg/ml Sigma) for 4 h at 52°C and extracted twice with phenol chloroform. DNAs was precipitated with ethanol in the presence of 20 µg of glycogen (Roche Diagnostics, Manheim) and dissolved in TE.

Recovered DNA was quantified by real-time PCR (see Table 1 for the list of primers). At least three dilutions of input DNA were run to generate the standard curve. DNA recovery was measured as input DNA fraction.

2.5 RNA immunoprecipitation (RIP).

For RIP experiments, $5x10^7$ cells were cross-linked with formaldehyde at a final concentration of 1% added directly to the medium for 15 min. The reaction was stopped with 0.125 M glycin. Cells were washed twice with cold PBS, scraped and collected. Cell pellets were resuspended in 1 ml of RIPA buffer [50 mM Tris-HCl (pH 7.5), 1% NP-40, 0.5% sodium deoxycholate, 0.05% SDS, 1 mM EDTA, 150 mM NaCl] containing protease inhibitors [aprotinin, leupeptin and pepstatin (Sigma) 10 µg/ml] and RNasin [(Promega) 50U/500ml]. Extracts were sonicated with a Branson 250 sonifier (Branson Ultrasonic Corp., Danbury, CT) at 30% amplitude with 8 x 10-s bursts and 30-s pauses, resulting in an average fragment size of 1,000 nucleotides. Chromatin was then centrifuged for 10 min at 15,000 g to remove insoluble materials. Amounts of chromatin equivalent to 4 OD at 260 nm were taken for each immunoprecipitation.

Chromatin samples were then pre-cleared for 30 min with 20 µl of 50% suspension of a 1:1 mix of Protein A- and Protein G-Sepharose beads. Then, chromatin was separated by beads and recovered by centrifugation for 3 min at 1,000g. One tenth of the supernatants were saved as "10% input." Supernatants were incubated overnight with 10 µg of specific antibody or nonimmune rabbit IgG (to measure background recovery) and immunoprecipitate by addition of 60 µl of Protein A-/Protein G-Sepharose beads blocked with DNase-free BSA (9.95 mg/ml) and salmon testes DNA (10.5 mg/ml). The final pellets were washed by rocking for 4 min once in each of the following buffers: low-salt immune complex wash buffer [0.1%

SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl, (pH 8.1), 150 mM NaCl], high-salt immune complex wash buffer (the previous buffer with 500 mM NaCl) and LiCl immune complex wash buffer [0.25 M LiCl, 1% NP-40, 1% deoxycholic acid, 1 mM EDTA, 10 mM Tris-HCl (pH 8.1)]; and twice in TE [10 mM Tris-HCl, 1 mM EDTA]. The immunocomplexes were eluted in TE-SDS 1% and digested with proteinase K for 1 hour at 52°C. Cross-links were reversed at 65°C for 5 h. RNA was extracted twice with phenol chloroform and treated for 2 h with 4 Units of DNase I (Fermentas). RNA was purified with phenol chloroform, precipitated with ethanol in the presence of 20 µg of glycogen (Roche) and treated again with 4 Units of DNase I for 2 h. After a second round of phenol extraction and precipitation cDNA was then prepared from 66% of the RNA sample using SuperScript III and random primers (Invitrogen). The conditions used for retrotranscription were as follows: 5 min at 65°C, 5 min at 25°C, and 60 min at 50°C. RNA was hydrolyzed with 0.2 N NaOH and 0.1 M EDTA at 65°C for 7 min. HEPES buffer was then added to a final concentration of 0.3 M.

cDNA was analyzed by quantitative PCR with primer pairs (see Table 1 for the list of primers) spanning the HIF-1α gene (results, Figure 1). As negative controls, we used cDNA prepared with no primers, and non-retrotrascribed samples. At least three dilutions of genomic DNA were run to generate the standard curve. For all specific antibodies, the RNA recovered values were at least 10-fold more enriched than non-immune controls, and background levels set by NP and α-satellite DNA (results, Figure 5). Moreover, recovered RNA values were at least 100-fold more enriched than non-retrotranscribed samples. Final results are the means of at least 4 determinations from at least two independent RIP experiments.

2.6 Quantitative Real-Time PCR (qPCR).

Real-time PCR was performed by using the LightCycler and the FastStart DNA Master SYBR Green I kit (Roche Diagnostics, Mannheim). PCR reactions contained 1x FastStart DNA SYBR Green I Master Mix, 2.08 mM $MgCl_2$ and 350 nM each of primers. Specificity of PCR products was routinely controlled by melting curve analysis and agarose gel electrophoresis.

Quantification and melting curve analyses were performed using the Roche LightCycler software by the crossing point method as indicated by the supplier. The PCR reaction performed can be described as

$$C_T = C_O * E^n$$

Where

- C_T is the concentration of a sample which is identified by a given cycle number
- C_0 is the starting concentration of the sample
- **E** is the overall reaction efficiency
- **n** is the number of cycles

The crossing point method generates a standard curve by plotting the crossing cycle number versus the logarithm of the concentration of each given standard sample.

2.7 Total Protein Extraction.

Cells were washed twice with ice-cold PBS and lysed with RIPA buffer with protease and phosphatase inhibitors (2 mM DTT, 1 mM Pefabloc, 1 mM NaVanadate, 1 mM NaF, 4 μg/ml pepstatin, 4 μg/ml leupeptin, and 4 μg/ml aprotinin). The cell suspension was briefly sonicated to disrupt membranes and cellular debris, which were then pelleted by centrifugation at 15,000 g for 15 min at 4°C. The supernatant was saved for Western blotting. Protein concentration was measured with the DC Protein Assay (Biorad) according to the manufacture's instructions.

2.8 Histone Protein Extraction.

Cells were washed twice with ice-cold PBS scraped and lysed with Lysis buffer [10 mM Tris-HCl (pH 7.5), 1 mM $MgCl_2$, 0.5% NP-40] containing protease inhibitors [aprotinin, leupeptin, pepstatin (10 μg/ml) and Phosphatase Inhibitor Cocktail II (Sigma)]. Cellular debris were pelleted by centrifugation at 16,000 g for 1 min at 4°C, resuspended and incubated for 15 min with Lysis buffer with 400 mM NaCl. Samples were briefly centrifugated, and pellets were incubated for 10 min at 4°C with 5 volumes of Extraction solution [220 mM H_2SO_4, 20% Glycerol, 10 μg/mL 2-mercaptoethanolammine]. The histone-containing supernatant was obtained by centrifugation at 16,000 g for 10 min. Histones were pelleted from the supernatant by

adding 20% trichloroacetic acid, and centrifugation. Then the pellets were resuspended in 100% ethanol and centrifuged again at 16,000 g for 20 min. The pelleted histones were store at -80 °C. Protein concentration was measured with the DC Protein Assay (Biorad).

2.9 Western Blotting Analyses.

Whole cell lysates, corresponding to 20 μg of proteins were separated by 6.5% SDS PAGE. Proteins were then blotted onto a Hybond ECL-nitrocellulose membrane. Histone aliquots, corresponding to 10-15 μg were loaded onto a 12% SDS-PAGE gel, and then transferred to a Hybond ECL-nitrocellulose membrane. Equal loading was checked by Ponceau staining.

Abs specific for Top1 or β-Actin were purchased from Santa Cruz Biotechnology (Santa Cruz, CA), and used at a 1:400 and 1:2,000 diluition, respectively. Abs specific for total p53 and CHK2 were used at a 1:1,000 dilution (Cell Signaling Technology Inc., Beverly, MA).The Abs anti γ-H2AX or H1 (Santa Cruz Biotechnology) were used at a 1:1,000 and 1:500 dilution, respectively.

Specific bands were then detected with with ECL Plus Western blot imaging system (GE Healthcare). Horseradish peroxidase-conjugated mouse and rabbit IgG (1:2,000 and 1:5,000 dilution respectively) were purchased from GE Healthcare. Horseradish peroxidase-conjugated goat (1:40,000 diluition) was from Santa Cruz Biotecnology.

2.10 TOP 1 cleavage sites mapping (“Topo-seq”).

2.10.1 Preparation of DNA for sequencing.

To prepare high molecular weight DNA $8x10^7$ cells were treated with or without CPT (20 μM Sigma), washed twice with ice-cold PBS and lysed with 10 mL of lysis buffer [10mM Tris-HCl (pH 8.0), 100mM EDTA (pH 8.0), 0.5%SDS]. Lysates were then collected and digested overnight with proteinase K (200 μg/ml Roche) at 55°C. DNA was extracted twice with phenol and once with phenol chloroform, and ethanol precipitated in the presence of 2 M ammonium acetate (Sigma). The sample was then digested with 5 μg of pancreatic RNase (Roche) for 1 hour at 37°, then adjusted to 0.5% SDS and incubated for 1 hour at 55°C with

proteinase K (200 μg/ml). DNA was extracted twice with phenol and precipitated with ethanol in the presence of 2 M ammonium acetate. Finally the pellet was resuspended in 1 mL of TE [10 mM Tris-HCl, 1 mM EDTA (pH 8.0)] and incubated at roon temperature for 12 h with gentle rotation.

The presence of CPT induced single strand DNA breaks (SSBs) was established by sonicating (Bioruptor /Diagenode) for 10s 1 μg of the sample at low potency and analyzed by electrophoresis on 0.6% agarose gel.

To remove the TOP 1 covalently linked to the DNA, the sample was treated with 30 μg of Tyrosyl-DNA phosphodiesterase 1(TDP1, gently offered by Yves Pommier/Center for Cancer Research, NCI, Bethesda, MD) in TDP1 buffer [PBS 1x, 80 mM KCl, 0.01% tween] for 2 h at 37°C, extracted with phenol and precipitated with ethanol and ammonium acetate. The DNA was incubated for 4 h with 500 units of T4 polynucleotide kinase (New England Biolabs) in the presence of 1mM ATP. The DNA was twice extracted with phenol and precipitated with ethanol and ammonium acetate to completely remove free ATP.

To label TOP 1 induced SSBs we performed nick translation. With this purpose, 500 μg of DNA was incubated for 40 s at 16°C with a mixture of 200 μM of dATP, dGTP, dCTP and 20 μM of digoxigenin-11-dUTP (Roche), 117 μM of ddATP, ddGTP, ddCTP (Roche) and 1,000 units of DNA Polymerase I (from *E. coli*-New England Biolabs). The reaction was stopped with 50 μM EDTA and extracted with phenol and the sample was precipitated twice in the presence of ethanol and 2.5 M ammonium acetate.

DNA was then sheared by sonication using a Bioruptor to an average DNA fragment size of 200–300 bp. One tenth of the supernatant was saved as "10% input" and the rest of the sample was incubated at 4°C overnight with 10 μg of Anti-digoxigenin antibody (Roche) with gentle rotation.

Immunocomplexes were recovered by addition of 60 μl of Protein G-Sepharose beads (Roche) and incubated for 4 h at 4°C. The beads were washed once with PBS, thrice with NP-40 buffer [20 mM Tris-HCl (pH 8.0), 137 mM NaCl, 10% Glycerol, 1% NP-40, 2 mM EDTA (pH 8.0)]; twice with TE [10 mM Tris-HCl, 1 mM EDTA (pH 8.0)]; and finally resuspended in 200 μl of TE. Each wash was performed for 10 minutes by gentle agitation followed by 4 minutes of centrifugation at 1,500g.

The pellet was adjusted to 0.5% SDS and digested with proteinase K (200 µg/ml) at 65°C overnight. DNA was purified using QIAquick PCR Purification Kit Protocol (QIAGEN) according to the manufacture's instructions and finally quantified.

Before Solexa sequencing the sample was analyzed by qPCR using primer sets reported in table 1.

2.10.2 Preparation of the library.

The DNA recovered from the previous steps was taking into a library to be sequenced on the 1G Genome Analyzer. Briefly, this is the procedure to prepair the library:

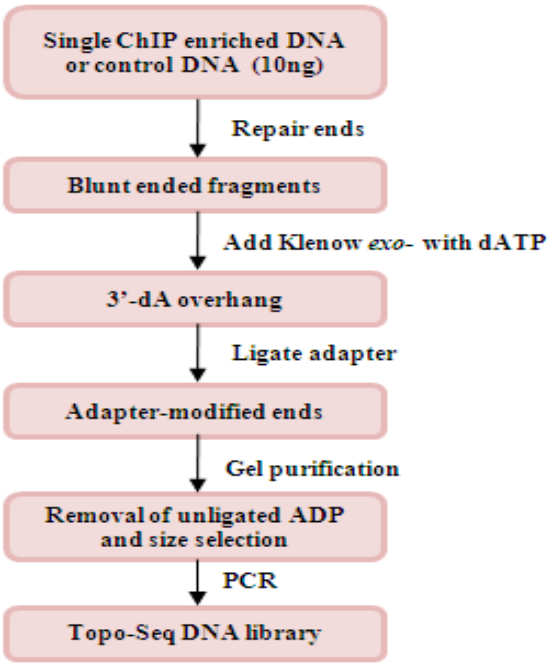

Figure 1. Sample preparation Workflow. See text for details.

The Epicentre DNA END-Repair kit (Epicentre Biotechnologies) was used to generate blunt-ended DNA. 0.1 to 5 µg of DNA were incubated for 45 min at room temperature with a mixture of End repair buffer [33 mM Tris-acetate (pH 7), 66 mM potassium acetate, 10 mM magnesium acetate, 0.5 mM DTT], 0.25 mM of each dNTPs 1 mM ATP, 1 µl End-Repair Enzyme mix (T4 DNA polymerase + T4 PNK). DNA was extracted with phenol chloroform, precipitated and resuspended with 30 µL of TE.

To add “A-base” to 3` ends of DNA fragment the sample was incubated for 30 min at 70°C with a mixture of 1x Tag buffer, 0.2 mM dATP and 15 units of Taq DNA polymerase. DNA was extracted with phenol chloroform, precipitated and finally resuspended in 10 µL of TE.

To ligate adapters to DNA fragments two mixed linkers were used. Ligation was carry out with 300ng of DNA which were incubated with T4 DNA ligase buffer, 0.1 µl Adaptor oligo mix and 1,000 units of T4 DNA ligase The reaction was performed at room temperature for 30 min and then over night at 16°C. After two steps of DNA purification using QIAquick PCR Purification Kit the DNA was eluted in 30 µl of elution buffer. A size selection of the linker ligated DNA was performed through 2% E-Gel (Invitrogen) electrophoresis. The gel was sliced around 200-400 bp region, then DNA was extracted using minelute gel extraction kit (QIAGEN) in a final volume of 12 µl elution buffer.

The DNA was then amplified in a total volume of 25 µl using Solexa primers (Fw: 5'-aca ctc ttt ccc tac acg acg c-3'/ Rv: 5'-caa gca gaa gac ggc ata cga gc-3') and enzyme mix with reaction buffer as suggested by the manufacturer. The DNA amplification was performed for 18-21 cycles according to the following protocol: denaturation at 98°C for 30 s; 98°C for 10 s; 65°C for 30 s; 72°Cfor 30 s. The amplified PCR product is obtained by excising a 220 bp band from a 2.5% agarose gel and purifying it through Qiagen gel extraction kit (QIAGEN).

2.10.3 Cluster Generation.

DNA is denatured into single strands with 0.1 N NaOH to a final DNA concentration of 0.5 nM. The incubation goes for 5 min at room temperature which makes the DNA suitable to hybridize the flowcell during cluster generation.

To generate the clusters different solution are prepared and inoculated into 0.2 ml eight strip tube or reagent tubes in various sizes, ready to be used by the machine.

Amplification Pre-Mix is prepared by mixing 15 ml of water, 3 mL of Cluster buffer and 12 mL of 5M Betaine.

Initial Extension Mix is prepared by mixing 975 µl of Amplification Pre-Mix, 20 µl of 10 mM dNTPs, 5 µl of Taq DNA Polymerase.

Amplification Mix using Bst DNA Polymerase is prepared by mixing 12 ml of Amplification Pre-Mix, 240 µl of 10 mM dNTPs, 120 µl of Bst DNA Polymerase.

Linearization Mix is done by mixing 1518 µl of water with Sodium periodate, and adding to it 60 µl 1.0 M Tris, 1,500 µl formamide and 2.3 µl of 3APL.

Blocking Mix is made with 1510.2 µl of Blocking Buffer, 1X 30.1 µl of 130 µM ddNTPs 19.7 µl of Terminal Transferase**.**

Sequencing Primer Mix is made with 1313.4 µl of Hybridization buffer and 6.6 µl Sequencing primer.

Cluster generation consists of the following steps:

1. Hybridize template DNA—Hybridize template molecules onto the oligonucleotide-coated surface of the flow cell.

2. Amplify template DNA—Isothermally amplify the molecules (for 25 cycles) to generate clonal DNA clusters.

3. Linearize—Chemically linearize the dsDNA clusters. This is the first step of converting dsDNA to ssDNA that is suitable for sequencing.

4. Block—Block the free 3'-OH ends of the linearized dsDNA clusters. This prevents nonspecific sites from being sequenced. After this process, the flow cell is stable and can be stored.

5. Denature—Convert the dsDNA to ssDNA.

6. Hybridize sequencing primers—Hybridize a sequencing primer onto the linearized and blocked clusters.

After this step, DNA clusters will be generated and the flow cell will be ready for sequencing.

2.10.4 Genome Analyser.

The resulting high-density array of templates on the flow cell surface is sequenced with the fully automated Solexa 1G Genome Analyzer. Templates undergo sequencing by synthesis in parallel using proprietary fluorescent labeled reversible terminator nucleotides. Briefly, the sequencing of DNA clusters bound to the flow cell consists of the following steps (Figure 2):

1) Determine first base- All four labeled reversible terminators, primers and DNA polymerase enzyme are added to the flow cell. The first cycle of sequencing consists first of the incorporation of a single fluorescent nucleotide.

2) First base recording- After laser excitation, the high resolution image of the emitted fluorescence from each cluster on the flow cell is captured. The identity of the

first base for each cluster is recorded. These images represent the data collected for the first base. Any signal above background identifies the physical location of a cluster, and the fluorescent emission identifies which of the four bases was incorporated at that position.

3) Determine second base- To initiate the next sequencing cycle, all four labeled reversible terminators and enzyme were added to the flow cell.

4) Second chemistry cycle- After laser excitation the image data is collected as before. The identity of the second base for each cluster is recorded.

5) Sequence reads over multiple chemistry cycle- Repeated cycles of sequencing to determine the sequence of bases in a given fragment a single base a time is performed.

6) Align the data- Illumina's ELAND alignment algorithm is used for downstream match reads to the genome

2.10.5 Solexa Pipeline Analysis.

Sequence tags were obtained and mapped to the human genome using the Solexa Analysis Pipeline. The output of the Solexa Analysis Pipeline was converted to browser extensible data (BED) files for viewing the data in the UCSC genome browser. Data for TOP 1 cleavage sites are presented in BED files detailing the genomic coordinates of each tag as well as in BED files detailing summary windows displaying the number of tags in 400 bp windows.

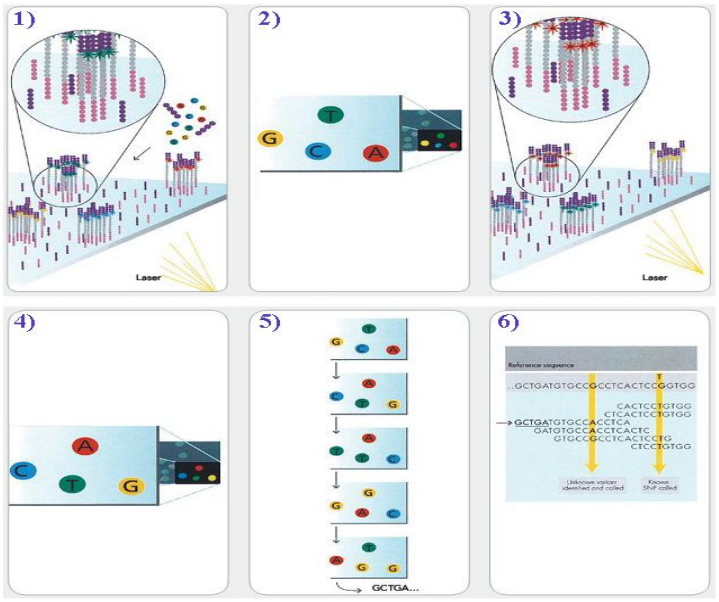

Figure 2. Sequencing by synthesis. See text for details.

Position	Forward primer	Reverse primer
HIF-1α -0.5 kb	TGAACAGAGAGCCCAGCAGAG	CCTGGTCCCAAACATGCATC
HIF-1α 0.05 kb	AGCTCCTCAGTGCACAGTGC	AGACTAGAGAGAAGCGGGCG
HIF-1α 0.1 kb	AGGATCACCCTCTTCGTCGC	AAGGCAAGTCCAGAGGTGGG
HIF-1α 0.2 kb	AGGATCACCCTCTTCGTCGC	CCGAGGGAATGGGCTTACTT
HIF-1α 2.1 kb	CTCTTAGATTTTCCTCAGCC	GCTGAGTAACCACCACTTAT
HIF-1α 7.9 kb	GAGGGGAAAATGTAGTCATTGGC	CGTCCTCTTCCACACCATACAGA
HIF-1α 8.4 kb	CCAAGTGTAGTCATTCTGCAAGC	CTTACTTTGTGCCAAGAGCTGTTC
HIF-1α 24.7 kb	TCTCTCCAATTACATATGCTGG	TCTCGAGACTTTTCTTTTCG
HIF-1α 25 kb	AGCCAGATCTCGGCGAAGTA	CCAGAAGTTTCCTCACACGC
HIF-1α 45kb	CCAGTTACGTTCCTTCGATCAGT	TTTGAGGACTTGCGCTTTCA
3'aHIF-1α	TTTGTGTTTGAGCATTTTAATAGGC	CCAGGCCCCTTTGATCAGCTT
c-MYC 0 kb	AGAAGGGCAGGGCTTCTCAGA	TCTGCCTCTCGCTGGAATTAC
c-MYC 1.9 kb	TAGCTTCACCAACAGGAACT	AGCTCGAATTTCTTCCAGAT
c-MYC 4.7 kb	AGCCACAGCATACATCCTGTC	CTCAGCCAAGGTTGTGAGGTT
GAPD 0 kb	TAGCTCAGGCCTCAAGACCTT	AAGAAGATGCGGCTGACTGTC
GAPD 2 kb	CTTGCCTCTTGTCTCTTAGAT	TGTAGCACTCACCATGTAGTT
α-sat	CTTTTTCATCATAGGCCTCAA	AGCTCACAGAGCTGAAACATT
LMNB2-B48b	CCAGAATCCGATCATGCACC	TCCGTTTTTGCAGGTTGTGCT
LMNB2-B48	AGATGCATGCCTAGCGTGTTC	TAGCTACACTAGCCAGTGACCTTTTTCCT

Table 1. List of Primers.

Chapter 3

Results

3.1 CPT affects the distribution of RNAP II at gene promoter in a TOP 1 dependent manner.

Previous results from our lab showed a novel effect of TOP 1 inhibition by CPT on RNAP II in the cell [35]. Early, following drug addition to the medium, campthotecin affected the distribution of RNAP II around the 5` end of several transcribed genes. Time course experiments followed by ChIP analysis revealed that upon 10 minutes of CPT treatments RNAP II levels decrease at promoter-proximal pausing regions and increase in the body of the gene (introduction, Figure 10). This effect is completely reversed by the Cdk inhibitor DRB showing a correlation between drug inhibition of TOP 1 and pathways that regulates transcriptional pausing [35].

As the findings apparently suggest that TOP 1 inhibition enhances RNAP II escape from pausing sites, and since promoter pausing is a fundamental step of the transcription cycle, we decided to test this hypothesis by analyzing RNAP II dynamics at the human HIF-1α gene locus. We selected this genetic locus as findings by G. Melillo et al. showed that TOP 1 inhibition by CPT fully abolish HIF-1α protein accumulation during hypoxia but not the level of gene transcription [101]. Thus, we hypothesized that if CPT induces a deregulation of RNAP II at the promoter pausing site of the HIF-1α gene, then the deregulation of RNAP II pausing might result in an aberration of mRNA maturation leading to lack of translation of the messenger into an active protein.

The Human HIF-1α locus is 52.738 kb long and is constituted of 15 exons. The intron size ranges from 0.084 to 4.8 kb, with the exception of the first intron, which is 24.542 kb [105] (Figure 1). First, we determined the presence of a RNAP II pausing site at the promoter of HIF-1α and then we assessed CPT influence on RNAP II at promoter and in the body of the gene.

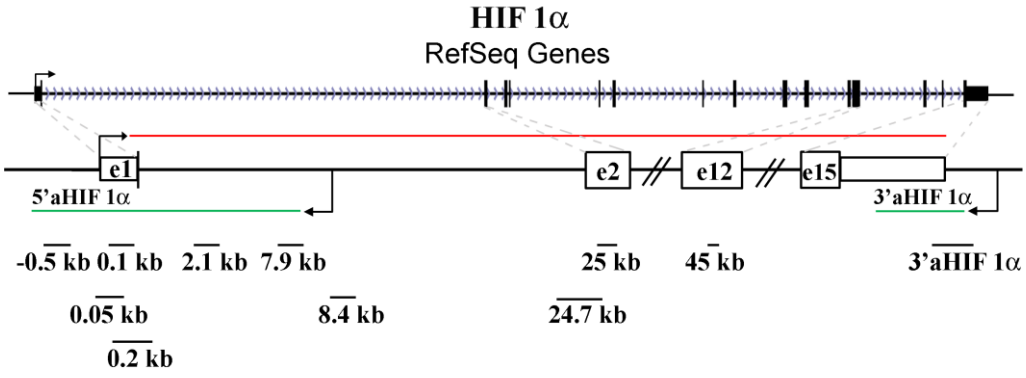

Figure 1. The human HIF-1α gene locus. The HIF-1α gene is indicated with exons (filled boxes) and introns (lines). The lower map is not to scale. Green and red lines are antisense transcripts detected at the 5' and 3' ends and the mRNA, respectively, of the HIF-1α gene. Horizontal lines and numbers indicate the amplicons used in this work, with numbers showing the average distance from the mRNA transcription start.

We performed chromatin immunoprecipitation (ChIP) experiments using two different colorectal cancer cell lines: HCT 116 and HCT 116 Top1siRNA. The latter cells have TOP 1 gene silenced by a siRNA which reduces TOP 1 content by almost 5 fold when compared with control cells (Figure 2). Cells were exposed to 10 μM of CPT for 1 hour and the levels of RNAP II were quantified using quantitative real time PCR (qPCR) on several genes including HIF-1α and GAPDH (Figure 3).

In untreated HCT 116 cells the amount of RNAP II found at the promoter regions of the analyzed genes was markedly larger than that found at regions along the transcribed sequence in agreement with the presence of a pause site at the promoter. After camptothecin treatments, RNAP II levels were found reduced at promoters by 70-80% and increased at sites along the transcribed template (Figure 3a). In HCT 116 Top1siRNA cells, we found that RNAP II likely pauses at the same promoter's site as in HCT 116 cells. After CPT treatments, we detected a lower decrease of 25-30% at promoter sites and we did not observe any increase of RNAP II levels in the body of the gene in cells with silenced TOP 1 gene.

We also verified by ChIP if CPT could influence the binding of the TATA-binding protein (TBP) to the studied gene promoters, however the drug did not affect TBP binding showing a selective effect on RNAP II (Figure 3).

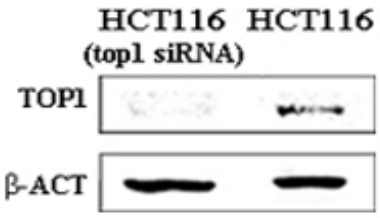

Figure 2. TOP 1 contents in HCT 116 and HCT 116 (Top1siRNA) cells. Total cellular protein extracts were analyzed by SDS gel electrophoresis and Western blots with specific antibodies. β-Actin (β-ACT) is a loading control. TOP 1 in HCT 116 (Top1siRNA) cells was about 20% of the content in control HCT 116 cells.

The negative control for our ChIP experiments was represented by the levels of RNAP II and TBP at the repressed chromatin of the centromeric α-satellite DNA. qPCR analyses did not detect any signal at this regions that is not transcribed in HCT 116 cells (Figure 3 and not shown).

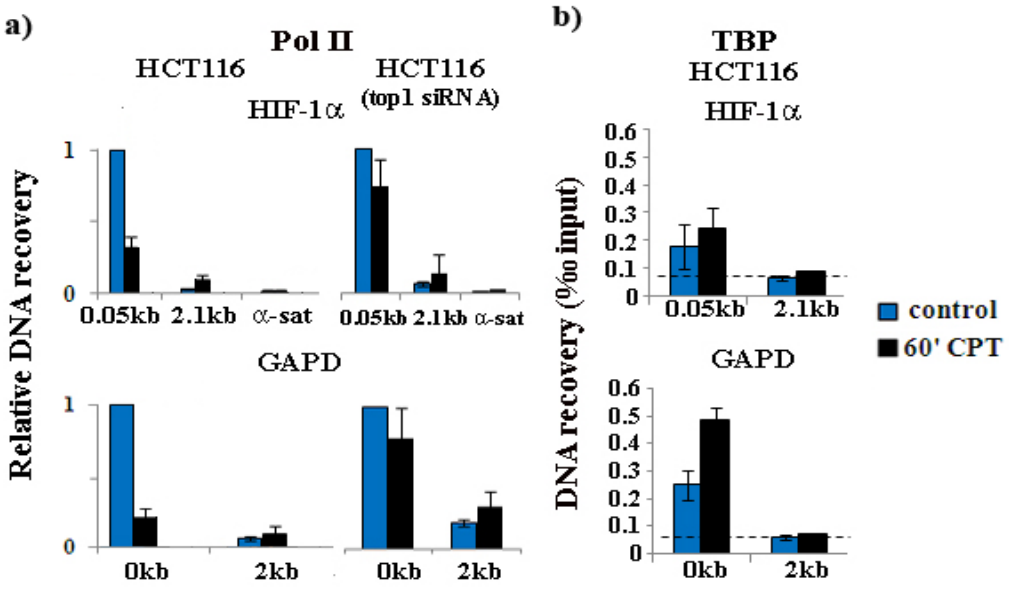

Figure 3. TOP 1 inhibition affects RNAP II escape from the HIF-1 α promoter pausing site. a) HCT 116 cells were treated for 1 hour with CPT (10 μM) and chromatin immunoprecipitation was performed with anti RNAP II Abs. Levels of RNAP II were also determined in HCT116 (Top1siRNA) cells exposed or not to CPT treatment (10 μM, 1 hour). DNA recoveries were determined at promoter-proximal (0.05 or 0 kb) and -distal regions (about 2 kb from mRNA start) of the HIF-1α and GAPD genes. α-sat, negative control region. Mean recovery with non-immune IgG is 0.08. b) Levels of TBP at promoter-proximal regions and at distal regions after CPT treatments. HCT116 cells were treated with 10 μM CPT for 0 and 60 minutes. The dashed line shows the mean DNA recovery with non immune IgG.

Thus, RNAP II likely pauses at the promoter proximal site of HIF-1α and GAPDH gene loci and TOP 1 is required for the specific reduction of polymerase density induced by CPT. Because this was shown for different genes in several cell lines ([35] and not shown) the findings also support the idea of a general effect of TOP 1 inhibition by CPT at transcribed loci.

We further evaluated the specificity of CPT effects on RNAP II distribution by comparing CPT with other DNA-damaging agents: VM26, a t
TOP 2 inhibitor and cisplatin (CisPt) responsible for interstrand crosslinks formation. Both of the two agents promote DNA damage that is independent from TOP 1.

Human MRC5 cells were treated for 1 hour with increasing concentration of each agent (0.4, 2, and 10 μM of CPT; 2, 10, and 50 μM of VM-26 and CisPt). The RNAP II occupancy at promoter regions and at sites around 2 kb downstream from transcription start site was studied through chromatin immunoprecipitation. We compared the relative levels of RNAP II at different gene regions for the three drugs. In case of CPT treatments, with respect to the untreated sample, the data showed that RNAP II density is decreased even with the lowest CPT dose (0.4 μM) at both pausing sites of c-MYC and GAPDH, while it increases at internal regions (Figure 4).

In contrast, cisplatin and VM26 caused a reduction of RNAP II levels both at promoters and along the transcribed gene (Figure 4). This non-specific decrease of RNAP II density occurs only at higher concentration (50 μM not shown). Thus, the interference with pausing regulation is likely specific for TOP 1 inhibition by camptothecin and is a very sensitive effect since it occurs even at relatively low concentrations

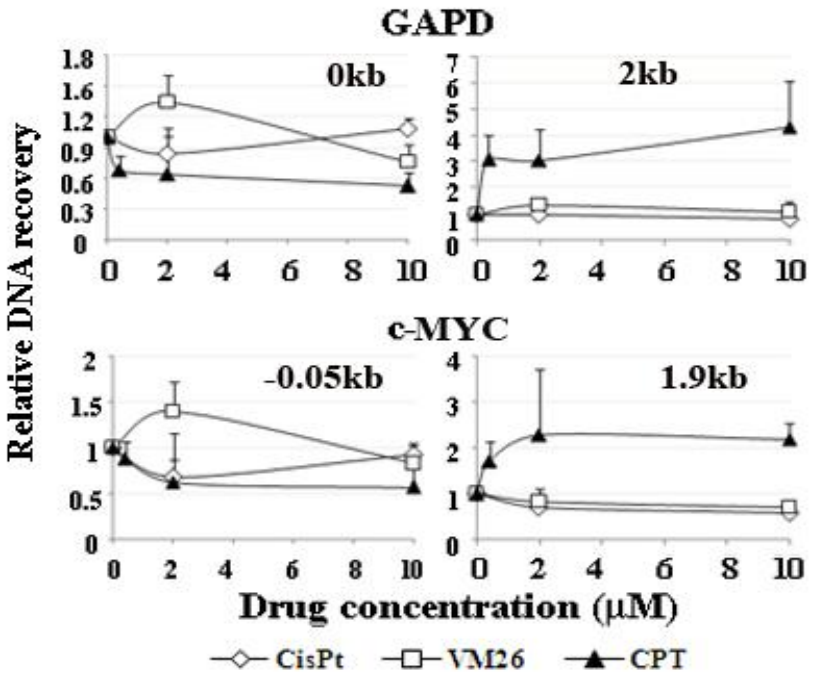

Figure 4. Effects of CPT, teniposide (VM-26) and cis-platin (Cis-Pt) on chromatin-bound RNAP II at the GAPD and c-Myc genes in human MRC5 cells. MRC5 cells were treated with CPT at 0.4, 2, and 10 μM (filled triangles), VM-26 (open squares) and Cis-Pt (open diamond) at 2, and 10 μM. RNAP II levels were determined at the gene promoter (0 kb), and an internal region (2 and 1.9 kb, for GAPD and c-MYC, respectively). Values in all panels are normalized to the recovery of the promoter-proximal region in untreated cells, and are means ± S.D. of at least 4 determinations of 2 independent experiments.

Because the polymerase escape from transcriptional pausing is highly dependent on Cdk7 and Cdk9 activities, we next evaluated whether these kinases were involved in the observed CPT effect on RNAP II accumulation around pausing regions. Treatments of mammalian cells with the nucleotide analog DRB (5, 6-dichloro-1-b-D-ribofuranosylbenzimidazole) inhibit the nuclear Cdk9 and Cdk7 subunits of P-TEFb and TFIIH complexes, respectively. We take advantage of this compound to address the dependence of CPT effects from kinase activity.

HCT 116 cells were treated with CPT (10 μM) or co-treated with CPT and DRB (50 μM) for 1 hour, and levels of RNAP II were evaluated at HIF-1α and GAPDH genes by ChIP and qPCR. The results showed that RNAP II levels at promoter regions were somewhat increased in cells treated only with DRB (Figure 5). By blocking Cdk activity, DRB itself impairs RNAP II escape from promoter pausing site and induces an accumulation of hypophosphorylated polymerase around the 5`

ends of genes [88]. Notably, RNAP II levels registered in DRB-CPT co-treated sample were comparable to those in untreated sample, strongly highlighting that DRB is able to reverse the effect of CPT at the studied genes (Figure 5). This clearly demonstrates that Cdks are required for the altered distribution of RNAP II along the studied genes induced by TOP 1cc formation.

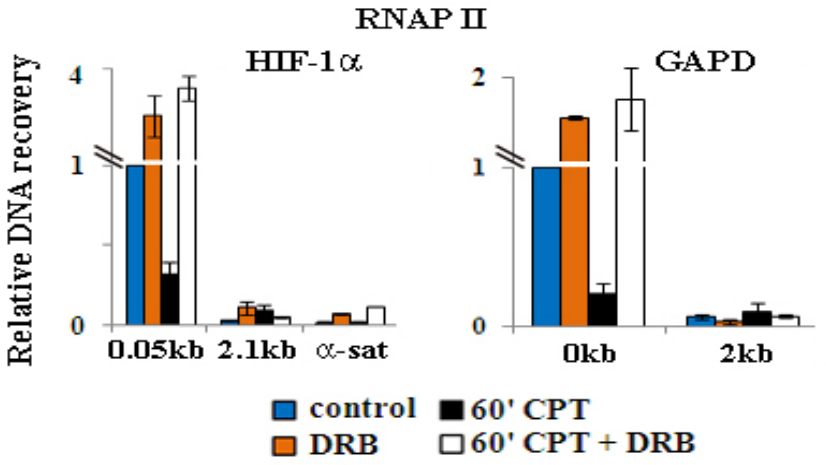

Figure 5. CPT affects RNAP II escape from the HIF-1 α pausing site in a Cdk 7 and Cdk 9 dependent manner. HCT 116 cells were treated for 1 hour with CPT (10 μM) and/or DRB (50 μM). In case of co-treatment DRB was added 15 minutes before CPT, and chromatin immunoprecipitation was performed with anti RNAP II Abs. DNA recoveries were determined at promoter-proximal (0.05 or 0 kb) and -distal regions (about 2 kb from mRNA start) of the HIF-1α and GAPD genes. α-sat, negative control region. Mean recovery with non-immune IgG is 0.08.

3.2 TOP 1cc enhances RNAP II escape from promoter-proximal pausing site.

The previous findings suggest that TOP 1 may be involved in RNAP II pausing regulation and that TOP 1 inhibition likely promotes polymerase escape at the studied genes. It is also well established the importance of promoter-proximal pausing within the transcription cycle as a means of coupling elongation to RNA maturation.

For many years, promoter-proximal pausing was considered a rare phenomenon characterizing only few genes. Recently, genome-wide location analyses of RNAP II have challenged this view by demonstrating that >1000 of 18,000 *Drosophila* genes exhibit hallmarks of RNAP II pausing within their promoter-proximal regions [106, 107]. Moreover, similar ChIP–chip analyses of RNAP II distribution in human cells suggested that regulation of transcription elongation represents a rate-limiting step in the expression of a significant fraction of genes [108].

Another interesting report about promoter pausing was reported by Gilchrist *et al* in 2008 [109]. The authors demonstrated that NELF-dependent RNAP II stalling can both enhance and repress gene expression *in vivo*, suggesting a role for promoter-proximal paused polymerase in transcription activation. It appears that the presence of stalled RNAP II enhances gene expression by maintaining permissive chromatin architecture around the promoter-proximal regions allowing the recruitment of further RNAP II at the promoter sites [109].

Thus, it is critical to establish whether and how TOP 1ccs affect polymerase pausing at promoter proximal sites. If TOP I inhibition by CPT enhances RNAP II escape from these regions, then we should expect increased transcript levels downstream to the pausing site. To test this hypothesis, an experiment approach was designed in which active chromatin is immunopurified and the attached nascent RNA is amplified by qPCR. We named the assay RNA immunoprecipitation (RIP). Antibodies specific for acetylated histone H4 (AcH4) and the N-terminal domain of RNAP II Rpb1 subunit were used to immunoprecipitate cross-linked extracts of HCT 116 cells treated with and without increasing concentration of camptothecin.

The levels of chromatin bound RNAs were detected along HIF-1α gene locus and allowed us to make some general considerations on the studied system.

- In absence of CPT treatment, nascent RNAs were at least 3 logs higher in exon regions than α-satellite regions (Figure 6a) and background levels as established by samples not retro-transcribed (Figure 6b). Differences among exons are likely due to different primer efficiency in qPCR and have already been reported for other genes [110].
- The quantification of nascent transcripts at intronic regions was around 1 log lower than exonic region, showing intermediate levels between exon and background regions (Fig 6).

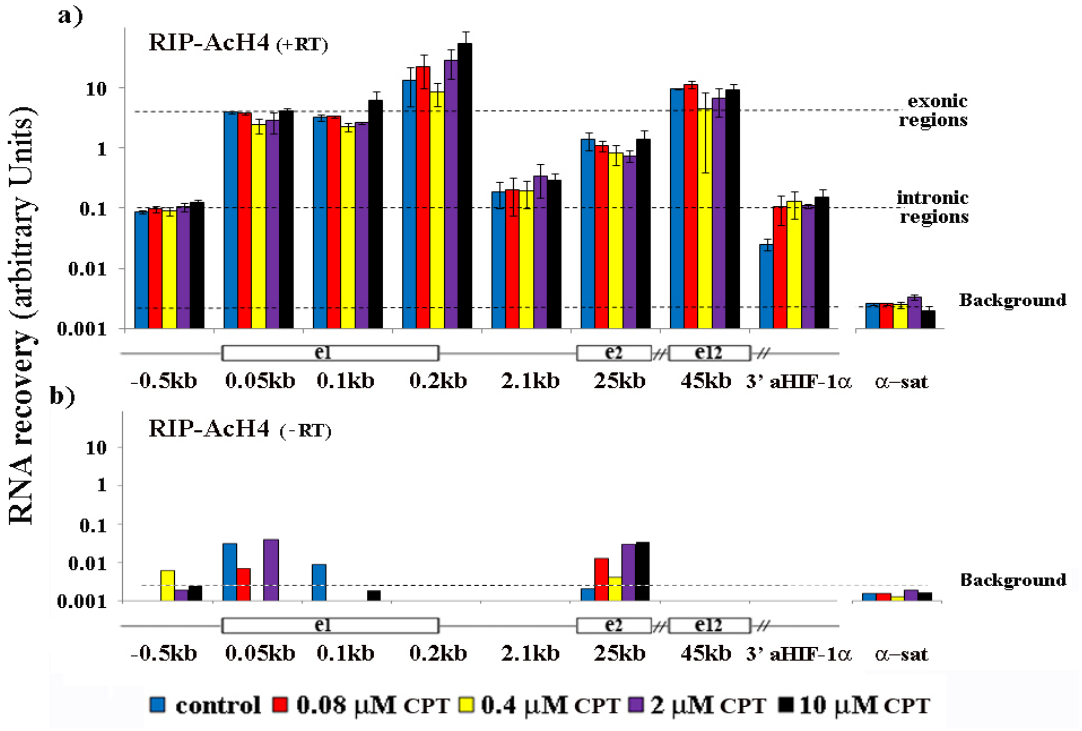

Figure 6. Chromatin-associated RNA at the HIF-1α gene locus. HCT 116 cells were treated for 1 hour at different CPT concentrations (0.08, 0.4, 2, and 10 μM), and indicated by colors of the columns. RIP was performed with Abs against anti acetylated histone H4 (AcH4). Pelleted RNA was retrotranscribed with (+RT/ a) or without (-RT/ b) SuperscriptIII and random primers, and RNA recovery was determined with quantitative real-time PCR using specific primer pairs corresponding to the indicated regions along the HIF-1α gene (see Figure 1). A diagram of the HIF-1α gene (not in scale) is shown under the graph. The graph reports the levels of RNA recovery from a representative experiment. The broken line indicates average recovery of exonic, intronic and centromeric α-sat regions. The background level was determined with cDNAs not retrotranscribed (b) and α-sat recovery levels (a).

The observed differences in RNA recoveries among exon, intron and non transcribed regions may be due to changes in AcH4 distribution along the studied area, thus the density of acetylated histone H4 along HIF-1α locus was analyzed by ChIP experiments using the same CPT treatments performed on cells for RIP experiments. In Figure 7, qPCR results clearly reveal an unchanged distribution of the acetylated histone after cells were exposed to increasing concentration of CPT, and only smaller differences among exonic and intronic regions than those observed with RIP experiments (Figure 6). Thus, the RIP data actually reflect the nascent RNA distribution along the gene rather than an acetylated histone H4 pattern. As intronic RNAs were lower than exonic RNAs, our data are consistent with a co-transcriptional splicing profile of the studied mRNA. This phenomenon has previously been

described by several groups which indicate that pre-mRNA splicing occurs in the context of transcription-unit activity and may be regulated within chromatin by diverse cellular mechanisms [111].

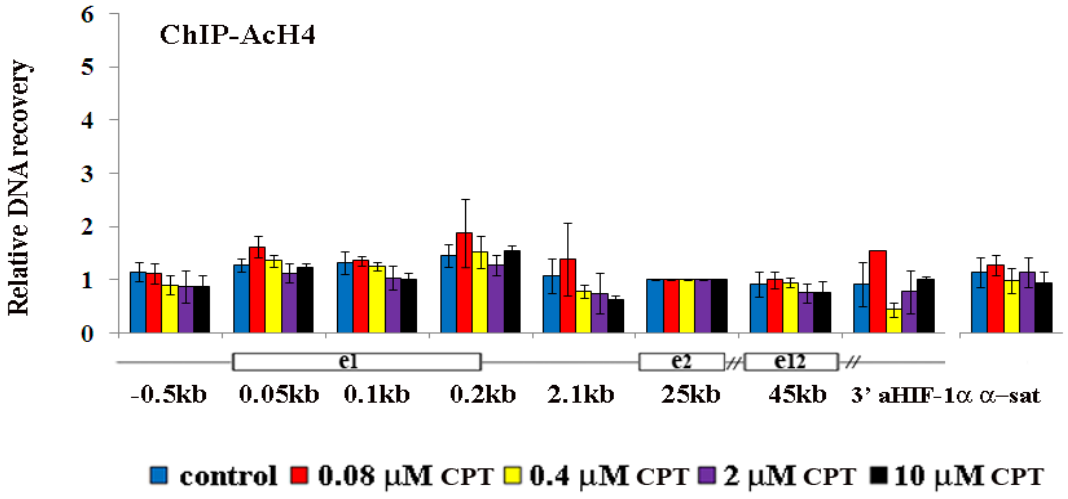

Figure 7. Levels of AcH4 along the HIF-1α gene after CPT treatments. HCT 116 Cells were CPT-treated (0.08, 0.4, 2, and 10 μM) for 1 hour, and ChIP was performed with either anti- acetylated histone H4 (AcH4) or non-immune IgG Abs. Values are normalized to exon 2 and are means ±S.D. of 4 to 6 determinations from two independent experiments.

Notably, upon 1 hour of CPT treatments, we do not register a dramatic change in chromatin-RNA association suggesting that TOP 1ccs are highly reversible in nuclear chromatin and do not block completely transcription of nascent RNAs. This finding is in agreement with the lack of influence of CPT on transcription factories by immunofluorescence microscopy [35]. As previously reported by us, within 1- or 2-hour of incubation, camptothecin did not affect the morphology or intensity of nuclear RNAP II foci, whereas replication factories were totally destroyed by the drug [35]. Unaffected nuclear transcriptional foci could therefore indicate that major destructive collisions do not often occur *in vivo*.

To address whether CPT increases nascent RNAs downstream to promoter proximal pausing site we compared relative levels of transcription at three overlapping regions spanning exon1 and the exon 1/intron 1 border of the HIF-1α gene (Figure 8). The primer sets used for these analyses were designed considering the pausing site usually localized after 20-50 bases from transcription start site [88] and RNAP II density in the region as analyzed by ChIP (see above). RNA extraction from anti-

AcH4 immunoprecipitates was followed by qPCR analyses. The results indicated that CPT increases transcription levels at the 3'-end of exon 1 in a dose-dependent manner, while leaving unchanged levels at the 5'-end (Figure 8).

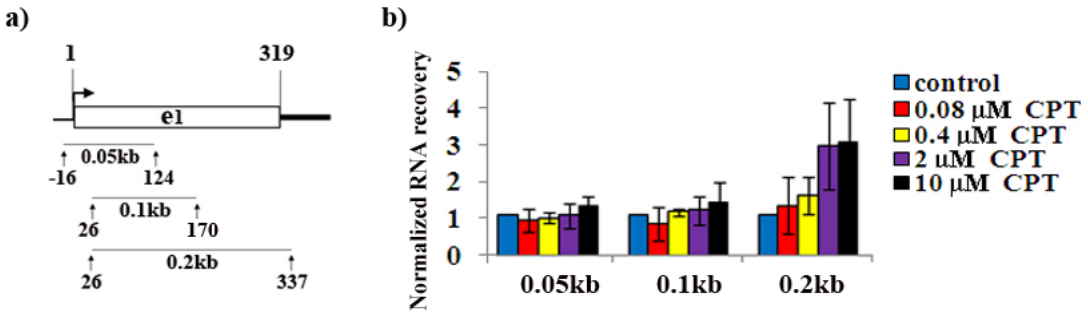

Figure 8. TOP 1 inhibition induces RNAP II escape from HIF-1α promoter pausing site. a) Schematic of HIF-1α exon 1 is shown, with black lines indicating gene regions amplified by primer sets. Numbers indicate positions from mRNA transcription start site. The map is to scale. b) HCT 116 cells were treated for 1 hour with CPT and RIP was performed with anti- acetylated histone H4 (AcH4) Abs. Pelleted RNA was retrotranscribed with random primers and cDNA was analysed with qPCR with the indicated primer sets. Values are normalized to exon 2 (25 kb) and control samples.

By exposing HCT 116 cells to treatments with both CPT and DRB we next assessed if CPT affects nascent RNA transcription in a manner related to Cdk7 and Cdk9 activities. The kinases inhibitor DRB was given to the cells 15 minutes before CPT, and then drug treatments were performed for 1 hour. The graph in Figure 9 shows that 2 and 10 µM of CPT increase the relative transcription levels downstream to promoter pausing in a dose dependent manner while DRB fully abolishes this effect (Figure 9). Consistent with the ChIP data on RNAP II (Figure 5), the results strongly highlight the dependence of CPT effects from Cdk7 and Cdk9 activities suggesting a possible link between TOP 1 and kinases in regulating the transcriptional pausing of the polymerase.

A broad and general inhibition of transcription elongation is an immediate effect of CPT in cultured cells likely due to the stalling of elongating polymerase by TOP 1ccs [34] and to the accumulation of DNA supercoils [112]. Nevertheless, our data provide strong evidence that camptothecin increases transcription downstream to the studied promoter pausing sites, and thus suggest that transcription inhibition by camptothcin may be caused by other, not necessarily alternative, mechanisms. Furthermore given the highly reversible state of TOP 1–camptothecin–DNA

complexes, an encounter between a trapped TOP 1 and an elongating RNAP II could rather be a transient event in chromatin of living cells and less frequent than expected.

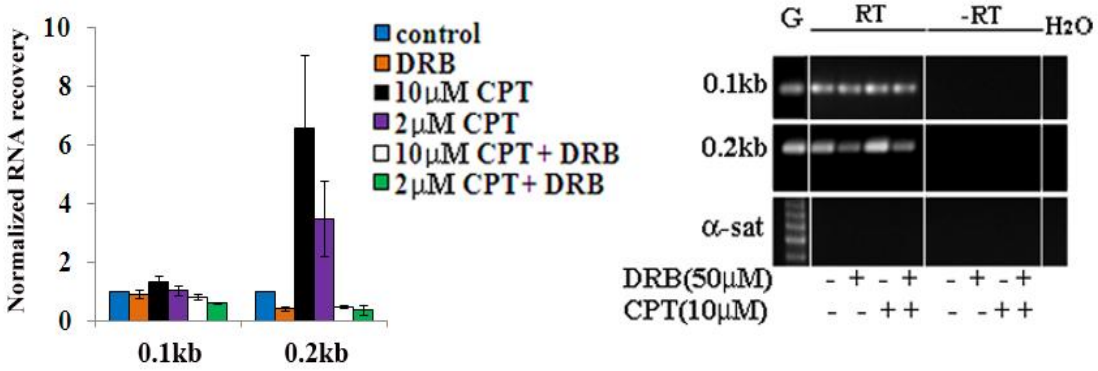

Figure 9. CPT-induced transcription downstream to the pause site is dependent on Cdk activity. Left panel. HCT116 cells were treated for 1 hour with CPT (2 and 10 μM) and/or DRB (50μM). DRB was added 15 minutes before CPT, and RIP was performed with anti acetylated histone H4 (AcH4) antibody. Nascent RNA levels were determined with qPCR at the indicated positions (see also Figure 8a). Right panel. A representative gel showing PCR products using genomic DNA (G) and RNA immunoprecipitated (+ RT) with anti-AcH4 Abs. –RT indicates no reverse transcription. α-sat amplicon is used to check for genomic DNA contamination. Symbols are: G, genomic DNA; RT and –RT show PCR reactions with or without retrotranscriptase enzyme, respectively. α-sat primers show no genomic DNA contamination.

To address if CPT may alter transcription at gene regions further downstream from promoter pausing site, we investigated the matured mRNA of HIF-1α in RIP samples by using two different antibody against acetylated histone H4 and RNAP II Rpb1 subunit. We analyzed transcript levels at a region spanning exon 1 and exon 2 and, as a control, a region spanning exons 10-12 (Figure 10). The qPCR analyses revealed some unexpected findings. First, we found that CPT increases transcription levels of the region spanning exon 1-exon 2 (Figure 10). Because we detected the increase only in chromatin-bound RNAs, but not in total RNA samples (Input RNA), the findings show that CPT affects specifically a subpopulation of RNA, the nascent RNA fraction, and that the drug likely increases transcription downstream to the promoter pausing site.

Second, the results showed that transcript levels were not impaired by CPT treatments at a region further downstream. However, in this case we observed a marked alteration of intron 10 and 11 splicing. Notably, we detected the presence of two additional

bands of about 500 bp and 310 bp in agarose gels, besides the expected spliced product of 442 bp (Figure 10). Interestingly, both the splicing variants were also detected with the lowest CPT concentration (0.08 μM), and the aberrant spliced products seem to increase in a dose-dependent manner (Figure 10b).

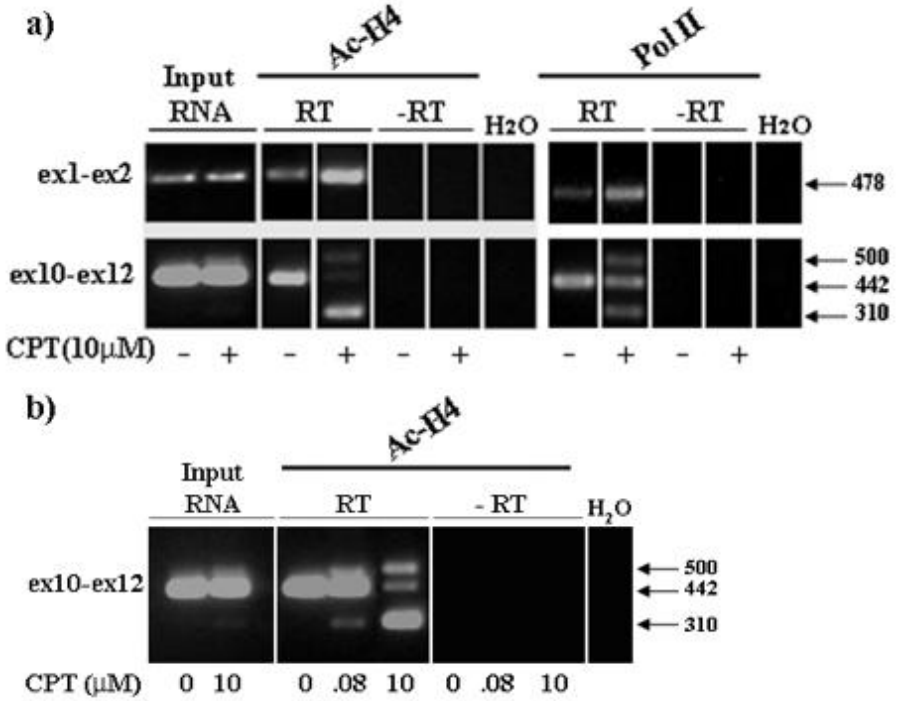

Figure 10. TOP 1 inhibition effects on transcript elongation and alternative splicing of HIF-1α mRNA. a) HCT116 cells were treated with CPT 10 μM for 1 hour and RIP was performed with Abs against AcH4 or RNAP II. Transcription levels downstream to the pausing site of the HIF-1α were measured with primers spanning exons 1 and 2 and exon 10 and 12. b) HCT 116 cells were treated with different CPT concentration (0.08, 10 μM) for 1 hour and RIP was performed with antibody against acetylated histone H4. Splicing of introns 10 and 11 was determined by PCR analysis with specific primers at exon 10 and exon 12. The observed extra band of 310 bases may correspond to a transcript in which exon 11 has been spliced out. The 500-bp band might instead correspond to the inclusion of an intron. Input RNA indicates non immunoprecipitated RNAs. –RT indicates no retrotranscription. On the right of both a) and b), molecular weight in bp of PCR products.

Thus, it appears that CPT alters the splicing of introns 10 and 11 (Figure. 10), as the observed extra band of 310 bases may correspond to a transcript in which exon 11 has been spliced out. Interestingly, exon 11 encodes for the oxygen-dependent degradation (ODD) domain of the protein, localized between amino acid residues 401 and 603 [113]. The ODD domain is required for the ubiquitin/proteasome-dependent degradation of HIF-1α. Under normoxic conditions, HIF-1α interacts with the von

Hippel-Lindau tumour suppressor protein (pVHL) through two hydroxylated proline residues (Pro-402 and Pro-564) of the ODD domain. The ubiquitin E3 ligase pVHL mediates HIF-1α ubiquitylation, and thus its proteasomal proteolysis [100, 114]. Park *et al* showed a new alternatively spliced variant of the human HIF-1α mRNA, which lacked both exons 11 and 12, producing a frame shift and giving a shorter form of HIF-1α in which part of the ODD domain is missing. The expressed HIF-1α variant competed with the endogenous HIF-1α and suppressed HIF-1 activity, resulting in the down-regulation of mRNA expression of hypoxia-inducible genes [114].

To evaluate the generality of CPT effects on RNAP II, we performed the same RIP experiment on the human c-MYC gene locus. We selected this gene according to our previous studies that showed an apparent enhancement of RNAP II escape triggered by CPT from the well-known pausing site near the P2 promoter of the c-MYC gene [35].

RNA immunoprecipitation experiments made with anti-AcH4 and anti-RNAP II showed that 1 hour of CPT exposure increases nascent RNA levels downstream to the P2 promoter leaving unchanged the levels at the 5' end of exon 1 (Figure 11). Furthermore, co-treatments with DRB completely suppress the observed CPT effect as the levels of nascent RNAs remain close to the control recovery (Figure 11c).

The findings unequivocally indicate that CPT-induced escape of RNAP II is not a specific response at the HIF-1α promoter, but it may rather represent a general mechanism of response to CPT interference with transcriptional regulation.

The findings demonstrated that TOP 1ccs can increase transcriptional elongation immediately downstream to pausing sites at least in HIF-1α and c-MYC genes. Similar results were previously reported by Listerman *et al* for other genes. By using the RNA immunoprecipitation approach, the authors showed that FOS pre-mRNA splicing is co-transcriptional and strongly increased by camptothecin treatment [115]. The data were not compared to the levels of RNAP II bound to the promoter. To explain the data Listerman and co-workers concluded that co-transcriptional FOS splicing is favoured by obstruction of RNAP II elongation by camptothecin. These observations provide evidence that the kinetics of transcription by the polymerase can influence the co-transcriptional maturation of mRNAs. Our results, however, support a

different hypothesis, i.e.: we suggest that co-trascriptional splicing is promoted by enhanced RNAP II escape from pausing sites.

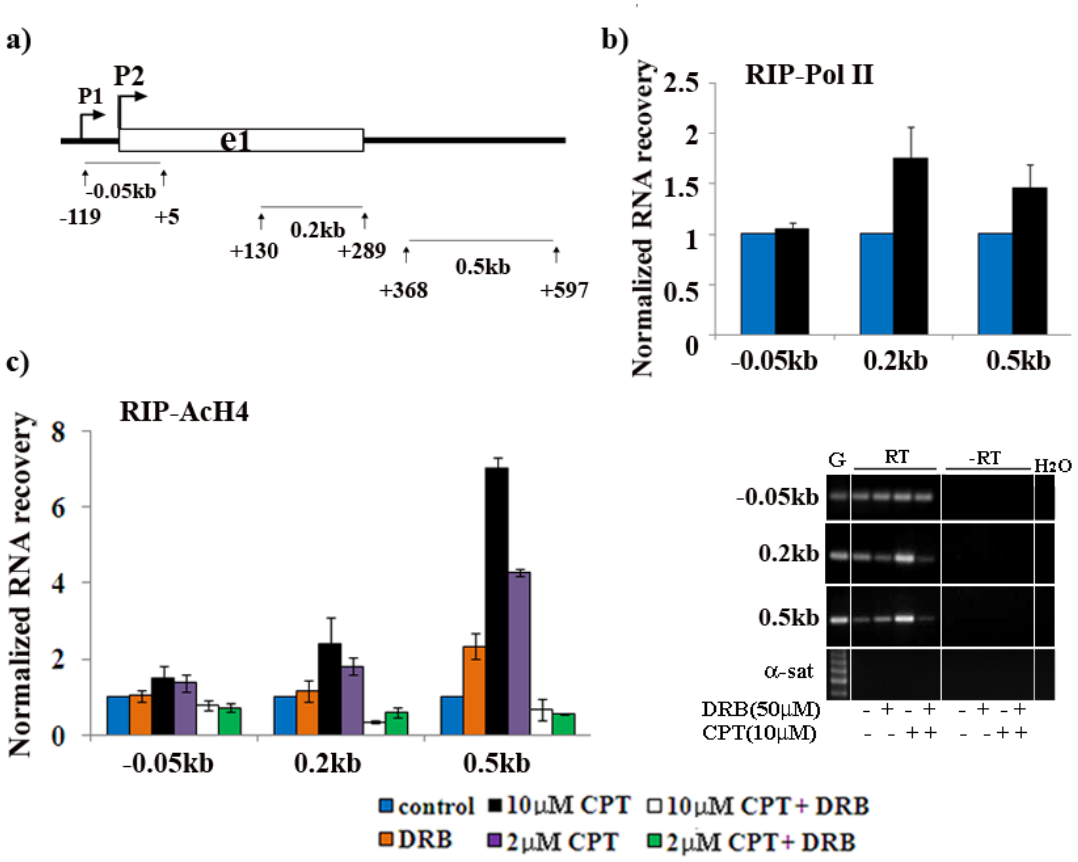

Figure 11. CPT-induced TOP 1ccs promote RNAP II escape from the c-MYC P2 promoter proximal pausing. a) Schematic of c-MYC exon 1 region is shown, with lines indicating the studied amplicons. Positions are relative to mRNA transcription start (in base pairs). b) HCT 116 cells were treated for 1 hour with CPT and RIP was performed with anti-RNAP II c) Cells were exposed to CPT and/or DRB for 1 hours at the indicated concentrations and anti- acetylated histone H4 (AcH4) Abs was used to run RIP protocol. On the right: PCR products of RIP samples. Controls are: genomic DNA (G), no reverse transcription (–RT) and water. α-satellite DNA (α-sat) reactions show no genomic DNA contamination. A representative experiment is reported.

3.3 TOP 1ccs promote transcription of nascent RNAs at both termini and inside the first intron of the HIF-1α gene.

It is known that polymerases do not elongate at constant rates: site specific pauses have been described at splicing and polyadenylation sites, and during early elongation [112]. Using fluorescence recovery after photobleaching at a reporter gene, Darzaq *et al* [112] analyzed the transcriptional process *in vivo*, and demonstrated that a fraction of RNAP II at any moment are paused during elongation for cumulatively

long periods. At steady state, they account for about one-fourth of all polymerase signals [112]. Notably, this pausing is different from the promoter-proximal pausing observed in several genome-wide investigations [12] and in our studies [35] using chromatin immunoprecipitation of RNAP II, as it occurs 1 kb downstream from the promoters. Moreover, RNAP II density has a peak at mRNA termination regions suggesting a pausing site at 3' end of genes [99].

Because our data clearly demonstrate that TOP 1ccs alter promoter proximal pausing, we next aimed at investigating if it may also alter pausing regulation in the body of the gene. To this end, we exposed cells to increasing concentrations of CPT for 1 hour, and quantified levels of nascent RNAs at internal regions and at 5' and 3' adjacent regions of the HIF-1α gene. We first observed that RNA levels upstream to transcription start site and downstream to mRNA termination site were not at background levels (as measured with α satellite DNA recovery) in the untreated cells, but they were more similar to intronic regions (Figure 6a). Surprisingly, CPT treatments increase the nascent transcripts in a dose dependent manner starting from 0.4 μM up to 10 μM (Figure 12).

The transcripts enhancement occurs both when we use antibody against anti-acetylated histone H4 and anti-Rpb1 for RIP experiments. Interestingly, we do not detect any increase around the mRNA termination site of the gene when using Abs against RNAP II. We then noticed that the drug induced a similar enhancement at the 5'-end of the first intron of the HIF-1α gene whereas the intron 3'-end and other internal gene regions remain unchanged (Figure 12c). Furthermore, co-treatments with DRB totally reverse the increase both at the ends of the gene and at intronic regions (Figure 13), demonstrating again the clear dependence of CPT effects from Cdk activities (Figure 13). These observations cannot easily be interpreted as an increase of mRNA transcription due to CPT-promoted alterations of RNAP II pausing at promoter sites. Moreover, the transcript recovery around the 3'-end of HIF-1α was higher than in other regions, and the transcripts were not associated to RNAP II since we did not detect any enrichment in RIP experiments performed with anti-Rpb1 antibody. Therefore, the nascent RNAs in these specific regions of the locus do not likely correspond to the pre-mRNA of the studied gene.

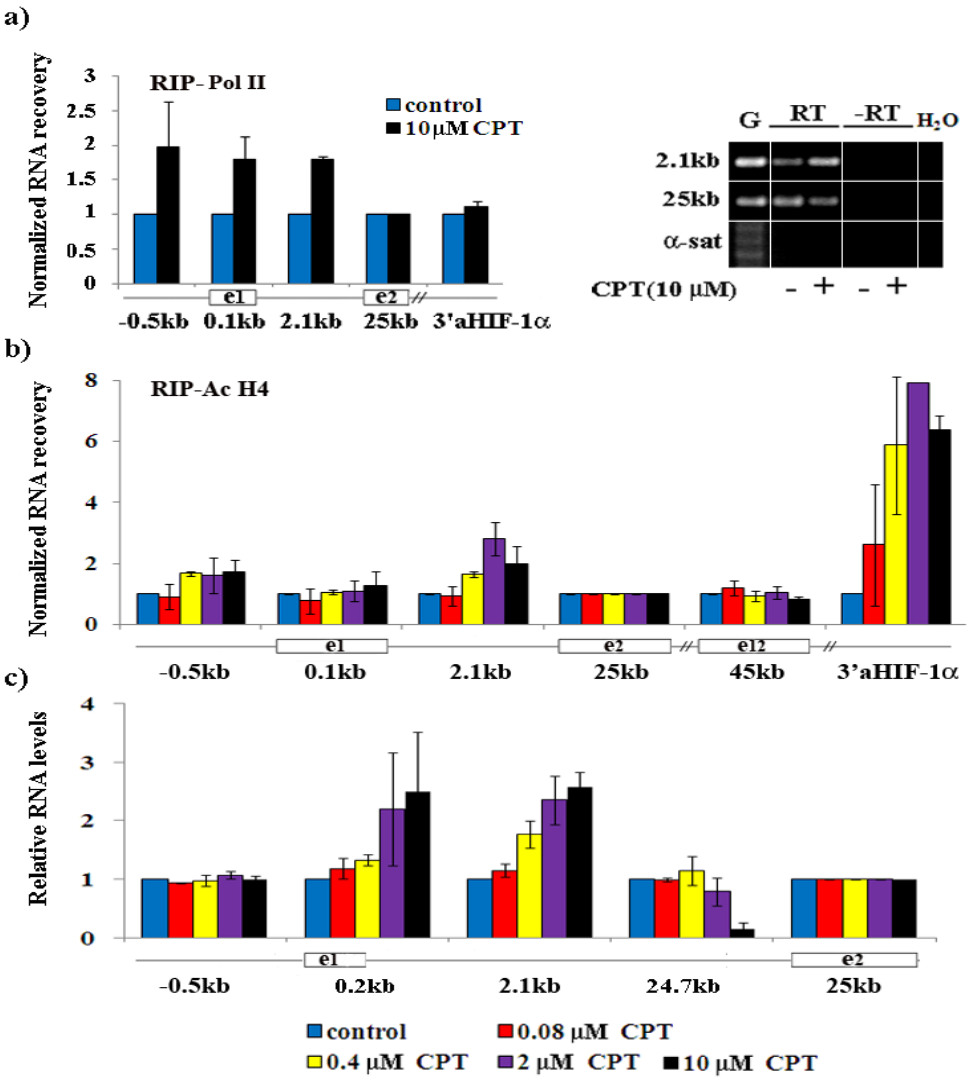

Figure 12. Levels of specific RNAs are increased by TOP 1 inhibition. a) HCT 116 cells were treated for 1 hour with 10 μM of CPT, and RIP was performed with anti-RNAP II Abs. Pelleted RNA was retrotranscribed with random primers and cDNA was analysed with qPCR with the indicated amplicon. Values are normalized to exon 2 (25 kb) and control samples. On the right, a representative gel showing the increased transcript levels after CPT treatments in the 2.1 kb region. b) HCT 116 cells were treated for 1 hour with increasing concentration of CPT as reported. Levels of nascent RNAs were quantified at HIF-1α gene locus by using an anti-acetylated histone H4 Abs (Ac-H4) in the RIP protocol. c) Levels of transcript fragments in total RNAs purified from HCT116 cells treated with the indicated CPT concentrations for 1 hour. Levels were normalized to the recoveries of exon 2 (25 kb) and untreated samples. α-Satellite was always 10^3-fold lower than the lowest determination of HIF-1α regions. In all panels values are means ±S.D. of 4 to 6 determinations from two independent experiments.

A different explanation of our data could be that the polarity of the CPT-increased nascent RNAs was antisense to the mRNA. Even though this interpretation seemed initially strange, however an article published in 1999 reported the discovery

of a natural antisense transcript complementary to the 3' untranslated region of the HIF-1α messenger RNA and overexpressed specifically in nonpapillary kidney cancers [103]. We then hypothesized that the observed enhanced nascent RNAs at the 3' and 5' ends the HIF-1α gene corresponded to antisense transcripts. Thus, the experimental data raised the intriguing possibility that TOP 1ccs may increase the transcription of low abundance antisense RNAs at the human HIF-1α gene locus.

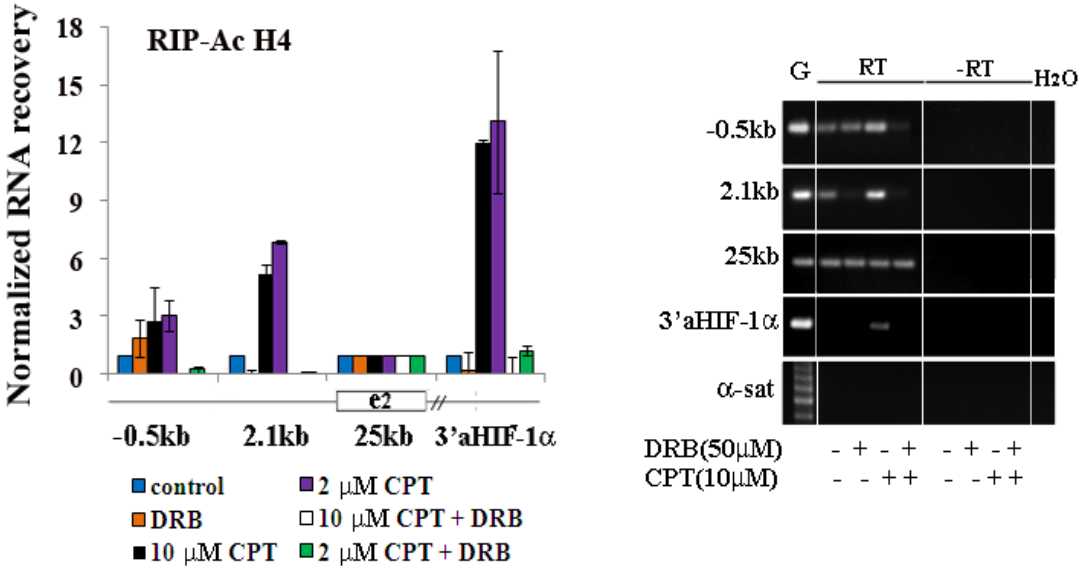

Figure 13. Activation of nascent RNAs is dependent on cdk activity. HCT116 cells were treated for 1 hour with CPT and/or DRB at the indicated concentrations. DRB was added 15 minutes before CPT, and RIP was performed with anti-AcH4 Abs. Nascent RNA levels were determined at different positions along the HIF-1α gene. Values are normalized to exon 2 (25 kb) and control samples. A representative gel showing PCR products using genomic DNA (G) and RNA immunoprecipitated (+ RT) with anti-AcH4 Abs. –RT indicates no reverse transcription. α-sat amplicon is used to check for genomic DNA contamination. Values are means ±S.D. of 4 to 6 determinations from two independent experiments.

3.4 5'aHIF-1α: a novel antisense transcript activated by CPT inhibition of TOP 1.

We first aimed at verifying whether TOP 1ccs may induce the transcription of the 3'aHIF-1α antisense RNA. HCT 116 cells exposed to CPT for 4 hours were subjected to total RNA extraction and PCR analyses with specific primers for exon 2 and 3'aHIF-1α. The latter primers were designed in a region spanning 3`UTR and untranscribed regions of HIF-1α locus in order to catch only antisense transcript [116]. The cDNA retrotranscription was performed with strand specific primers in order to

get only the antisense transcript. Strikingly, we detected the fragment corresponding to the antisense RNA only in CPT-treated sample while exon 2 levels were reduced upon CPT treatments (Figure 14a).

Thus, we effectively proved that TOP 1ccs can increase the level of the natural antisense localized at the 3'-end of the studied gene. These findings prompted us to investigate whether CPT could also activate antisense transcription around the 5'-end of HIF-1α where we found increased nascent RNAs upon CPT treatments.

To this end, we developed a precise strategy consisting of: 1) cell treatments with CPT for 4 hours, 2) total RNA extraction, 3) cDNA retrotranscription with strand specific primers spanning the first part of HIF-1α gene (horizontal line of Figure 14b), 4) PCR analyses with primer sets spanning the entire HIF-1α gene (vertical line of Figure 14b). As during retrotranscription reverse primers target the sense strand and forward primers target the antisense strand, this strategy allows to study both sense and antisense transcripts separately. PCR analyses on total RNA retrotranscribed with forward primers surprisingly reveal the presence of antisense transcripts in samples from cells treated with 10 μM of CPT for 4 hours. This is also confirmed by PCRs on RNAs retrotranscribed with random primers since amplicons increased in antisense cDNAs were also enhanced in this cDNA population. According to PCR data, the antisense transcripts should span from an intronic region at about 8 kb from mRNA start site to a region more than 0.5 kb upstream to the mRNA start (Figure 14b). PCR analyses in the second part of the first intron did not detect any antisense transcription.

Sense transcript analyses revealed that pre-mRNA detected at exon 1 and intron 1 regions (amplicons 0.1 and 2.1 kb, respectively) were unchanged after 4 hours of CPT. Finally, we did not detect any PCR product at 0.5 kb upstream to transcription start site. As expected after 4 hours of CPT, we observed a reduction in exon 2 levels in total cDNA retrotranscribed with random primers (Figure 14b).

This strategy led us to demonstrate the activation of antisense transcription at the 5' and 3' ends of the HIF-1α locus upon CPT treatments.

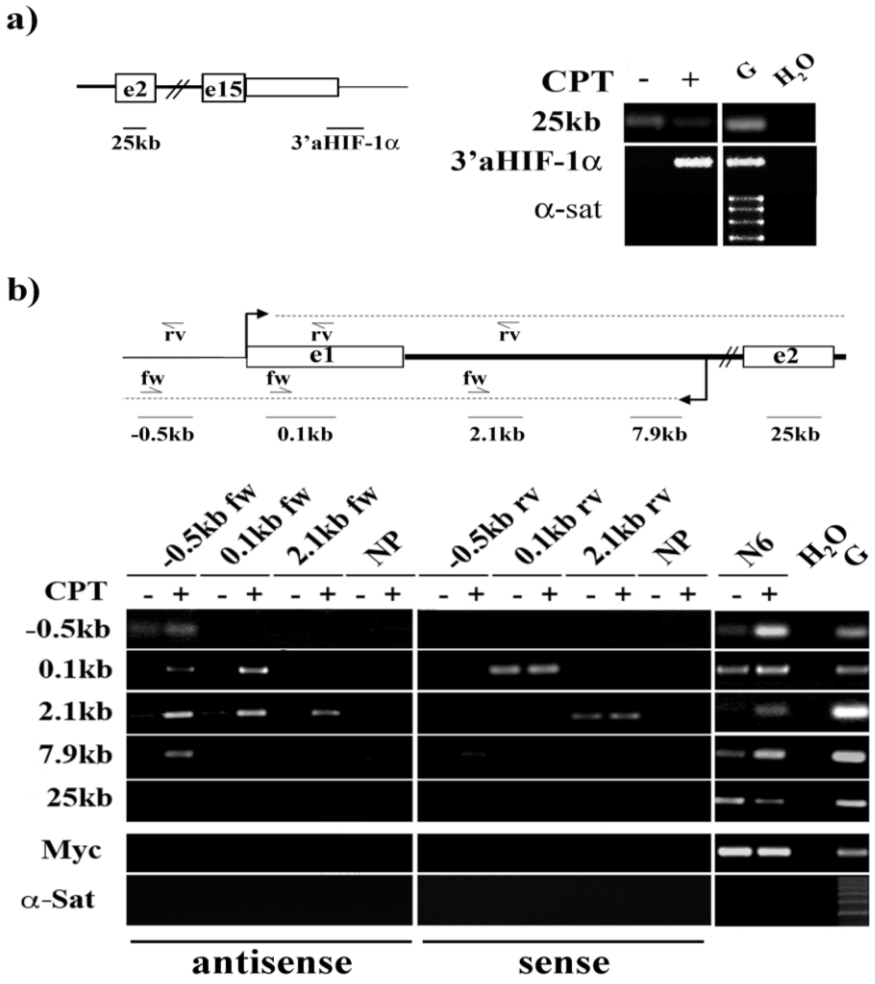

Figure 14. Antisense transcript at the HIF-1α gene locus is increased by TOP 1cc. a) Total RNAs were extracted from control (-) or 10 μM CPT-treated (+) cells for 4 hours. PCR analyses were performed with primer indicated in the left map. b) Upper panel. Schematic of HIF-1αgene region analized by PCR analysis. Dashed lines are antisense transcript detected at the 5' and the mRNA of the HIF-1α gene. Arrows indicate specific primers used for retrotranscription. Black lines indicate gene regions amplified by primer sets. Numbers indicate positions from mRNA transcription start. The map is not to scale. Lower panel. PCR analysis on strand-specific retrotranscripts from HCT116 cells. Rv primers were designed to target the HIF-1α mRNA, while fw primers targeted antisense transcripts. Total RNAs were extracted from cells untreated (-) or treated with 10 μM CPT for 4 hours (+). Specific primers used for retrotranscription are indicated on top of the gel, and amplicons used for PCR analyses are indicated on the left of the gel. Negative and positive controls of retrotranscription were no primers (NP) and random primers (N6), respectively. Negative and positive controls of PCR were water (H_2O) and genomic DNA (G), respectively. α-satellite DNA (α-sat) was used to check for genomic DNA contamination.

To better investigate these novel antisense RNAs, we performed Northern blot analyses. Total RNA extracted from HCT 116 cells treated with or without CPT for 1, 2 and 4 hours was hybridized with a strand specific oligoprobe corresponding to the

2.1 kb amplicon of HIF-1α intron 1 (Figure 15a). The probe was designed complementary to the antisense strand so that it can hybridize only with antisense transcripts. We detected a band corresponding to an RNA (5'aHIF-1α) approximately 12 kb long that increased over 4 hours of CPT treatments. Interestingly, when we repeated the same experiment on HCT 116 Top1siRNA cells, which expressed reduced TOP 1 levels, we did not see any significant increase of the antisense RNA. The weak signal detected and corresponding to the 5'aHIF-1α could rather be due to low TOP 1 levels still expressed in these cells (Figure 15a). No transcript was detected using a strand specific oligoprobe complementary to the same intronic region of sense RNAs (not shown). Thus, the data strongly demonstrate the presence of a non-coding antisense RNA, 12 kb long, complementary to the first region of HIF 1α mRNA, the induction of which is dependent from TOP 1cc formation.

We also assessed the presence of the novel 5' antisense transcript both in normoxic and in hypoxic conditions and in a different human cell line, the leukemia Jurkat cells. qPCR analyses were made with total RNA extracted by CPT treated Jurkat cells using primers corresponding to exon 2 and amplicon 2.1 kb of intron 1 (Figure 15c). Since intron 1 mRNA is usually barely detectable in total RNA samples because it is mostly spliced out, its detection in RNA samples indicates the presence of 5'aHIF-1α transcript. In the CPT untreated sample, we detected poor levels of the 2.1 kb amplicon that likely correspond to pre-mRNA subpopulation or spliced out introns.

By comparing relative levels of the 2.1kb fragment with respect to exon 2 levels, we registered a time dependent increasing in the antisense transcript after CPT treatments. Moreover, since the relative levels of the 2.1 kb amplicon are higher during hypoxia than in normoxia, we conclude that CPT has a more pronounced effect on the induction of the antisense under hypoxic conditions (Figure 15c).

Finally, by performing a Northern blot analysis using a probe complementary to exon 2, we investigated the CPT effect on mRNA transcription. CPT treatments were for 1 to 4 hours on Jurkat cells in both normoxic and hypoxic conditions. Figure 15b shows the detection of a transcript approximately 4 kb long, which corresponds to the HIF-1α mRNA, that is reduced by CPT in a time dependent fashion. Thus, CPT impairs HIF-1α sense transcription in agreement with established knowledge [34], but

strikingly, it has also an unexpected effect, the enhancement of antisense transcription at both 5' and 3' sites of the studied gene locus.

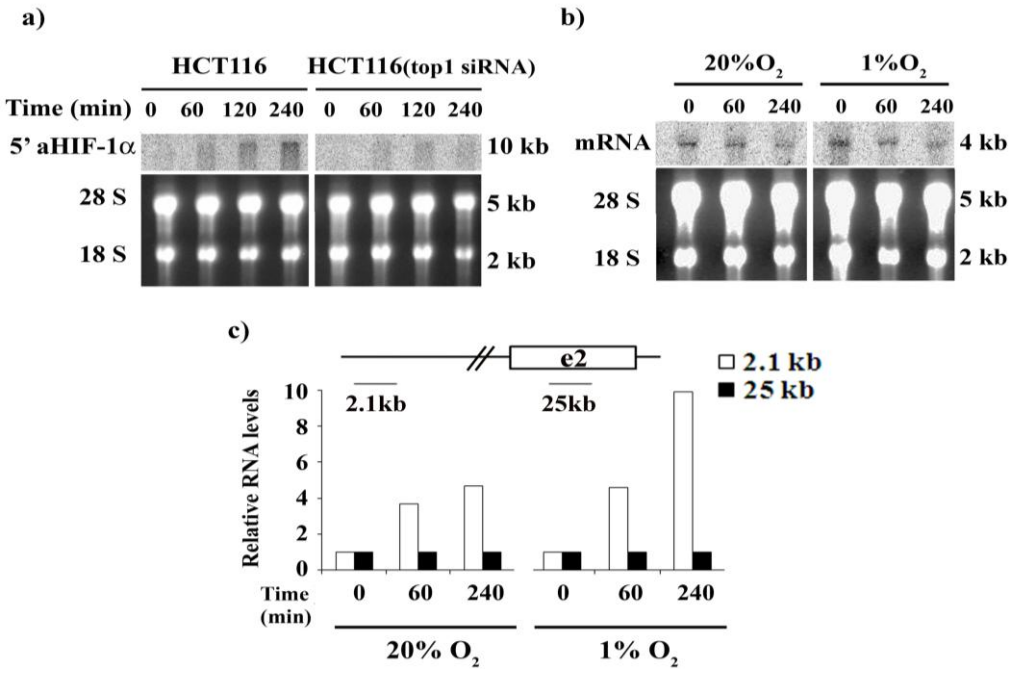

Figure 15. Antisense transcription was induced by CPT treatment and dependent on TOP 1 presence. a) Levels of 5'aHIF-1α antisense RNA in cells treated with 10 μM CPT for 1, 2 and 4 hours. Northern hybridization was with a 40-nt oligomer corresponding to the 2.1 kb amplicon and complementary to antisense RNA. Agarose staining of rRNAs was used to check for equal loading and RNA integrity. HCT116 top1siRNA cells had TOP 1 silenced with siRNA. b) Levels of HIF-1α mRNA in Jurkat cells treated with 10 μM CPT for 1 and 4 hours. Northern blot of total RNAs was carried out with a probe corresponding to exon 2. Agarose staining was used to check for equal loading and RNA integrity. c) Levels of transcribed fragments in total RNAs purified from Jurkat cells treated with 10 μM CPT for 1 and 4 hours in normoxia (20%) or hypoxia (1%). Fragment levels were normalized to exon 2 (25 kb) and untreated samples. DNA contamination (as established with α-sat) were 10^3-fold lower than lowest determinations of HIF-1α regions. A representative experiment is shown.

TOP 1ccs trapped by camptothecin generates replication-mediated DNA double-strand breaks. Accordingly, blockage of DNA synthesis can prevent the cytotoxicity of camptothecin and the activation of the classical DNA-damaging responses [34].

To investigate whether CPT-activated antisense transcription was caused by replication-mediated DNA damage, we performed experiments using the DNA replication inhibitor aphidicolin, and the ATM (ataxia-telangiectasia mutated) and ATR (ATM-RAD3 related) checkpoint kinases inhibitor caffeine. Both these

compounds were added to HCT 116 cells with or without CPT for two hours. Levels of exon 2 (amplicon 25 kb) and intron 1 (amplicon 2.1 kb) were quantified through qPCR after total RNA extraction.

As already indicated above the 2.1 kb amplicon is monstly spliced out in total RNA samples, thus its detection likely indicates the presence of the 5'aHIF-1α transcript.

Figure 16a shows that CPT activates antisense transcription in the absence as well as in the presence of aphidicolin and caffeine. We also checked in parallel the cellular contents of early hallmarks of DNA double-stranded breaks as a consequence of drug treatments [117]. By western blot analyses, we compared the cellular contents of the histone variant γ-H2AX and phosphorylated p53 in HCT 116 cells after treatments with CPT, and with or without aphidicolin or caffeine (Figure 16b). We detected increased levels of γ-H2AX and p53 in the presence of CPT alone and the increases were fully abolished by both Caffeine and aphidicolin. The findings indicate that double strand break formation is associated with CPT action, but also that replicating DNA breaks do not occur in the presence of aphidicolin. Thus, because we are able to detect the increase of antisense transcription also in cells co-treatment with aphidicoline or caffeine, we infer that the increase is independent from replication and/or DNA breaks at replication forks, and is not correlated to the activation of checkpoint pathways.

Finally, we assessed the TOP 1 dependence of antisense activation by performing the same experiment on HCT 116 top1siRNA (Figure 16a). The data showed that, in the absence of TOP 1, CPT does not alter the level of antisense transcripts and co-treatment with aphidicolin does not change this situation in comparison to untreated sample. The results confirm that the CPT activation of antisense transcription requires the presence of TOP 1 in the cells.

3.5 TOP 1 inhibition by CPT induces a more accessible chromatin structure.

Non-coding RNA molecules of various sizes appear to play a broad role in the regulation of chromosome behavior. For example, it has been shown that p15 antisense RNA induces p15 gene silencing through heterochromatin formation [118]. Studies in the fission yeast *Schizosaccharomyces pombe* also reveal that histone deacetylases, a specific class of enzymes that negatively regulate DNA accessibility to

transcription factors, are able to affect transcription by repressing antisense activation from cryptic start sites [119].

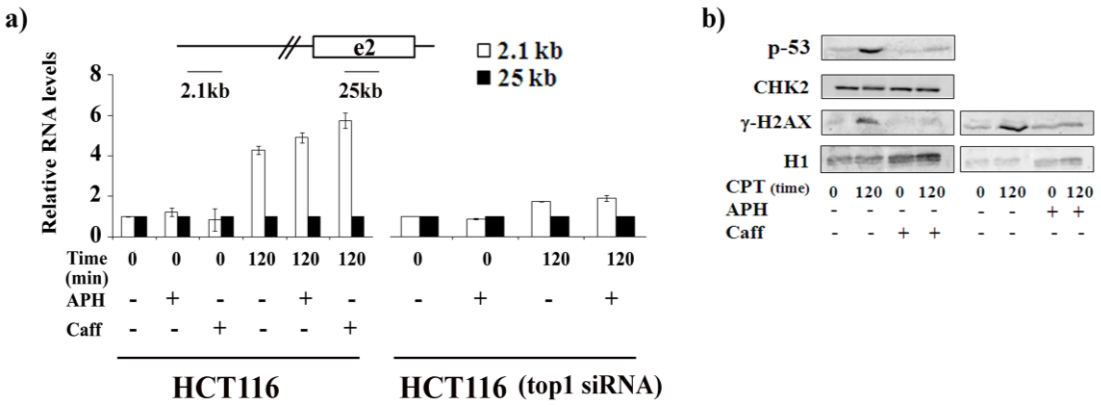

Figure 16. Activation of antisense transcription was independent from replicative DNA damage and/or DNA damage checkpoint kinases. a) Transcription levels of the indicated HIF-1α regions were quantified by real-time PCR and normalized to the levels of exon 2 (25 kb) and untreated samples. Total RNA was extracted from untreated cells (0), and cells treated with 10 μM CPT for 2 hours with or without 5 μM aphidicoline (APH) or 5 mM caffeine (Caff). Left and right panels show the findings in the HCT116 and HCT116 (top1siRNA) cells, respectively. All values are means ±S.D. of 4 determinations from 2 independent experiments. b) HCT 116 cells were treated with CPT (10 μM) or co-treated with CPT and Caff or APH for 120 minutes as indicated. Total cell proteins (upper two panels) or histones (lower panels) were then separated with SDS gel electrophoresis. The indicated proteins were then revealed by Western blots with specific antibodies. CHK2 and H1 histone were used as loading controls. CHK2 phosphorylation (a light smear above the band) was detected in the CPT-treated cells only.

As previous data from our lab showed that TOP 1 gene deletion affects telomere-proximal gene expression and histone modifications in *S. cerevisiae* [120], we asked the question of whether CPT could also affect histone modifications at the human HIF-1α gene locus. To this purpose, we performed chromatin immunoprecipitation studies in HCT 116 cells treated with CPT, using antibodies directed against different histone modification markers. We analyzed histone H4 acetylation at lysines 5, 8, 12 and 16, and histone H3 acetylation at lysines 9 and 14, which correlate with more accessible chromatin structures, and histone H3 dimethylation of lysine 9, which is indicative of repressed chromatin. We normalized the DNA recovery values of each histone marker with that of total histone H4 that is correlated to the nucleosome density at the studied DNA fragment.

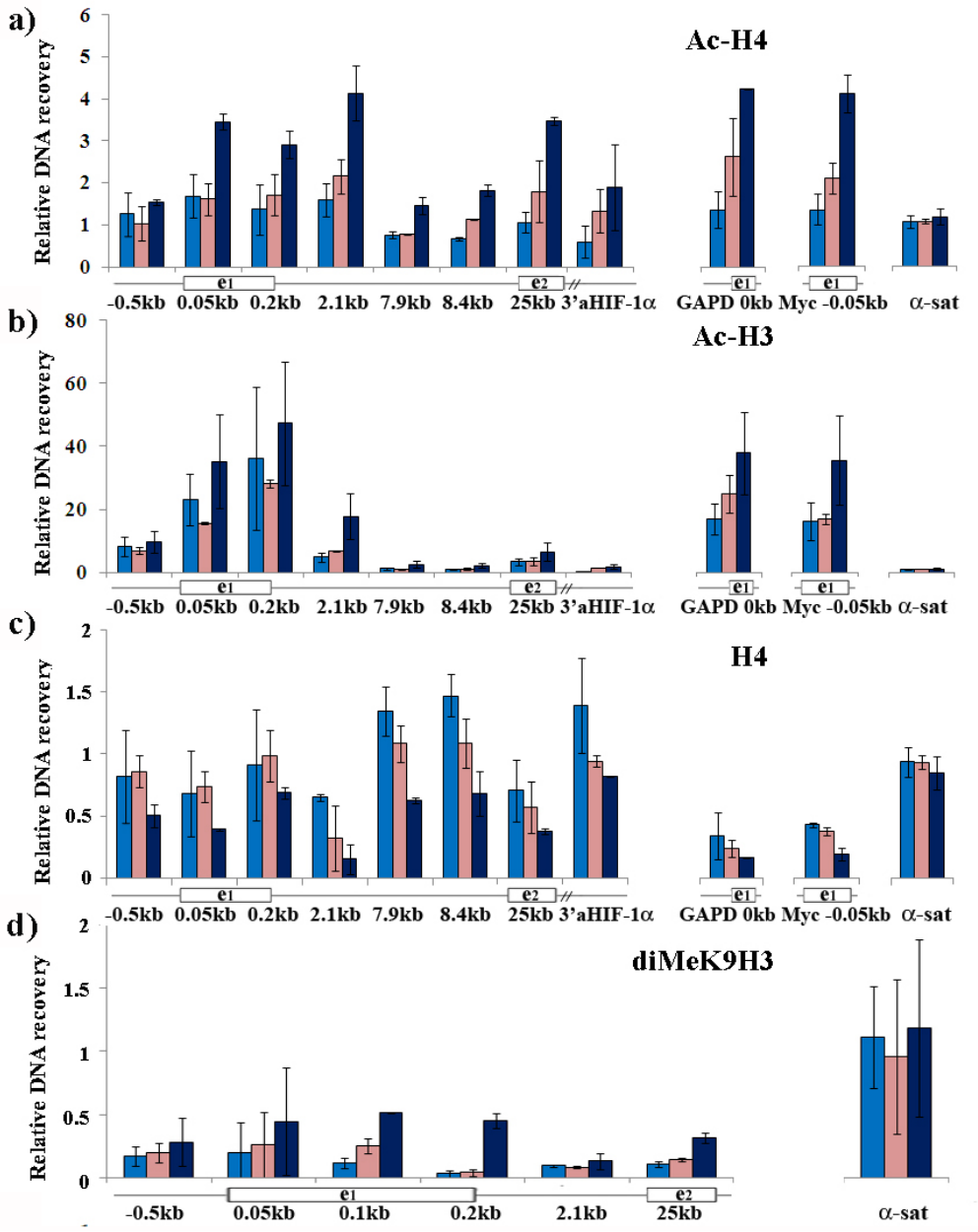

Figure 17. Late effects of CPT on chromatin structure. HCT116 were treated with 10 μM CPT for 0, 120 and 240 minutes. Chromatin was collected and immunoprecipitated with Abs against acetyl -K5, -K8, -K12 and K16 of histone H4 (a) or acetyl -K9 and -K14 of histone H3 (b). Histone acetylation levels were determined along the HIF-1α gene, at GAPD promoter (GAPD 0kb) and at c-MYC P2 promoter (Myc 0kb). Mean DNA recovery with non-immune IgG is 0.41. Values are means ±S.D. of 4 to 8 determinations from 2 to 4 independent experiments. c) Chromatin immunoprecipitation was performed using Abs against total histone H4 in cells treated with CPT for different time. d) Levels of dimetil -K9 of histone H3 after CPT treatments. A representative ChIP experiment is shown, and values are means ±S.D. of two to three determinations. Mean DNA recovery with non-immune IgG is 0.59. In panels a), b) and c) the values are normalized to the recoveries of α-sat, as internal standard, and to that of total histone H4 (shown in c).

The results thus showed that CPT promotes a more accessible chromatin structure at transcriptionally active regions while leaving acetylation unchanged at non transcribed regions (Figure 17c).

Finally, we assessed if CPT effects on chromatin structure was related to Cdk activity by co-treating cells with DRB and CPT for 4 hours and by analyzing acetylation levels of histone H3 and H4. Figure 18 shows that DRB totally suppresses the CPT increase of acetylation levels of histone H3 and H4 in active loci. Thus, CPT affects histone acetylation in a Cdk-dependent manner.

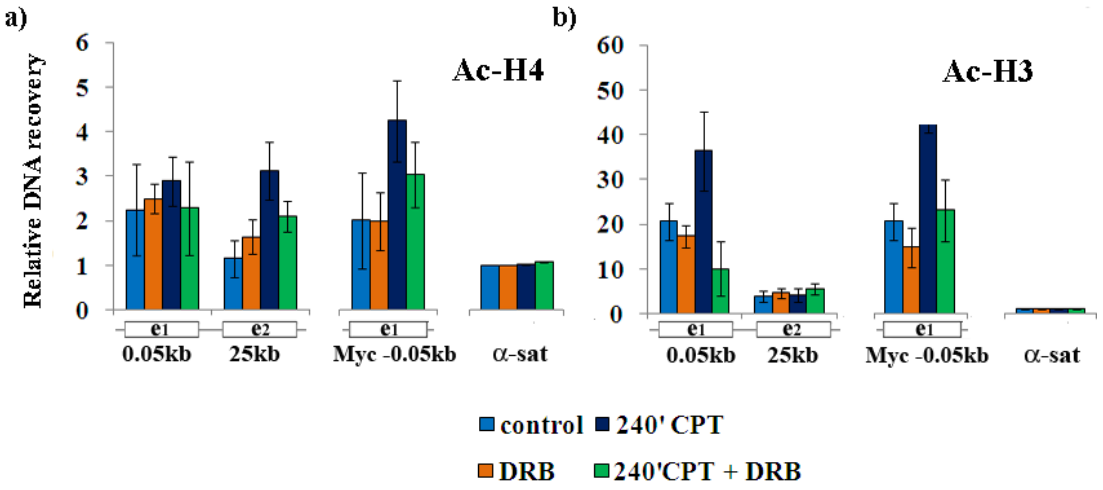

Figure 18. DRB suppress CPT effect at chromatin level. HCT116 cells were treated with 10 μM CPT and/or DRB 50 μM for 4 hours. In case of co-treatment DRB was added 15 minutes before CPT. Chromatin immunoprecipitation with anti-acetylated histone H4 (a) or H3 (b) abs was performed followed by qPCR analysis. Values are means ±S.D. of four to six determinations of two independent experiments. Values are normalized to the recoveries of α-sat, as internal standard, and to that of total histone H4 (shown in Figure 17c).

3.6 "Topo-Seq": a new strategy to map TOP 1 cleavage sites across the genome in human cell lines.

The transcriptional consequences of TOP 1cc stabilized by camptothecin at relatively low concentrations are wider than expected. Overall, our findings show new unexpected roles for TOP 1 during transcription since its inhibition affects transcript elongation, splicing, chromatin structure and activation/de-repression of antisense RNAs. We still ignore if TOP 1 catalytic activity is involved in these functions, but we

certainly demonstrated the independence of these effects from replication-generated double-strand breaks and from checkpoint pathway activation.

Since the evidence for many TOP 1 functions are fragmentary or indirect, we decided to develop a strategy to get a general picture of the protein's role in the cellular environment. Thus, we aimed at defining TOP 1 binding site across the genome in a human cell line to build a solid ground to the understanding of the enzyme roles in each specific context. In addition, to gain information not only about TOP 1 localization but also about its catalytic activity, we planned a strategy to map the genomic distribution of TOP 1 cleavage sites. This will reveal the sites at which TOP 1 is bound and exerts its nicking-rejoining activity.

The process to develop "Topo-seq" strategy basically consisted of: first, an *in vitro* phase which helped to validate the main steps of the protocol and, second, *in vivo* cell phase that allows testing the accuracy of the strategy with endogenous topoisomerases 1. Finally, after verifying with the appropriate controls the success of the protocol, we moved to the most important step that consists of sequencing and analyzing the DNA fragments associated with TOP 1 cleavage sites.

3.7 *In vitro* nicking-labeling strategy.

The mapping strategy takes advantage of the TOP 1 catalytic cycle. Since the enzyme has to introduce a nick into the double helix to exert its activity, we decided to use the nick translation method to map the TOP 1 associated nicks. In order to optimize this technique, we set up an *in vitro* assay using a plasmid DNA (pUC19) and a nicking endonuclease that cleaves only one strand of DNA on a double strand DNA substrate (Nt-BbvCI). We choose this enzyme in order to produce a single nick into the plasmid. This was then labeled by nick-translation using digoxigenated nucleotides (DIG-dUTP). To avoid the generation of a long labeled fragment, which may lead to loss of resolution in the mapping, we used increasing concentration of dideoxynucleosides to inhibit excessive chain elongation by DNA polymerase. After nick-translation the plasmid was digested with a restriction enzyme which gave three different fragments (the 896-bp fragment spans the nick site), immunoprecipitated with anti-digoxigenin antibody and subjected to gel electrophoresis (Figure 19). Appropriate controls of the nick translation and of the immunoprecipitation step were a non-labeled sample (NL) and a non-immune IgG immunoprecipitated sample,

respectively. Figure 19 shows that in comparison to the total DNA sample (Input), the anti-digoxigenin immunoprecipitates are enriched in the 896 bp bands, whereas the unbound DNA, which was recovered from the supernatant of immunoprecipitated sample, was depleted of that specific band. No bands were observed after the immunoprecipitation in non-labeled sample and in samples immunoprecipitated with non-specific antibody (IgG). We are not able to detect a dose dependent effect of dideoxynucleosides. Thus, the gel clearly shows the success of this technique both for labeling and recovering nicked fragments. Since during the catalytic cycle, TOP 1 is covalently linked to the 3` end of the nicked strand, we optimize the conditions of TDP1 treatments to remove TOP 1 molecules covalently linked to DNA, and polynucleotide kinase treatments to reconstitute a free 3'-OH and a 5'-phosphate needed for nick translation (not shown).

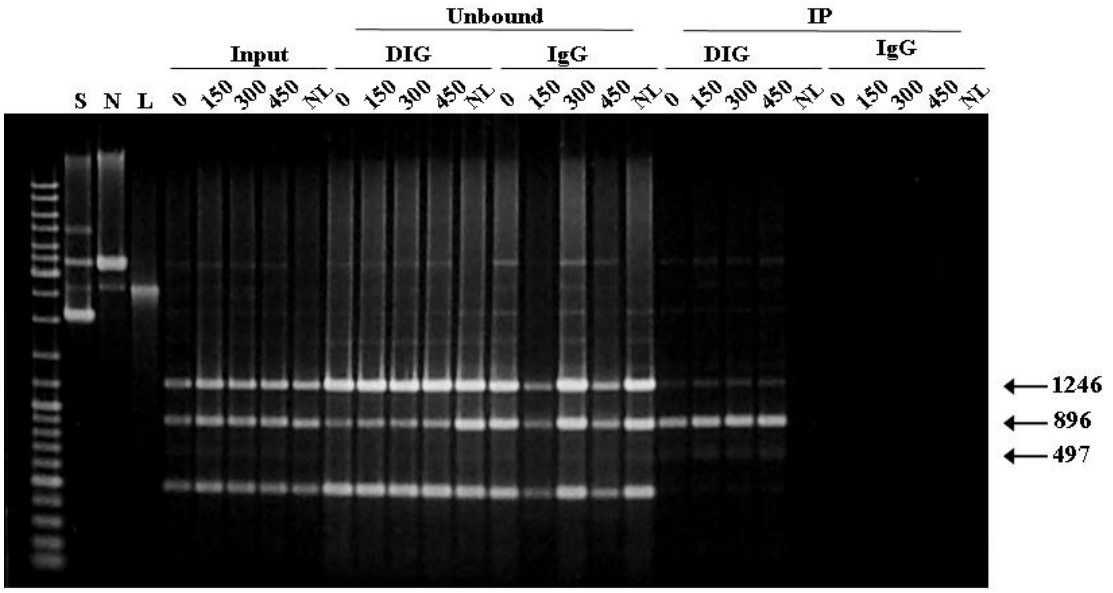

Figure 19. In vitro nicking-labeling strategy. The pUC 19 plasmid was digested with Nt-BbvCI in order to generate a single nick. The plasmid was subjected to nick translation using DIG-11-dUTP and increasing concentrations of dideoxynucleosides (0, 150, 300 and 450 μM) and then digested with Apa I which generates 3 bands: 896bp containing the nick site, 497bp and 1246bp. Immunoprecipitation was performed using antibody against digoxigenin and non immune IgG. Negative control of nick-translation was non-digioxigenine labeled sample (NL). S, N, and L indicate respectively pUC19 (supercoiled), pUC19+ Nt-BbvCI (nicked) and pUC19+Hind III (linearized).

3.8 *In vivo* mapping of TOP 1 cleavage site.

Since *in vitro* tests gave us promising results, we moved on to the next phase by applying the optimized protocol to an *in vivo* system: a colorectal cancer cell line. To block TOP 1 bound to the DNA, asynchronous HCT 116 cells were treated with 20 μM CPT for 10 minutes, the DNA was then purified and checked for the presence of nicks. We performed a short cell treatment in order to prevent the activation of checkpoint pathway and/or damage response that could repair TOP 1 induced nicks. To test for the presence of CPT generated single strand breaks we take advantage of some interesting unpublished data collected in David Levens' lab about how nicks could influence the sonication of DNA. Following these findings we sonicated for 10s a small amount of CPT treated and untreated DNA. Then, we checked the fragment size of DNA on an agarose gel. Figure 20a shows a clear difference in the average fragment size between the untreated and CPT-treated DNA samples, the latter being more degraded. It appears that the presence of a nick in a double helix represents a weak point for nucleic acids that become more sensitive to break when subjected to shearing mechanical forces (Figure 20a).

By performing nick translation we next labeled the CPT generated nicks, and by doing immunoprecipitation of sheared DNA we separated the labeled fragments from the unlabeled ones. Finally the total DNA recovery was quantified relative to the starting amount of DNA and then qPCR analyses were performed as a final control for the experiment. Since in a previous paper Falaschi *et al* showed that the origin of replication of Laminin B2 and in particular the position B48 (see below) is a preferential binding site for TOP 1 [121] we designed primer sets around this region. The qPCR analyses showed an enrichment of DNA recovery at B48 position in the CPT treated sample in comparison to the untreated one (Figure 20c). We detected a very low signal at the repressed chromatin of centromeric α-satellite DNA chosen as a negative control site. Furthermore the DNA subjected to nick translation without labeled nucleotide (NL) used to set the background levels of the experiment showed very low quantification (Figure 20c). We also found that the total DNA recovery in CPT treated sample was increased with respect to the untreated one further confirming the goodness of the technique.

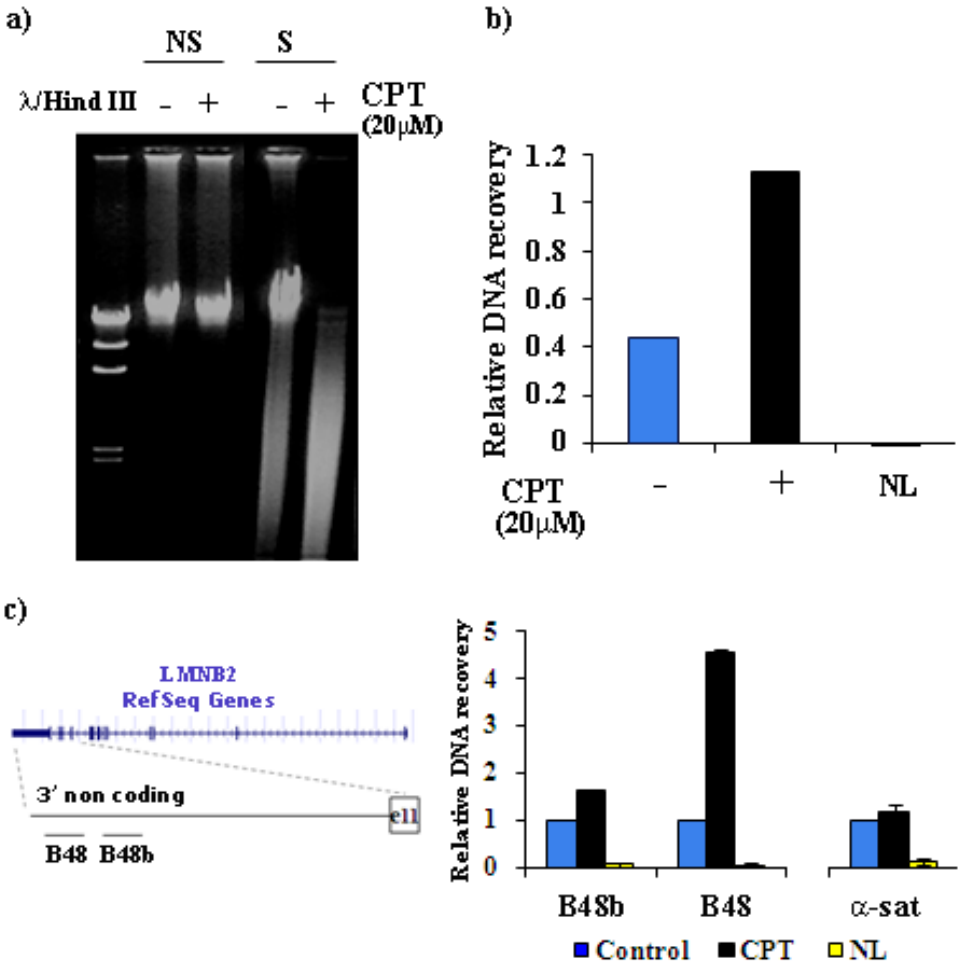

Figure 20. DNA recovery in CPT treated sample was increased with respect to the untreated one. a) HCT 116 cells were treated with 20 μM CPT for 10 minutes, the DNA was digested with proteinase K and SDS and purified with phenol extraction. 1 μg of DNA was sonicated (S) for 10s at low potency and run on a 1% agarose gel together with the non sonicated DNA (NS). λ/HindIII indicates λ DNA digested with Hind III used as marker. b) Nanodrop quantification of DNA recovery after immunoprecipitation with Anti-DIG Abs. Negative control of nick-translation is non digioxigenin labeled DNA (NL). Values are normalized for the starting amount of DNA. c) Left. The LMNB2 gene is indicated with exons (filled boxes) and introns (lines). The lower map is not to scale. Horizontal lines indicates primer in the 3' non coding region of the LMNB2 gene, localized approximatively 2.5 kb from exon 11. Right. Levels of DNA recovery at positions B48b and B48 and at centromeric α-satellite DNA. Values are normalized for the recovery in untreated sample.

We also applied the same protocol to CPT treated and untreated HCT 116 Top1siRNA cells. The reduced levels of TOP 1 in this system would be a good control for our technique since it would reveal the real TOP 1 cleavage sites. After immunoprecipitation with anti-DIG total DNA recovery was quantified and normalized to the starting DNA amount (Input). The graph in Figure 21a shows that in comparison to HCT 116 cells the DNA recoveries in HCT 116 Top1siRNA were at

least 10 times less along with the non labeled sample. As expected there is no difference between CPT treated and untreated sample. Moreover, levels of TOP 1 cleavage site at 3` non coding region of Laminin B2 gene quantified by qPCR were lower than in HCT 116 cells and were unchanged after CPT treatment (Figure 21b). These results showed a good level of confidence about the success of the technique and prompted us to perform the Solexa sequencing of the recovered DNA.

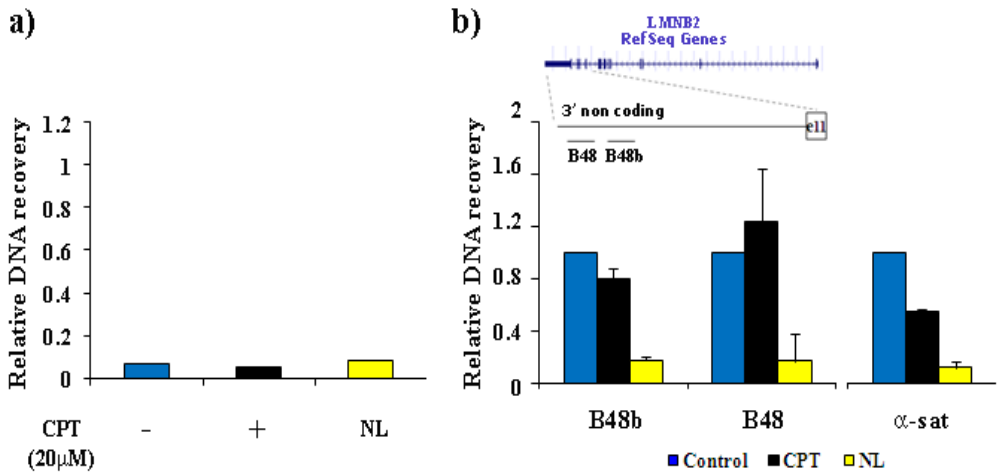

Figure 21. DNA recovery in HCT 116 (Top1siRNA) was unchanged between CPT treated and untreated sample. a) HCT 116 (Top1siRNA) cells were treated with 20 μM CPT for 10 min, the DNA was digested with proteinase K and SDS and purified with phenol extraction. After immunoprecipitation with Anti-DIG Abs total DNA recovery was quantified. Negative control of nick-translation is non digoxigenin labeled DNA (NL). Values are normalized for the starting amount of DNA. b) Levels of DNA recovery at two sites of 3` non coding region of Laminin B2 gene (B48b and B48 showed in the map) and at centromeric α-satellite DNA. Values are normalized for the recovery in untreated sample.

3.9 Comprehensive Genome-Wide Identification of TOP 1 cleavage sites.

To perform "Topo-Seq", we generated a library with the DNA recovered after immunoprecipitation and sequenced the sample using Illumina platform (Figure 22a). The sequencing procedure requires a one-step adaptor ligation and limited PCR amplification (18-21 cycles) of DNA molecules. The DNA is bound to the inside surface of the flow cell channels and amplified (cluster generation); then the sequences are read over 25 chemistry cycles. The image files generated by the analyzer were processed to produce DNA sequenced data (tags) using the Solexa Pipeline Analysis

(Figure 22a). The data were aligned to the NCBI Human Genome build 36 and only the ones with unique genomic position were used for analysis.

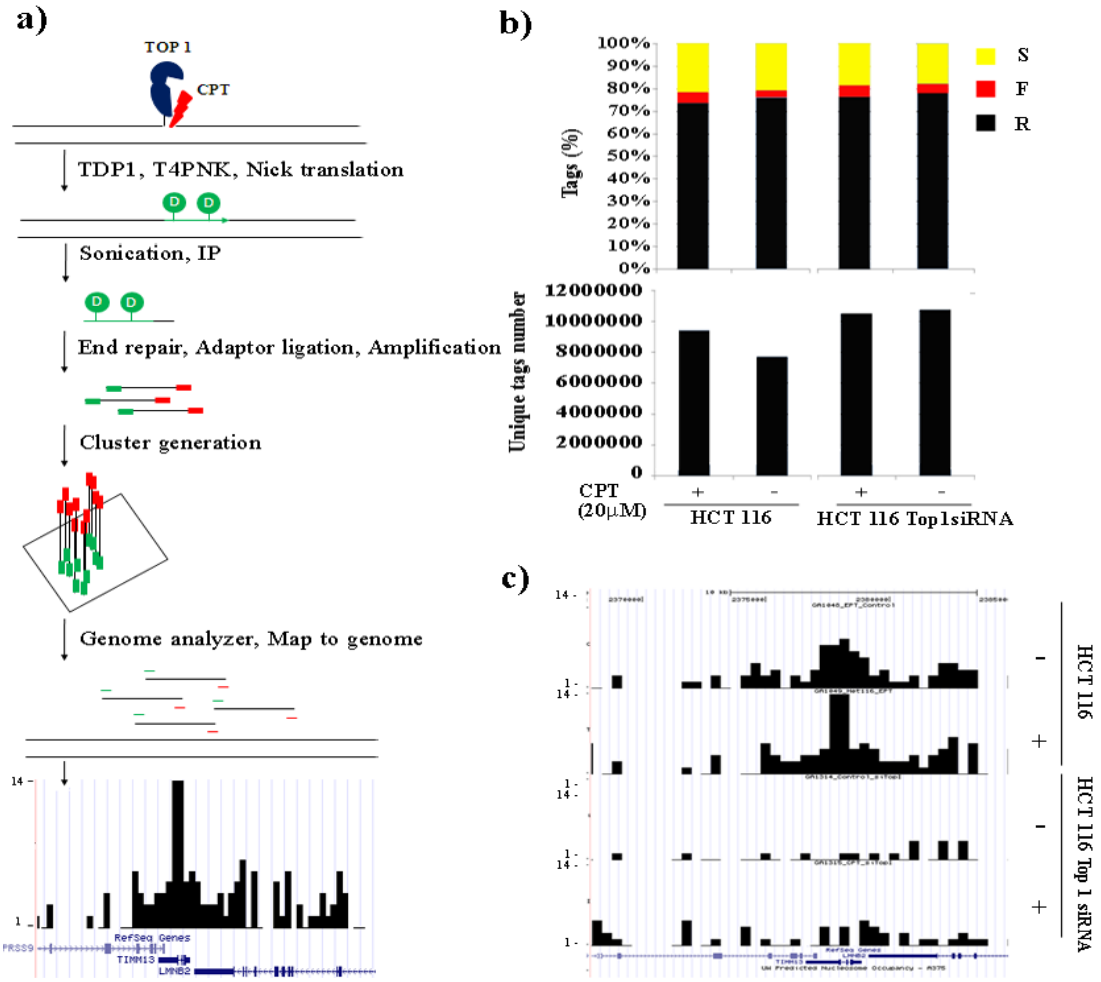

Figure 22. "Topo-Seq" strategy. a) The flow chart of "Topo-Seq". See materials and methods for details. b) "Topo-Seq" was performed on HCT 116 and HCT 116 (Top1siRNA) treated with 0 and 20 μM of CPT for 10 min. The image file generated by the analyzer was processed to produce DNA sequence data (tags). Tags were aligned to the NCBI Human Genome build 36. Upper panel. The percentage of tags aligned with repetitive sequence (suppressed-S), non aligned tags (failed-F) and tags aligned with unique genomic position (reported-R) is shown. Lower panel. Total number of uniquely aligned tags (R) is reported. c) Comparison of TOP 1 associated tags at 3' non coding region of Laminin B2 in different cell lines and in different drug treatments. The data are displayed as custom tracks on the UCSC genome browser.

Approximatively 75% of all the generated sequences were unique, i.e. they were mapped in a unique location in the NCBI genome sequence assembly (Figure 22b). About 20% of tags aligned with repetitive sequences and around 5% of them

failed to align. These data were discarded for any further analyses. The unique tags number generated in all different samples were around 9 million (Figure 22b)

To confirm the qPCR analysis at 3' non coding region of Laminin B2 gene (Fig 20c), we studied TOP 1 associated tags along this area. Figure 22c shows the distribution of tags in both HCT 116 and HCT 116 Top1siRNA cells treated with or without CPT. The data indicate the presence of a large peak around the origin of replication of Laminin B2 in both HCT 116 untreated and CPT treated cells. Notably, in the latter case the peak is higher and is likely due to the ability of the drug to block TOP 1 at a specific site. By blocking its target CPT increases the probability to find TOP 1 at a certain position. We did not observe any peak in the cell line with reduced levels of TOP 1 strongly suggesting that the tags are really associated to TOP 1 cleavage sites (Figure 22c).

Since TOP 1 is mainly involved in resolving topological stress that arise during transcription [34], to further determine the reliability of this method, we compared the distribution of tags with respect to the distribution of genes.

We retrieved all the human genes from Ensembl database (version 57) and we divided the whole genome into 4 regions: i) intragenic regions, defined by all genes location (which includes introns and exons); ii) regions 5kb upstream from transcription start site (TSS); iii) regions 5kb downstream from transcription termination (TT) and iv) intergenic regions, the rest of the genome. Next, for each data set, we counted the number of TOP 1 associated tags that overlaps with those regions. The results in Figure 23 showed that TOP 1 cleavage sites are generally enriched in intragenic regions. We found that about 60% of all TOP 1 cleavage site overlap with genes, upstream and downstream regions of the genome. This is in accord with previous data which establish that TOP 1cc occurs primarily in transcribed regions [35] and indicates that the "Topo-seq" is indeed a reliable method for capturing TOP 1 cleavage sites. Since tags distribution was unchanged between CPT treated and untreated samples we hypothesized that also in absence of the drug TOP 1 cleavage sites can be detected on the genome.

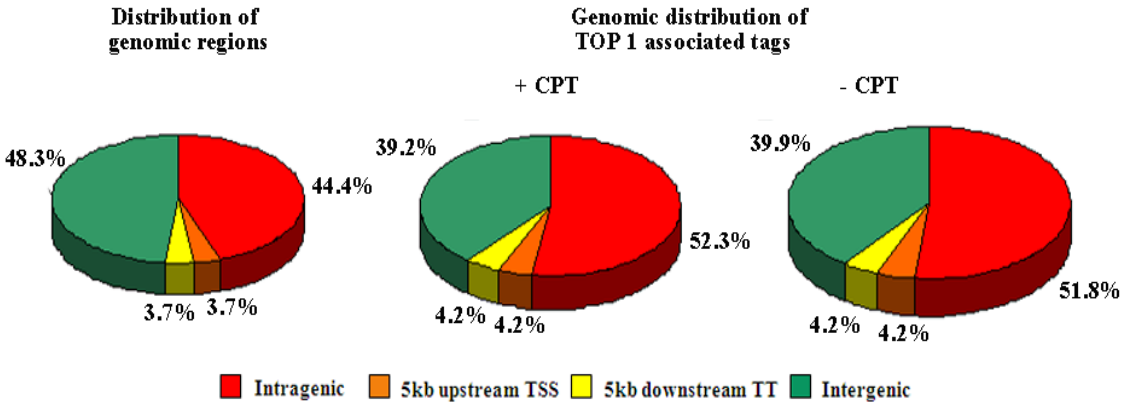

Figure 23. TOP 1 cleavage sites are enritched in intragenic regions Left. By using Ensembl database (version 57) the whole genome was divided into 4 regions: i) intragenic, defined by all genes location; ii) 5kb upstream TSS and not intragenic, iii) 5kb downstream TT, and not intragenic; and iv) intergenic, the rest of the genome. Right. Percentage of tags that overlaps with those regions in CPT treated and untreated HCT 116 cells.

In an attempt to gain more information about cleavage sites localization at a gene level, we next studied tags distribution at three separate gene position: around transcription start sites (TSS), in the middle of genes and around transcription terminations (TT). From Ensemble database we selected genes (approximately 21,000) long more than 5 kb. We next computed tags distribution near TSS and TT of those genes. Finally we aligned all these regions according to TSS, and TT, and compute the average signal. The tags profile showed in Figure 24 allowed us to make some considerations on the studied system:

- The amount of tags in the intragenic regions is higher than in the intergenic regions in both cell lines. As the data are normalized to the total number of tags in order to have the same amount of signal, the findings indicate that TOP 1 preferentially localizes in intragenic regions with respect to the rest of the genome.
- A big peak near TSS is observed in HCT 116 cells, with the untreated cells showing a peak higher than the treated one. The peak is centered approximately 100 bp downstream from TSS. We do not detect any peak in HCT 116 Top1siRNA at the same position, but rather a depletion of tags.

- HCT 116 cells also show a peak before TT which is lower than the one observed at TSS. Interestingly further downstream from the peak we observe a big depletion of tags in both treatment conditions. It is interesting to note that after CPT treatment the signal in the body of the gene and at TT increase with respect to the untreated sample where it is lower at TSS. This probably reflects the contribution of TOP 1 in removing positive supercoils which accumulate downstream the transcription machinery. Surprisingly in correspondence to the depletion of tags observed in the WT cells we detect a peak in HCT 116 Top1siRNA which is higher after CPT treatment.

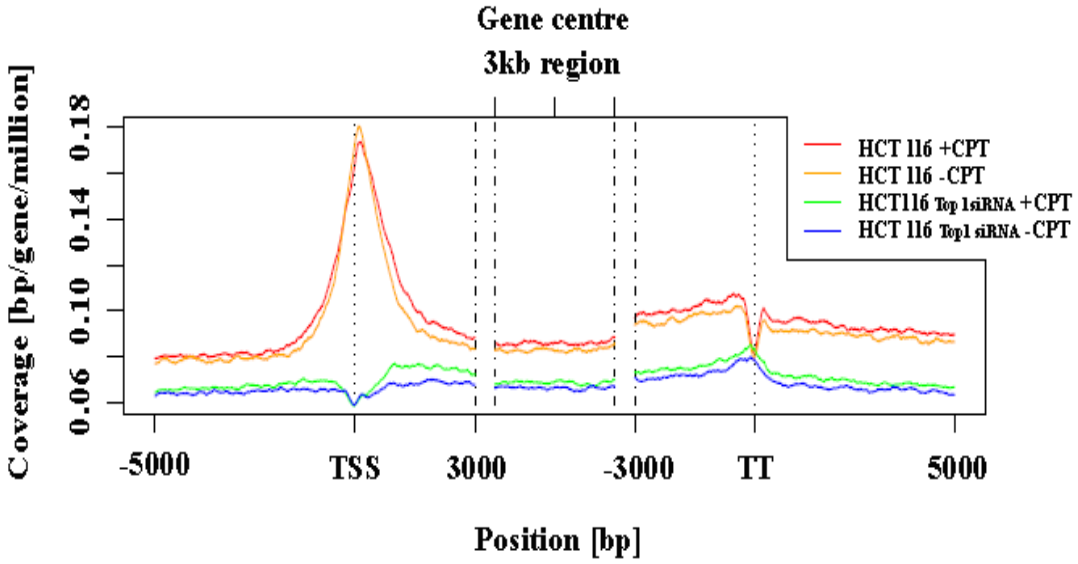

Figure 24. TOP 1 preferentially localizes around 5' and 3' ends of genes. Tags distribution around TSS, in the middle of genes and around TT in HCT 116 and HCT 116 (Top1siRNA) cells treated or not with 20 μM CPT. Approximately 21,000 genes were aligned relative to each positions and the detected tag numbers were calculated. The data are normalized for the total number of tags and multiplied by million.

Upon siRNA knockdown of TOP 1 much less signal corresponding to TOP 1 cleavage sites is detected at a gene level. However TOP 1 is not completely eliminated from the studied cells and in accord with this we observe an increase in tags number inside the gene and at the 3' ends in comparison to intergenic regions. Even if we still need further confirmation the data seems really intriguing since they reveal very different distribution profile of TOP 1 cleavage sites in the studied cell lines (Figure 24).

To evaluate the role of TOP 1 in transcription we also compared TOP 1 levels around TSS for each gene with the mRNA expression data of HCT 116 cells [49]. The raw microarray data for HCT116 cells (four replicates) were downloaded from GEO (accession number GSE7161)[49]. The average value over four replicates was computed to get a single expression value for each Ensembl gene. According to the expression value all genes were split into 3 categories: present (P), if the gene was unambiguously detected in all replicates and absent (A), if it was not detected in all replicates otherwise the gene was called marginal (M). Figure 25a shows the boxplot of final data. To evaluate the influence of gene expression on the TOP 1 activity we used the Mann-Whitney test for assessing whether two sample of observations (TOP 1 levels for expressed genes (P) and non-expressed (A) genes) come from the same distribution. Figure 25b shows a boxplot for the TOP 1 associated tags in CPT treated cells. The statistical test shows that TOP 1 levels associated to non-expressed and expressed genes are different with high levels of tags in the latter subpopulation. The correlation between the two data sets is weak, but it is statistically significant. Recently, it has been proposed that transcribed units are organized in architectural chromatin domains containing one or more gene. Taking account of this scenario it is probable that the amount of supercoils which accumulate during transcription is influenced by the length of the gene with short genes accumulating more supercoils than long ones. Hence, it would be helpful to evaluate the TOP 1 activity relative to high expressed short genes and high expressed long genes.

We still need to complete the picture with further analysis at a gene expression level but these preliminary findings mainly indicate that the developed procedure is a pioneer tool for the detection of TOP 1 cleavage sites across the genome and open the way to further investigation of TOP 1 role in different nuclear processes.

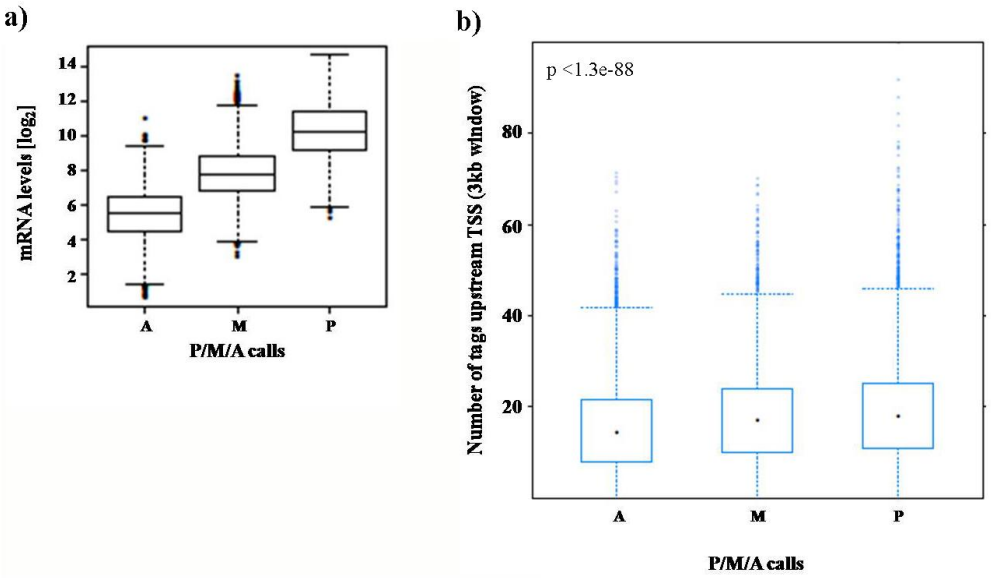

Figure 25. TOP 1 cleavage sites distribution correlate with mRNA expression levels of HCT 116 cells. a) The raw microarray data (platform: Affymetrix Human Genome U133 Plus 2.0 Array) for HCT 116 cells were downloaded from GEO (accession number GSE7161). CEL files were analysed using R environment with Bioconductor package. The raw probe level data were converted to expression values that correspond to those from MAS 5.0 (using mas5 and mas5calls methods implemented in affypackage). Normalized expression values, present (P), absent (A) and marginal (M) calls and their associated Wilcoxon p-values were calculated (see text for details). b) Boxplot of TOP 1 levels near TSS for P, A and M calls of mRNA expression levels The data on Y axis correspond to TOP 1 associated tags in CPT treated HCT 116 cells. p-value<1.3e-88.

Chapter 4

Discussion

The classical view of the role of DNA topology in transcription regulation has been recently challenged by several reports, showing that mechanical stresses, constrained by architectural features of DNA and chromatin, may broadly contribute to gene regulation. For instance, in a recent paper published by Vijayan and colleagues in 2009, it was shown that the topological status of the chromosome of the cyanobacterium *Synechococcus elongatus* is highly correlated with circadian global gene expression [122]. By measuring genome wide gene expression and superhelicity of an endogenous reporter plasmid, the authors demonstrated that each topological state correspond to a distinct state in gene expression. By modulating the stability of interactions between RNA polymerase and transcription factors with promoter or coding regions, DNA torsional stress strongly affects transcription and may induce activation or repression of gene expression. In eukaryotes at least for the c-MYC gene, it has been demonstrated that local DNA torsional tension can significantly regulate gene expression [123]. Using a plasmid system with two divergent promoters Kouzine *et al*. have shown that the degree of dynamic supercoiling increases as transcription intensifies. Notably dynamic negative supercoiling upstream of the c-MYC gene promoter can promote the formation of non-B-DNA structures in the susceptible FUSE (far upstream element) sequence, favoring the recruitment of transcription factors [123].

Thus, DNA topology has gained much more importance than previously thought, and the interest in enzymatic activities able to control superhelicity has been growing as well.

Our work is focused on the role of DNA topoisomerase 1 during transcription and on the transcriptional consequences associated with TOP 1 inhibition by CPT in human cell lines. Although the established cellular effects of camptothecin are peculiar of DNA damage responses, TOP 1cc occurs primarily in actively transcribed regions, but the transcription-dependent effects of TOP 1cc are not yet fully known. A broad and general inhibition of transcription elongation is an immediate effect of camptothecin in cultured cells, possibly due to the stalling of elongating RNA

polymerases by TOP 1ccs [34] and/or by persistent transcription-generated DNA supercoils [47, 112].

However, what we have found was quite unexpected since the transcriptional stress induced by CPT treatments can lead to increased RNAP II escape from promoter proximal pausing, marked alteration of HIF-1α co-transcriptional splicing, increased chromatin accessibility and activation/derepression of antisense transcripts. These events occur at an actively transcribed region and are not equal in time since for instance, some occur early and others later during drug treatments.

We have previously reported that short cell exposures to CPT affect RNAP II by promoting the hyperphosphorylation of its Rpb1 subunit and a specific reduction of RNAP II density at promoter by ChIP [35]. Here, we have further investigated this phenomenon by quantifying nascent RNAs downstream to the pausing site. As we found increased chromatin-bound RNAs after CPT treatment downstream to the pausing sites, we have thus demonstrated that an enhanced RNAP II escape from pausing sites can occur after TOP 1 inhibition at the human Hypoxia Inducible Factor 1α and c-MYC genes (results, Figure 8). Remarkably, the drug effect can be reversed through the inhibition of Cdk7 and Cdk9 kinase activities (results, Figure 9) supporting the idea that TOP 1cc can increase the activity of Cdks which in turn phosphorylate the Rpb1 subunit of RNAP II. Interestingly, camptothecin can disrupt the large inactive P-TEFb complex, thus releasing a free active P-TEFb complex (containing the Cdk9 subunit), which may then contribute to camptothecin-increased phosphorylation of Rpb1 [76].

A hyperphosphorylated polymerase is not competent for recruitment to promoters leading then to a reduction in RNAP II density at promoter sites of the studied genes (results, Figure 3) (Figure 1). Interestingly, Gilchrist *et al.* in 2008 demonstrated that a paused RNAP II enhances gene transcription by maintaining permissive chromatin architecture around the promoter-proximal regions allowing the recruitment of further RNAP II at the promoter sites [109]. This suggests that a promoter-proximal paused polymerase can activate further transcription. If this is the case, then our findings may suggest that, by enhancing polymerase escape from pausing sites, CPT could inhibit the recruitment of new enzymes at promoters therefore leading to reduced gene transcription.

Such interference with initiation regulatory mechanisms is likely specific for TOP 1 and camptothecin, as VM-26 (a TOP 2 poison) and cisplatin (which promotes the formation of DNA interstrand crosslinks) caused a decrease in RNAP II density both at promoters and along the transcribed gene (results, Figure 4).

The data also show another interesting aspect of the TOP 1 poison: in contrast to other transcription inhibitors, such as DRB and α-amanitin, CPT does not cause a complete block of transcription elongation since 1 hour of CPT treatments does not reduce the overall level of chromatin-associated RNAs along the studied loci (results, Figure 6). Our findings are in agreement with other published data. In an effort to perform a kinetic analysis of RNAP II transcription Darzacq *et al.* showed that transcription can be inefficient and that polymerase often pauses during elongation at a gene-array locus [112]. Interestingly, while leaving active the entire population of RNAP IIs, camptothecin increased the efficiency of intragenic pausing but not the pause time, resulting in a reduction of the elongation rate to a ¼ of the normal rate [112]. Thus, the authors concluded that camptothecin is not able to fully block transcription at the studied gene array [112].

Interestingly, previous findings from our lab showed a lack of influence of CPT on transcription factories under conditions that completely destroy replication factories as shown by fluorescence microscopy [35]. This likely suggests that TOP 1ccs is highly reversible in nuclear chromatin and do not arrest completely transcription of nascent RNAs. Unaffected nuclear transcriptional foci could therefore indicate that major destructive collisions do not often occur *in vivo* since an encounter between a trapped TOP 1 and an elongating RNAP II could transiently block polymerase movement without leading to irreversible strand cut and to polymerase disassembly from the template. Nevertheless, at high CPT concentrations, the collisions may be more frequent leading to irreversible single-strand cuts in living cells. Other studies have shown that TOP 1-dependent single-strand DNA damage is recognized and repaired in a transcription-dependent manner [70]. The repair mechanism involves removal of TOP 1ccs and DNA break processing, and is coupled to ubiquitination and degradation of TOP 1 and RNAP II through the 26 S proteasome pathway [70]. However, transcription-dependent ubiquitination of the polymerase has been detected in studies in which drug concentrations were higher and time periods longer that those used by us [35].

An unexpected finding of the present work is the demonstration that CPT can affect splicing regulation of HIF-1α pre-mRNA in two ways: by increasing the pre-mRNA splicing of exon 1 and 2 and by severely affecting introns 10 and 11 splicing (results, Figure 10). Since splicing alterations were detected only in the nascent RNA subpopulation and not in total RNA after short cell treatments, we can infer that both these events occur co-transcriptionally. Listerman *et al.* [115] showed already an increase in co-transcriptional splicing of the first intron of the FOS gene after CPT treatments, but it was interpreted in terms of obstruction to and increased pause of RNAP II by camptothecin. Since we demonstrate that CPT increases transcription downstream to the promoter-proximal pausing site without reducing the overall level of chromatin associated RNAs, we may suggest that co-transcriptional splicing is influenced more by the alteration of RNAP II regulation rather than by collisions between TOP 1ccs and translocating RNAP II. Notably some lines of evidence show that *in vitro,* the hyperphosphorylated CTD can stimulate splicing more than the hypophosphorylated form [124]. Furthermore, Bird *et al.* demonstrated that co-transcriptional splicing requires hyperphosphorilation of RNAP II since treatment with DRB inhibits processing of human β-globin pre-mRNA transcribed from an injected plasmid in *Xenopus* oocyte. This requirement is not needed when the pre-mRNA processing occurs uncoupled from transcription [125]. Since CPT increases hyperphosphorylation of RNAP II and its escape from the promoter pausing it is possible that this can lead to alterations of HIF-1α pre-mRNA as well.

CPT also affects splicing further downstream from the HIF-1α promoter pausing site (results, Figure 10). The spliced variants detected in CPT-treated samples indicate splicing alterations of introns 10 and 11 which could severely affect the protein functions. Indeed, exon 11 encodes for the oxygen-dependent degradation (ODD) domain of the protein, which is required for the ubiquitin/proteasome-dependent degradation of HIF-1α. Even if the biological meaning of this alteration in term of protein functions is still unknown, some evidence state that a HIF-1α splicing variant lacking part of ODD acts as dominant negative factor that competes with the endogenous protein and suppresses its activity [126]. CPT has been shown to impair HIF-1α target genes expression [101], which could then be correlated to the HIF-1α splicing alterations.

It is difficult to explain how an enzyme that regulates DNA topology may affect RNA processing but some evidence state that besides its DNA relaxation activity, TOP 1 is also implicated in the regulation of mRNA splicing in higher eukaryotes [34]. This presumably happens through the direct phosphorylation of splicing factors of the serine/arginine (SR)-rich family [127] to which the alternative splicing factor/splicing factor 2 (ASF/SF2) belongs. Proteomic analyses of complexes containing TOP 1 have identified the N-terminal domain of the protein as the main region that can interact with protein partners. Interestingly, 10 of the 36 proteins identified as interacting with TOP 1 are involved in RNA splicing [128]. One of the splicing factors, PSF, has been shown to stimulate DNA relaxation activity of TOP 1 [129], in contrast to ASF/SF2 splicing factor that seems to inhibit enzyme activity [130]. In an interesting report Tuduri *et al.* demonstrated that TOP 1 has a fundamental role in coordinating DNA fork progression and ongoing RNA transcription. By monitoring subnuclear localization of ASF/SF2 in cells depleted of TOP 1, the authors found a profound alteration of ASF/SF2 speckles organization [127]. Moreover depletion of TOP 1 target ASF/SF2, induces replication fork arrest. Accordingly, TOP 1 seems to prevent R-loop formation by both relaxing DNA supercoils and promoting the ASF/SF2-dependent assembly of mRNA particle complexes [127].

Furthermore, alternative splicing has been proposed to play a key role in the responses to DNA damage. In *Drosophila* it was shown that the ATM (ataxia-telangiectasia mutated) and ATR (ATM-RAD3 related) signal transduction pathways regulate TAF1 pre-mRNA alternative splicing in response to DNA damage induced by CPT or ionizing radiation [131, 132]. Interestingly, only the knockdown of ATR had a significant inhibitory effect on CPT induced alternative splicing. In contrast, ATM knockdown abrogated IR- induced spliced variants. This might indicate that the CPT and IR activate two different damage pathways that may converge on a common way to elicit the same effect on TAF1pre-mRNA splicing. Recently, the correlation between UV induced DNA damage and alternative splicing has been demonstrated in an elegant report published by Munoz *et al*. By studying the alternative splicing of human fibronectin in several cancer cells, the authors found that low doses of UV upregulate the co-transcriptional inclusion of EDI (extra domain I) exon in the fibronectin mRNA. Notably upon UV exposition the phosphorylation of RNAP II CTD increases and the elongation rate of the enzyme decreases. Thus the reduction of

elongation rate is thought to favor the use of weak splice site by the spicing machinery by increasing the time for their recognition [65]. Since CPT promotes Rpb1 hyperphosphorylation and HIF-1α alternative splicing, TOP 1ccs may influence the elongation rate of RNAP II and splicing activity with a similar mechanism proposed for UV induced damage.

Intriguingly, our data intriguingly reveal that the transcriptional consequences of persistent TOP 1ccs stabilized by camptothecin at relatively low concentrations are wider than those discussed above. Camptothecin (2-10 μM) can activate antisense transcription at the human HIF-1α gene locus (results, Figure 13), and increase levels of histone modifications marks of open chromatin conformation (results, Figure 17). These events require Cdk7 and Cdk9 activities since co-treatments with CPT and kinases inhibitor DRB revert these phenomena (results, Figures 13 and 18, respectively). Thus, it appears that transcriptional stress triggered by sustained TOP 1 inhibition may induce a more accessible chromatin conformation and activation of antisense transcription (Figure 1).

Indeed, several *in vitro* and *in vivo* evidence prove that CPT locally alters DNA topology eventually leading to a change in gene expression. Duann and co-workers first [132] and Koster *et al.* later [133] showed data suggesting that TOP 1 may be actively involved in maintaining the negatively supercoiled state of DNA in the cell. Single molecule manipulation experiment to monitor the dynamics of TOP 1 in the presence of its inhibitor topotecan (TPT) revealed that the drug impedes effectively the relaxation of positive, but not negative, supercoils [133]. Moreover, inhibition of the enzyme results in more positively supercoiled DNA in the cell [132]. If we consider that positive supercoils are in the path of an advancing DNA polymerase or RNA polymerase, drug inhibition of TOP 1 can lead to a persistent overwinding of DNA that could result in fork collapse and formation of lethal DNA lesion at replicating DNA. The authors suggested a new hypothesis of a potential new mechanism for drug-induced cell death as the inhibition of the TOP 1 catalytic activity rather than DNA cleavage by the enzyme would be responsible for cell death.

As a result of the dynamic nature of DNA-nucleosome interaction, DNA supercoils may propagate across large distance through the genome and this may regulate interactions between proteins and DNA [123]. In agreement with this view, TOP 1 may regulate nucleosome remodeling by modulating the torsional tension

generated by the assembly and/or disassembly of nucleosomes [134]. A recent genome-wide investigation showed that deletion of TOP 1 gene or its mutation are associated with increased acetylation of core histones at telomeric and sub-telomeric regions in *S. cerevisiae* [120]. Thus, the regulation of DNA supercoiling by DNA topoisomerases can be critical for physiological processes to proceed properly in human cells.

A strong evidence of the present report is that upon siRNA knockdown of TOP 1 we do not observe increased RNAP II escape and antisense transcription at the HIF-1α locus, supporting the idea that CPT effects are dependent on the presence of TOP 1 in the cell [13]. In addition, the observed CPT-related events are not associated to replication dependent DNA damage as treatments with the replication inhibitor aphidicoline or the checkpoint pathway inhibitor caffeine do not prevent antisense transcription and increased polymerase escape (results, Figure 16).

In some cases, supercoiling alterations induced by CPT may lead to the formation of transcription-dependent double-strand breaks. In non-proliferating primary neurons, it has been recently shown that transcription arrest by stalled TOP 1ccs activates the DSB-ATM-DDR (DNA damage response) pathway which induces the formation of γ-H2AX foci [72]. By blocking TOP 1 activity, camptothecin may promote an increase of local negative supercoiling behind the transcriptional apparatus thus stabilizing R-loop structures, which may increase genome instability [72].

Thus, when the supercoiling imbalance promoted by CPT occurs at promoters, it may cause interference with the RNAP II pausing regulation or activation/derepression of antisense transcripts. Even if we still miss the molecular mechanism of these processes, we can speculate that Cdk activity, and PTEF-b in particular, is involved in these phenomena (Figure 1)

An important finding of this work is that CPT induces natural antisense transcripts at the HIF-1α gene locus which may influence cancer-related molecular pathway. Since genome-wide natural antisense transcription has been reported for many organisms, it has become clear that many genomic loci contain transcription units on both strands [135]. Often one strand encodes for a protein, whereas the transcript from the other strand is non-coding. Such natural antisense transcripts can negatively regulate the conjugated sense transcript in several ways. One intriguing mechanism is the “transcriptional interference” [135]. Two bulky RNA polymerase II

complexes on opposite DNA strands might collide with and stall one another. The interference occurs mostly in the elongation step, resulting in either transcription arrest or transcription in one direction (sense or antisense) only. Such a mechanism might occur in cases in which inverse expression is observed [135]. Since we detect increased HIF-1α antisense production and decreased mRNA transcription, such a mechanism may explain our data.

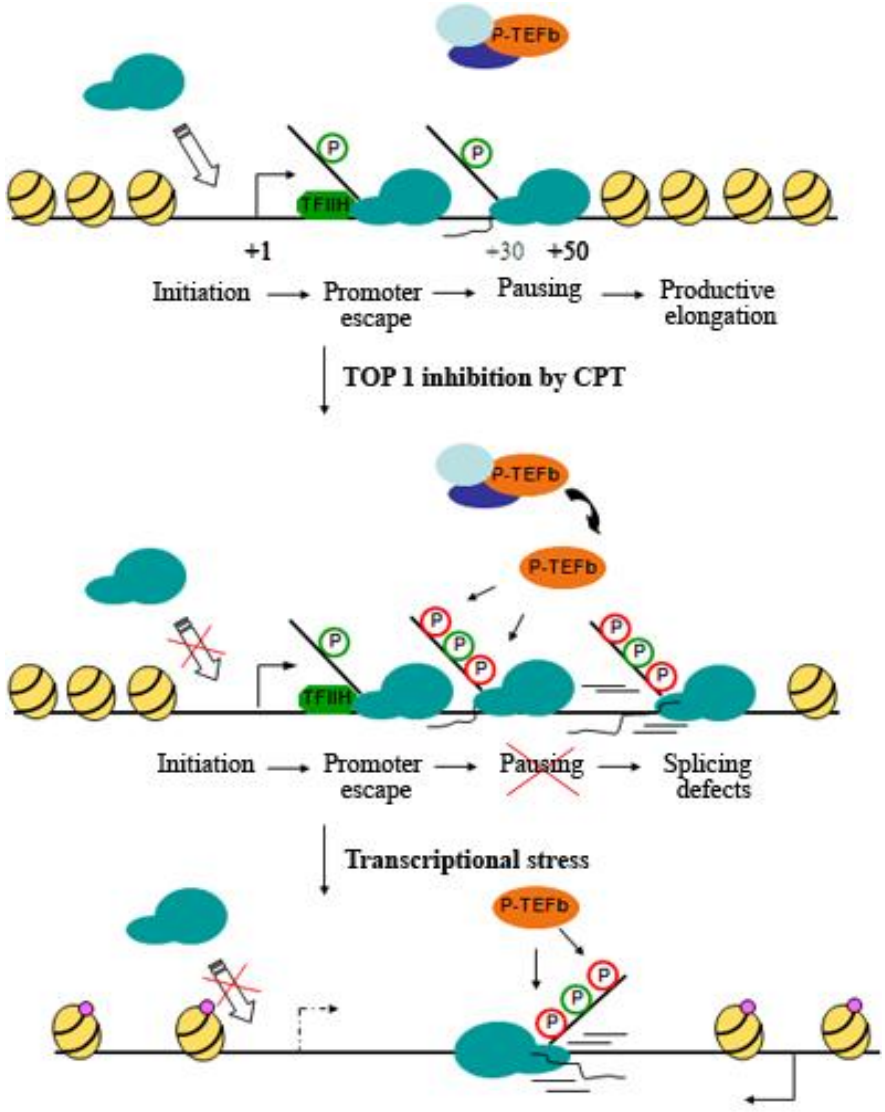

Figure 1. A model of the molecular mechanism triggered by CPT induced TOP 1ccs. Arrow indicate transcription start site and direction. The yellow sphere are nucleosome and the small pink sphere are histones acetylation markers. The cyan object is RNAP II.

Our findings demonstrate that HIF-1α gene locus is complex since at least two non-coding RNAs are transcribed from the locus, and one of them, the 3`aHIF-1α, was shown already to play an important role in transcriptional adaptations to hypoxia [116]. HIF-1 is an heterodimer constituted by an inducible subunit, HIF-1α or HIF-2α, and the constitutive HIF-1β subunit. When oxygen tension reduces, both HIF-1α and 2α are expressed but after prolonged hypoxia, the HIF-1α protein disappeared because of a reduction in its mRNA stability, whereas the HIF-2α protein remained at high levels in the cell. The prolonged hypoxia induces the natural antisense RNA (3`aHIF-1α) that destabilizes the HIF-1α mRNA, eventually decreasing HIF-1α protein and activity [116]. Even if we need to prove the biological function of the novel 5`aHIF-1α RNA, we speculate that it may have a role in transcriptional or post/transcriptional regulations of the HIF-1α mRNA.

In any case, camptothecin can impair the balance of cellular antisense and sense transcripts of the cancer-related HIF-1α gene. Interestingly, CPT has been shown to have an antiangiogenic activity in animal tumor models, and to markedly reduce HIF-1α protein accumulation in hypoxic cells. This phenomenon occurs in a transcription dependent fashion and is not correlated with VHL pathway, or CPT induced replicative damage [101, 136]. Thus, de-repression or activation of antisense RNAs by CPT may regulate the activity of HIF-1α and contribute to the control of tumor progression by TOP 1 poisons in animal models and human patients. This might constitute a different rational basis for the development of novel therapeutic approaches of human cancers.

Overall the findings revealed new unexpected functions of TOP 1 at gene level and prompted us to extend the investigations to the whole genome. “Topo-Seq” is a unique approach developed by us which offers a comprehensive view of TOP 1 cleavage sites across the genome. In comparison to a ChIP-Seq approach, which gives information only about protein localization, this novel strategy reveals the sites at which the enzyme is active. Taking advantage of the enzyme’s mechanism of action, we have optimized a specific procedure that allows the labeling of the TOP 1 associated nicks using nick translation and DIG-dUTP. Each step was thoroughly verified by both *in vitro* and *in vivo* cell tests in order to provide a population of DNA molecules that basically reflect the number of cleavage sites in the double helix.

Moreover, we have confirmed the reliability of the technique by comparing TOP 1 associated tags in two cell lines which is differing only in the amount of TOP 1 expression.

A global analysis of TOP 1 associated tags in HCT 116 cells, which express normal levels of TOP 1 protein, reveals that TOP 1 preferentially localizes at intragenic regions (results, Figure 23). By aligning the tags distribution near TSS and TT of almost 21,000 genes, we also found a high density of tags around 5'and 3'ends of genes (results, Figure 24). According to the "twin supercoiled model" this finding probably may reflect the involvement of the enzyme in removing transcription associated supercoils [11]. Moreover, since TSS is enriched in critical regulatory elements some of them sensitive to torsional stress [137], TOP 1 might be involved in promoter regulation. In the case of the c-MYC gene, for example, the far upstream FUSE element is sensitive to torsional stress and melts in the presence of negative supercoils. This occurs when c-MYC promoter is activated and negative supercoiling are generated by ongoing transcription. Melted FUSE recruits FBP, which enhances transcription by looping with TFIIH at the P2 promoter pausing site. Thus, the system regulates ongoing transcription by looping TFIIH at the major P2 promoter [137].

We may speculate that TOP 1 can also modulate DNA superhelicity at transcribed gene promoters in order to regulate gene activation. This could also explain why the peak at 5' is higher than the peak at 3' end. The observed data were further confirmed by the analyses that correlate mRNA expression levels with TOP 1 associated tags. The correlation was quite weak but statistically significant and will be better defined by further analysis aiming at linking TOP 1 activity with gene length. The large depletion of tags immediately downstream to TT is something difficult to explain and needs further investigation. Multiple causes might inhibit the binding of TOP 1 such as a particular chromatin structure or sequences which bind TOP 1 poorly.

A clarification needs to be done about how the analysis on tags distribution at the gene level (results, Figure 24) was performed: the data were normalized to the total number of tags (shown in results, Figure 22b) which is quite similar in the studied samples but it does not reflect the total amount of the DNA (results, Figure 20b) recovered after the immunoprecipitation. Therefore, if the peak at TSS is higher in condition A than in condition B it only means that TOP 1 prefers to localize near TSS in condition A relative to other places in the genome with respect to condition B.

However, the overall amount of TOP 1 near TSS in condition A might be lower than in condition B because it depends on the amount of DNA coming from each experiment. For instance this means that after CPT treatments TOP 1 prefers to stay in the body of the gene (or at the 3' end) with respect to the TSS.

A second information that emerges from the studied system is that TOP 1 cleavage sites can be caught even without CPT. This is in contrast with some early reports [138] which showed that TOP 1 associated nicks can be detected by using DNA denaturing alkaline elution only after CPT treatments. We still need to further confirm the data but we may argue that the sensibility and the resolution of our approach is higher than previous techniques aiming to detect the total distribution of TOP 1 induced cleavage sites. This finding also allowed us to make some considerations. According to single molecule nanomanipulation experiments performed by Koster *et al* to monitor the dynamics of TOP 1, the uncoiling rate of the enzyme is very rapid. In contrast, the presence of the CPT analog topotecan hinders TOP 1 mediated uncoiling [133]. Since the uncoiling rate somehow reflects the time TOP 1 is bound to the helix, our data suggest that most of TOP 1 molecules are covalently bound to the DNA at the time of harvesting cells, and the addition of CPT only increase the probability to find the enzyme in a certain area. Indeed total DNA recovery in CPT treated sample was increased with respect to the untreated one, confirming that after CPT treatment DNA has a larger amount of nick. These findings will be further confirmed by performing the same experiment without using the enzyme TDP1. Since TDP1 is required to label TOP1 associated nick, this experiment will show us nicks which are not generated by TOP 1.

Interestingly, we found a large difference in the amount and distribution of tags at gene levels in HCT 116 Top1siRNA cells (results, Figure 24). We detected less signal both at intergenic, intragenic regions and at the ends of genes, which strongly confirm the association between TOP 1 and the detected tags. Strikingly, upon TOP 1 siRNA which expression which impairs protein cellular contents by 80%, the peaks at TSS disappear and TOP 1 molecules are mostly localized around TTs. Thus, it seems that the shortage of TOP 1 in the cells induces the system to optimize its use in order to have TOP 1 only where it is really needed i.e. at 3'ends of genes. Several lines of evidence suggest a role for TOP 2 in facilitating transcription at specific promoters. The enzyme associates with promoter regions bearing breaks in double-stranded DNA

[139]. Moreover, TOP 1 and TOP 2 have redundant functions and at least in yeast they can compensate each other [13]. In higher eukaryotes *TOP 1* gene is essential [48] and we cannot completely abrogate its expression. Since an increase of TOP 2 has been observed in HCT 116 Top1siRNA [49] this might correspond to a compensatory mechanism of functional TOP 2 to TOP 1 reduction [24]. Taken together these data led us to speculate that TOP 1 is fundamental at TT but is not needed at the 5' end of genes where likely TOP 2 preferentially works. We aim to further confirm the hypothesis by both reverting HCT 116 Top1siRNA to the wt TOP 1 phenotype and by mapping TOP 2 distribution in HCT 116 Top1siRNA cells. In addition, this approach can be used to show the contribution of TOP 1 activity in different nuclear processes providing remarkable insights into the physiological and pathological roles of the enzyme.

Chapter 5

Bibliography

1. Watson, J.D., Crick, F.H.C., *Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid.* Nature., (1953). **171**(4356): p. 737–738.
2. Fogg, J.M., Catanese Jr. D. J., Randall, G. L., Swick, M. C., Zechiedrich, L. , *Differences Between Positively and Negatively Supercoiled DNA that Topoisomerases May Distinguish*, in *Mathematics of DNA Structure, Function and Interactions*. (2009), Springer New York. p. 73-121.
3. Liu, L.F., Wang, J.C., *DNA–DNA gyrase complex: the wrapping of the DNA duplex outside the enzyme.* Cell., (1978). **15** (3): p. 979–984.
4. Holmes, V.F., Cozzarelli, N. R., *Closing the ring: links between SMC proteins and chromosome partitioning, condensation, and supercoiling.* Proc.Natl. Acad. Sci. USA, (2000). **97**(4): p. 1322–1324.
5. Zechiedrich, E.L., Khodursky, A.B., Bachellier, S., Schneider, R., Chen, D., Lilley, D.M., Cozzarelli, N.R., *Roles of topoisomerases in maintaining steady-state DNA supercoiling in Escherichia coli.* J. Biol. Chem., (2000). **275** p. 8103–8113.
6. Crisona, N.J., Strick, T. R., Bensimon, D., Croquette, V., Cozzarelli, N. R., *Preferential relaxation of positively supercoiled DNA by E. coli topoisomerase IV in single-molecule and ensemble measurements.* Genes Dev, (2000). **14**(22): p. 2881-92.
7. Zechiedrich, E.L., Khodursky, A. B., Bachellier, S., Schneider, R., Chen, D., Lilley, D. M., Cozzarelli, N. R., *Roles of topoisomerases in maintaining steady-state DNA supercoiling in Escherichia coli.* J Biol Chem, (2000). **275**(11): p. 8103-13.
8. Champoux, J.J., *DNA topoisomerases: structure, function, and mechanism.* Annu. Rev. Biochem., (2001). **70**: p. 369–413.
9. Musgrave, D.R., Sandman, K. M., Reeve, J. N., *DNA binding by the archaeal histone HMf results in positive supercoiling.* Proc. Natl. Acad. Sci. U S A, (1991). **88**(23): p. 10397-401.
10. Wang, J.C., *Moving one DNA double helix through another by a type II DNA topoisomerase: the story of a simple molecular machine.* Q Rev Biophys., (1998). **31**(2): p. 107-44.
11. Liu, L.F., Wang, J.C., *Supercoiling of the DNA template during transcription.* Proc. Natl. Acad. Sci. USA, (1987). **84**(20): p. 7024-7.
12. Hiasa, H., Marians, K.J., *Two Distinct Modes of Strand Unlinking during theta-type DNA Replication.* J. Biol. Chem., (1996). **271**(35): p. 21529–21535.
13. Wang, J.C., *Cellular roles of DNA topoisomerases: A molecular perspective.* Nat. Rev. Mol. Cell. Biol., (2002). **3**(6): p. 430–440.
14. Schoeffler, A.J., Berger, J.M., *DNA topoisomerases: harnessing and constraining energy to govern chromosome topology.* Quart. Rev. Biophys., (2008) **41**(1): p. 41–101.
15. Bouthier de la Tour, C., Portemer, C., Huber, R., Forterre, P., Duguet, M., *Reverse gyrase in thermophilic eubacteria.* Journal of Bacteriology (1991). **173**(12): p. 3921–3923.

16. Koster, D.A., Croquette, V., Dekker, C., Shuman, S., Dekker, N. H., *Friction and torque govern the relaxation of DNA supercoils by eukaryotic topoisomerase IB.* Nature (2005). **434**(7033): p. 671–674.
17. Liu, L.F., Rowe, T. C., Yang, L., Tewey, K.M., Chen, G. L., *Cleavage of DNA by mammalian DNA topoisomerase II.* J. Biol. Chem., (1983). **258**(24): p. 15365–15370.
18. Crisona, N.J., Strick, T. R., Bensimon, D., Croquette, V., Cozzarelli, N. R. , *Preferential relaxation of positively supercoiled DNA by E. coli topoisomerase IV in single-molecule and ensemble measurements.* Genes Dev., (2000). **14**(22): p. 2881–2892.
19. Gellert, M., Mizuuchi, K., O`dea, M.H., Nash, H. A. , *DNA gyrase: an enzyme that introduces superhelical turns in DNA* Proc. Natl. Acad. Sci. USA, (1976). **73**(11): p. 3872–3876.
20. Nitiss, J.L., *Investigating the biological functions of DNA topoisomerases in eukaryotic cells.* Biochim. Biophys. Acta., (1998). **1400**(1-3): p. 63–81.
21. Spell, R.M., Holm, C., *Nature and distribution of chromosomal intertwinings in Saccharomyces cerevisiae.* Mol. Cell. Biol., (1994). **14**(2): p. 1465-76.
22. Shaiu, W.L., Hsieh, T. S., *Targeting to transcriptionally active loci by the hydrophilic N-terminal domain of Drosophila DNA topoisomerase I.* Mol. Cell. Biol. , (1998). **18**: p. 4358–4367.
23. Bermejo, R., Capra, T., Gonzalez-Huici, V., Fachinetti, D., Cocito, A., Natoli, G., Katou, Y., Mori, H., Kurokawa, K., Shirahige, K., Foiani, M., *Genome-organizing factors Top2 and Hmo1 prevent chromosome fragility at sites of S phase transcription.* Cell., (2009). **138**(5): p. 870-84.
24. Wang, J.C., Caron, P. R., Kim, R. A., *The role of DNA topoisomerases in recombination and genome stability: a double-edged sword?.* Cell. (1990). **62**(3): p. 403–406.
25. Zhu, Q., Pongpech, P., DiGate, R. J., *Type I topoisomerase activity is required for proper chromosomal segregation in Escherichia coli.* Proc Natl Acad Sci U S A, (2001). **98**(17): p. 9766-71.
26. Gangloff, S., McDonald, J. P., Bendixen, C., Arthur, L., Rothstein, R., *The yeast type I topoisomerase Top3 interacts with Sgs1, a DNA helicase homolog: a potential eukaryotic reverse gyrase.* Mol. Cell. Biol., (1994). **14**(12): p. 8391-8.
27. Castano, I.B., Heath-Pagliuso, S., Sadoff, B. U., Fitzhugh, D. J., Christman, M. F., *A novel family of TRF (DNA topoisomerase I-related function) genes required for proper nuclear segregation.* Nucleic Acids Res., (1996). **24**(12): p. 2404–2410.
28. Mirkovitch, J., Gasser, S. M., Laemmli, U. K., *Scaffold attachment of DNA loops in metaphase chromosomes.* J. Mol. Biol., (1988). **200**(1): p. 101-9.
29. Stewart, L., Ireton, G. C., Champoux, J. J., *The domain organization of human topoisomerase I.* J. Biol. Chem., (1996). **271**(13): p. 7602–7608.
30. Redinbo, M.R., Stewart, L., Kuhn, P., Champoux, J.J., Hol, W.G., *Crystal structures of human topoisomerase I in covalent and noncovalent complexes with DNA.* Science., (1998). **279**(5356): p. 1504-13.
31. Stewart, L., Redinbo, M.R., Qiu, X., Hol, W.G., Champoux, J.J., *A model for the mechanism of human topoisomerase I.* Science, (1998). **279**(5356): p. 1534-41.

32. Redinbo, M.R., Champoux, J.J., Hol, W.G., *Novel insights into catalytic mechanism from a crystal structure of human topoisomerase I in complex with DNA.* Biochemistry., (2000). **39**(23): p. 6832-40.
33. Muller, M.T., *Quantitation of eukaryotic topoisomerase I reactivity with DNA. Preferential cleavage of supercoiled DNA.* Biochim. Biophys. Acta., (1985). **824**(3): p. 263-7.
34. Pommier, Y., *Topoisomerase I inhibitors: camptothecins and beyond.* Nat. Rev. Cancer, (2006). **6**(10): p. 789-802.
35. Khobta, A., Ferri, F., Lotito, L., Montecucco, A., Rossi, R., Capranico, G. , *Early effects of topoisomerase I inhibition on RNA polymerase II along transcribed genes in human cells.* J Mol Biol., (2006) **357**(1): p. 127-38.
36. Shaiu, W.L., Hsieh, T.S., *Targeting to transcriptionally active loci by the hydrophilic N-terminal domain of Drosophila DNA topoisomerase I* Mol. Cell. Biol. , (1998). **18**(7): p. 4358-4367.
37. Brill, S.J., Sternglanz, R., *Transcription-dependent DNA supercoiling in yeast DNA topoisomerase mutants.* Cell., (1988) **54**(3): p. 403-11.
38. Drolet, M., *Growth inhibition mediated by excess negative supercoiling: the interplay between transcription elongation, R-loop formation and DNA topology.* Mol. Microbiol., (2006). **59**(3): p. 723-30.
39. Giaever, G.N., Snyder, L., Wang, J.C., *DNA supercoiling in vivo.* Biophys. Chem., (1988). **29**(1-2): p. 7-15.
40. Kretzschmar, M., Meisterernst, M., Roeder, R.G., *Identification of human DNA topoisomerase I as a cofactor for activator-dependent transcription by RNA polymerase II.* Proc. Natl. Acad. Sci. USA, (1993). **90**(24): p. 11508-12.
41. Merino, A., Madden, K.R., Lane, W.S., Champoux, J.J., Reinberg, D., *DNA topoisomerase I is involved in both repression and activation of transcription.* Nature (1993) **365**(6443): p. 227-32.
42. Shykind, B.M., Kim, J., Stewart, L., Champoux, J.J., Sharp, P.A., *Topoisomerase I enhances TFIID-TFIIA complex assembly during activation of transcription* Genes Dev., (1997). **11** (3): p. 397-407.
43. Carty, S.M., Greenleaf, A.L., *Hyperphosphorylated C-terminal repeat domain-associating proteins in the nuclear proteome link transcription to DNA/chromatin modification and RNA Processing.* Molecular & Cell Proteomics., (2002). **1**(8): p. 598-610.
44. Mizutani, M., Ohta, T., Watanabe, H., Handa, H., Hirose, S., *Negative supercoiling of DNA facilitates an interaction between transcription factor IID and the fibroin gene promoter.* Proc. Natl. Acad. Sci. USA (1991). **88** (3): p. 718–722.
45. Liu, J., He, L., Collins, I., Ge, H., Libutti, D., Li, J., Egly, J.M., Levens, D., *The FBP interacting repressor targets TFIIH to inhibit activated transcription.* Mol. Cell., (2000). **5**(2): p. 331-41.
46. Weber, A., Liu, J., Collins, I., Levens, D., *TFIIH operates through an expanded proximal promoter to fine-tune c-myc expression.* Mol. Cell. Biol., (2005). **25**(1): p. 147-61.
47. Capranico, G., Ferri, F., Fogli, M. V., Russo, A., Lotito, L., Baranello, L., *The effects of camptothecin on RNA polymerase II transcription: roles of DNA topoisomerase I.* Biochimie, (2007). **89**(4): p. 482-9.

48. Morham, S.G., Kluckman, K. D., Voulomanos, N., Smithies, O., *Targeted disruption of the mouse topoisomerase I gene by camptothecin selection.* Mol. Cell. Biol., (1996). **16**(12): p. 6804-9.
49. Miao, Z.H., Player, A., Shankavaram, U., Wang, Y. H., Zimonjic, D. B., Lorenzi, P. L., Liao, Z. Y., Liu, H., Shimura, T., Zhang, H. L., Meng, L. H., Zhang, Y. W., Kawasaki, E. S., Popescu, N. C., Aladjem, M. I., Goldstein, D. J., Weinstein, J. N., Pommier, Y., *Nonclassic functions of human topoisomerase I: genome-wide and pharmacologic analyses.* Cancer Res., (2007). **67**(18): p. 8752-61.
50. Nitiss, J.L., *Targeting DNA topoisomerase II in cancer chemotherapy.* Nat Rev Cancer, (2009). **9**(5): p. 338-50.
51. Pommier, Y., Barcelo, J.M., Rao, V.A., Sordet, O., Jobson, A.G., Thibaut, L., Miao, Z.H., Seiler, J.A., Zhang, H., Marchand, C., Agama, K., Nitiss, J.L., Redon, C., *Repair of topoisomerase I-mediated DNA damage.* Prog. Nucleic Acid. Res. Mol. Biol., (2006). **81**: p. 179–229.
52. Pourquier, P., Pommier, Y., *Topoisomerase I-mediated DNA damage.* Adv. Cancer Res., (2001). **80**: p. 189–216.
53. Hsiang, Y.H., Hertzberg, R., Hecht, S., Liu, L. F., *Camptothecin induces protein-linked DNA breaks via mammalian DNA topoisomerase I.* J. Biol. Chem., (1985). **260**(27): p. 14873–14878.
54. Nitiss, J., Wang, J. C., *DNA topoisomerase-targeting antitumor drugs can be studied in yeast.* Proc. Natl. Acad. Sci. USA, (1988). **85**(20): p. 7501–7505.
55. Giovanella, B.C., Stehlin, J. S., Wall, M. E., Wani, M. C., Nicholas, A. W., Liu, L. F., Silber, R., Potmesil, M., *DNA topoisomerase I--targeted chemotherapy of human colon cancer in xenografts.* Science., (1989). **246**(4933): p. 1046-8.
56. Kaufmann, S.H., Charron, M., Burke, P. J., Karp, J. E., *Changes in topoisomerase I levels and localization during myeloid maturation in vitro and in vivo.* Cancer Res., (1995). **55**(6): p. 1255-60.
57. Li, T.K., Houghton, P. J., Desai, S. D., Daroui, P., Liu, A. A., Hars, E. S., Ruchelman, A. L., LaVoie, E. J., Liu, L. F., *Characterization of ARC-111 as a novel topoisomerase I-targeting anticancer drug.* Cancer Res., (2003). **63**(23): p. 8400-7.
58. Hertzberg, R.P., Caranfa, M.J., Hecht, S.M., *On the mechanism of topoisomerase I inhibition by camptothecin: evidence for binding to an enzyme-DNA complex.* Biochemistry., (1989). **28**(11): p. 4629–38.
59. Hertzberg, R.P., Busby, R.W., Caranfa, M.J., Holden, K.G., Johnson, R.K., et al., *Irreversible trapping of the DNAtopoisomerase I covalent complex. Affinity labeling of the camptothecin binding site.* J. Biol. Chem., (1990). **265**(31): p. 19287–95.
60. Hsiang, Y.H., Lihou, M.G., Liu, L.F., *Arrest of replication forks by drugstabilized topoisomerase I-DNA cleavable complexes as a mechanism of cell killing by camptothecin.* Cancer Res., (1989). **49**(18): p. 5077–82.
61. Furuta, T., et al., *Phosphorylation of histone H2AX and activation of Mre11, Rad50, and Nbs1 in response to replication-dependent DNA-double-strand breaks induced by mammalian DNA topoisomerase I cleavage complexes.* J. Biol. Chem., (2003). **278**: p. 20303–20312.

62. Zhang, H., Wang, J.C., Liu, L.F., *Involvement of DNA topoisomerase I in transcription of human ribosomalRNAgenes.* Proc. Natl. Acad. Sci. USA, (1988). **85**(4): p. 1060–64.
63. Collins, I., Weber, A., Levens,D., *Transcriptional consequences of topoisomerase inhibition.* Mol. Cell. Biol., (2001). **21**(24): p. 8437–8451.
64. Sordet, O., Larochelle, S., Nicolas, E., Stevens, E. V., Zhang, C., Shokat, K. M., Fisher, R. P., Pommier, Y., *Hyperphosphorylation of RNA polymerase II in response to topoisomerase I cleavage complexes and its association with transcription- and BRCA1-dependent degradation of topoisomerase I.* J Mol Biol, (2008). **381**(3): p. 540-9.
65. Munoz, M.J., Perez Santangelo, M. S., Paronetto, M. P., de la Mata, M., Pelisch, F., Boireau, S., Glover-Cutter, K., Ben-Dov, C., Blaustein, M., Lozano, J. J., Bird, G., Bentley, D., Bertrand, E., Kornblihtt, A. R., *DNA damage regulates alternative splicing through inhibition of RNA polymerase II elongation.* Cell., (2009). **137**(4): p. 708-20.
66. Bendixen, C., Thomsen, B., Alsner, J., Westergaard, O., *Camptothecin-stabilized topoisomerase I–DNA adducts cause premature termination of transcription.* Biochemistry., (1990). **29**(23): p. 5613–5619.
67. Wu, J., Liu, L. F., *Processing of topoisomerase I cleavable complexes into DNA damage by transcription.* Nucleic Acids Res., (1997). **25**(21): p. 4181–4186.
68. Holm, C., Covey, J. M., Kerrigan, D., Pommier, Y., *Differential requirement of DNA replication for the cytotoxicity of DNA topoisomerase I and II inhibitors in Chinese hamster DC3F cells* Cancer Res. , (1989). **49**(22): p. 6365–6368.
69. Squires, S., Ryan, A. J., Strutt, H. L., Johnson, R. T., *Hypersensitivity of Cockayne's syndrome cells to camptothecin is associated with the generation of abnormally high levels of double strand breaks in nascent DNA.* Cancer Res., (1993). **53**(9): p. 2012-9.
70. Desai, S.D., Zhang, H., Rodriguez-Bauman, A., Yang, J.M., Wu, X., Gounder, M.K., Rubin, E.H., Liu, L.F., *Transcription-dependent degradation of topoisomerase I–DNA covalent complexes.* Mol. Cell. Biol., (2003). **23**(7): p. 2341–2350.
71. Desai, S.D., Li, T. K., Rodriguez-Bauman, A., Rubin, E. H., Liu, L. F., *Ubiquitin/26S proteasomemediated degradation of topoisomerase I as a resistance mechanism to camptothecin in tumor cells.* Cancer Res., (2001). **61**(15): p. 5926–5932.
72. Sordet, O., Redon, C. E., Guirouilh-Barbat, J., Smith, S., Solier, S., Douarre, C., Conti, C., Nakamura, A. J., Das, B. B., Nicolas, E., Kohn, K. W., Bonner, W. M., Pommier, Y., *Ataxia telangiectasia mutated activation by transcription- and topoisomerase I-induced DNA double-strand breaks.* EMBO Rep., (2009). **10**(8): p. 887-93.
73. Li, X., Manley, J. L., *Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability.* Cell., (2005). **122**(3): p. 365-78.
74. Soret, J., Gabut, M., Dupon, C., Kohlhagen, G., Stévenin, J., Pommier, Y., Tazi, J., *Altered serine/arginine-rich protein phosphorylation and exonic enhancer-dependent splicing in mammalian cells lacking topoisomerase I.* Cancer Res. , (2003). **63**(23): p. 8203–8211.
75. Sordet, O., Larochelle, S., Nicolas, E., Stevens, E.V., Zhang, C., Shokat, K.M., Fisher, R.P., Pommier. Y., *Hyperphosphorylation of RNA polymerase II in*

response to topoisomerase I cleavage complexes and its association with transcription- and BRCA1-dependent degradation of topoisomerase I. J. Mol. Biol., (2008). **381**: p. 540-9.
76. Amente, S., Gargano, B., Napolitano, G., Lania, L, Majello, B., *Camptothecin releases P-TEFb from the inactive 7SK snRNP complex.* Cell. Cycle, (2009). **8**(8): p. 1249-55.
77. Luger, K., et al., *Characterization of nucleosome core particles containing histone proteins made in bacteria.* J. Mol. Biol., 1997. **272**(3): p. 301-11.
78. Luger, K. and T.J. Richmond, *The histone tails of the nucleosome.* Curr Opin Genet Dev., 1998. **8**(2): p. 140-6.
79. Tsukiyama, T., et al., *ISWI, a member of the SWI2/SNF2 ATPase family, encodes the 140 kDa subunit of the nucleosome remodeling factor.* Cell., 1995. **83**(6): p. 1021-6.
80. Dirscherl, S.S. and J.E. Krebs, *Functional diversity of ISWI complexes.* Biochem Cell. Biol., 2004. **82**(4): p. 482-9.
81. Shilatifard, A., *Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression.* Annu Rev Biochem, 2006. **75**: p. 243-69.
82. Strahl, B.D. and C.D. Allis, *The language of covalent histone modifications.* Nature., 2000. **403**(6765): p. 41-5.
83. Jenuwein, T. and C.D. Allis, *Translating the histone code.* Science, 2001. **293**(5532): p. 1074-80.
84. Wang, Z., Zang, C., Cui, K., Schones, D. E., Barski, A., Peng, W., Zhao, K., *Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes.* Cell., (2009). **138**(5): p. 1019-31.
85. Pokholok, D.K., et al., *Genome-wide map of nucleosome acetylation and methylation in yeast.* Cell., 2005. **122**(4): p. 517-27.
86. Cairns, B.R., *The logic of chromatin architecture and remodelling at promoters.* Nature., (2009). **461**(7261): p. 193-8.
87. Fuda, N.J., Ardehali, M. B., Lis, J. T., *Defining mechanisms that regulate RNA polymerase II transcription in vivo.* Nature., (2009). **461**(7261): p. 186-92.
88. Sims III, R.J., Belotserkovskaya, R., Reinberg, D., *Elongation by RNA polymerase II: the short and long of it.* Genes Dev., (2004). **18**(20): p. 2437-2468.
89. Orphanides, G., Reinberg, D., *RNA polymerase II elongation through chromatin.* Nature., (2000). **407**(6803): p. 471-5.
90. Saunders, A., Core, L.J., Lis, J.T., *Breaking barriers to transcription elongation.* Nat. Rev. Mol.Cell. Biol., (2006). **7**(8): p. 557-567.
91. Maniatis, T., Reed, R., *An extensive network of coupling among gene expression machines.* Nature., (2002). **416**(6880): p. 499– 506.
92. Komarnitsky, P., Cho, E.J., Buratowski, S., *Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription.* Genes Dev., (2000). **14**(19): p. 2452–2460.
93. Rougvie, A.E., Lis, J.T., *The RNA polymerase II molecule at the 5′ end of the uninduced hsp70 gene of D. melanogaster is transcriptionally engaged.* Cell., (1988). **54**(6): p. 795-804.
94. Lis, J., *Promoter-associated pausing in promoter architecture and postinitiation transcriptional regulation.* Cold Spring Harb. Symp. Quant. Biol., (1998). **63**: p. 347–356.

95. Peterlin, B.M., Price, D.H., *Controlling the elongation phase of transcription with P-TEFb.* Mol. Cell., (2006). **23**(3): p. 297-305.
96. He, N., Jahchan, N.S., Hong, E., Li, Q., Bayfield, M.A., Maraia, R.J., et al., *A La-related protein modulates 7SK snRNP integrity to suppress P-TEFbdependent transcriptional elongation and tumorigenesis.* Mol. Cell., (2008). **29**(5): p. 588-99.
97. Wada, T., Orphanides, G., Hasegawa, J., Kim, D. K., Yamaguchi, Y., Fukuda, A., Hisatake, K., Oh, S., Reiberg, D., Handa, H., *FACT relieves DSIF/NELF-mediated inibition of transcriptional elongation and reveals functional differences between P-TEFb and TFIIH.* Mol. Cell., (2000). **5** (6): p. 1067-1072.
98. Dantonel, J.C., Murthy, K. G., Manley, J. L., Tora, L., *Transcription factor TFIID recruits factor CPSF for formation of 3' end of mRNA.* Nature., (1997). **389**(6649): p. 399-402.
99. Glover-Cutter, K., Kim, S., Espinosa, J., Bentley, D. L., *RNA polymerase II pauses and associates with pre-mRNA processing factors at both ends of genes.* Nat. Struct. Mol. Biol., (2008). **15**(1): p. 71-8.
100. Semenza, G.L., *Targeting HIF-1 for cancer therapy.* Nat. Rev. Cancer, (2003). **3**(10): p. 721-32.
101. Rapisarda, A., Uranchimeg, B., Sordet, O., Pommier, Y., Shoemaker, R.H., Melillo, G., *Topoisomerase I-mediated inhibition of hypoxia-inducible factor 1: mechanism and therapeutic implications.* Cancer Res., (2004). **64**(4): p. 1475-82.
102. Laughner, E.P., Taghavi, K., Chiles, P., Mahon, C., Semenza, G.L., *HER2 (neu) signaling increases the rate of hypoxia-inducible factor 1alpha (HIF-1alpha) synthesis: novel mechanism for HIF-1-mediated vascular endothelial growth factor expression.* Mol. Cell. Biol., (2001). **21** (12): p. 3995–4004.
103. Thrash-Bingham, C.A., Tartof, K.D., *aHIF: a natural antisense transcript overexpressed in human renal cancer and during hypoxia.* J. Natl. Cancer Inst., (1999). **91**(2): p. 143–151.
104. Rossignol, F., Vaché, C., Clottes, E., *Natural antisense transcripts of hypoxia-inducible factor 1alpha are detected in different normal and tumour human tissues.* Gene., (2002). **299**(1-2): p. 135-40.
105. Iyer, N.V., Kotch, L. E., Agani, F., Leung, S. W., Laughner, E., Wenger, R. H., Gassmann, M., Gearhart, J. D., Lawler, A. M., Yu, A. Y., Semenza, G. L., *Cellular and developmental control of O2 homeostasis by hypoxia-inducible factor 1 alpha.* Genes Dev, (1998). **12**(2): p. 149-62.
106. Muse, G.W., Gilchrist, D. A., Nechaev, S., Shah, R., Parker, J. S., Grissom, S. F., Zeitlinger, J., Adelman, K., *RNA polymerase is poised for activation across the genome.* Nat Genet, (2007). **39**(12): p. 1507-11.
107. Zeitlinger, J., Stark, A., Kellis, M., Hong, J. W., Nechaev, S., Adelman, K., Levine, M., Young, R. A., *RNA polymerase stalling at developmental control genes in the Drosophila melanogaster embryo.* Nat Genet, (2007). **39**(12): p. 1512-6.
108. Guenther, M.G., Levine, S. S., Boyer, L. A., Jaenisch, R., Young, R. A., *A chromatin landmark and transcription initiation at most promoters in human cells.* Cell., (2007). **130**(1): p. 77-88.
109. Gilchrist, D.A., Nechaev, S., Lee, C., Ghosh, S. K., Collins, J. B., Li, L., Gilmour, D. S., Adelman, K., *NELF-mediated stalling of Pol II can enhance*

gene expression by blocking promoter-proximal nucleosome assembly. Genes Dev, (2008). **22**(14): p. 1921-33.
110. Gilbert, C., Kristjuhan, A., Winkler, G. S., Svejstrup, J. Q., *Elongator interactions with nascent mRNA revealed by RNA immunoprecipitation.* Mol Cell., (2004). **14**(4): p. 457-64.
111. Pandit, S., Wang, D., Fu, X. D., *Functional integration of transcriptional and RNA processing machineries.* Curr. Opin. Cell. Biol., (2008). **20**(3): p. 260-5.
112. Darzacq, X., Shav-Tal, Y., de Turris, V., Brody, Y., Shenoy, S. M., Phair, R. D., Singer, R. H., *In vivo dynamics of RNA polymerase II transcription.* Nat. Struct. Mol. Biol., (2007). **14**(9): p. 796-806.
113. Huang, L.E., Gu, J., Schau, M., Bunn, H. F., *Regulation of hypoxia-inducible factor 1alpha is mediated by an O2-dependent degradation domain via the ubiquitin-proteasome pathway.* Proc Natl Acad Sci U S A, (1998). **95**(14): p. 7987-92.
114. Chun, Y.S., et al., *A dominant-negative isoform lacking exons 11 and 12 of the human hypoxia-inducible factor-1alpha gene.* Biochem J, 2002. **362**(Pt 1): p. 71-9.
115. Listerman, I., Sapra, A. K., Neugebauer, K. M., *Cotranscriptional coupling of splicing factor recruitment and precursor messenger RNA splicing in mammalian cells.* Nat. Struct. Mol. Biol., (2006). **13**(9): p. 815-22.
116. Uchida, T., Rossignol, F., Matthay, M.A., Mounier, R., Couette, S., Clottes, E., Clerici, C., *Prolonged hypoxia differentially regulates hypoxia-inducible factor (HIF)-1alpha and HIF-2alpha expression in lung epithelial cells: implication of natural antisense HIF-1alpha.* J Biol Chem., (2004). **279**(15): p. 14871-8.
117. Furuta, T., et al., *Phosphorylation of histone H2AX and activation of Mre11, Rad50, and Nbs1 in response to replication-dependent DNA double-strand breaks induced by mammalian DNA topoisomerase I cleavage complexes.* J Biol Chem, (2003). **278**(22): p. 20303-12.
118. Yu, W., Gius, D., Onyango, P., Muldoon-Jacobs, K., Karp, J., Feinberg, A. P., Cui, H., *Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA.* Nature., (2008). **451**(7175): p. 202-6.
119. Nicolas, E., Yamada, T., Cam, H. P., Fitzgerald, P. C., Kobayashi, R., Grewal, S. I., *Distinct roles of HDAC complexes in promoter silencing, antisense suppression and DNA damage protection.* Nat. Struct. Mol. Biol, (2007). **14**(5): p. 372-80.
120. Lotito, L., Russo, A., Chillemi, G., Bueno, S., Cavalieri, D., Capranico, G., *Global transcription regulation by DNA topoisomerase I in exponentially growing Saccharomyces cerevisiae cells: activation of telomere-proximal genes by TOP1 deletion.* J. Mol. Biol., (2008). **377**(2): p. 311-22.
121. Abdurashidova, G., Radulescu, S., Sandoval, O., Zahariev, S., Danailov, M. B., Demidovich, A., Santamaria, L., Biamonti, G., Riva, S. Falaschi, A., *Functional interactions of DNA topoisomerases with a human replication origin.* EMBO J, (2007). **26**(4): p. 998-1009.
122. Vijayan, V., Zuzow, R., O'Shea, E. K., *Oscillations in supercoiling drive circadian gene expression in cyanobacteria.* Proc Natl Acad Sci U S A, (2009). **106**(52): p. 22564-8.
123. Kouzine, F., Sanford, S., Elisha-Feil, Z., Levens, D., *The functional response of upstream DNA to dynamic supercoiling in vivo.* Nat. Struct. Mol. Biol., (2008). **15**(2): p. 146-54.

124. Hirose, Y., Tacke, R., Manley, J. L., *Phosphorylated RNA polymerase II stimulates pre-mRNA splicing.* Genes Dev, (1999). **13**(10): p. 1234-9.
125. Bird, G., Zorio, D. A., Bentley, D. L., *RNA polymerase II carboxy-terminal domain phosphorylation is required for cotranscriptional pre-mRNA splicing and 3'-end formation.* Mol. Cell. Biol., (2004). **24**(20): p. 8963-9.
126. Chun, Y.S., Choi, E., Kim, T. Y., Kim, M. S., Park, J. W., *A dominant-negative isoform lacking exons 11 and 12 of the human hypoxia-inducible factor-1alpha gene.* Biochem J., (2002). **362**(1): p. 71-9.
127. Tuduri, S., Crabbe, L., Conti, C., Tourriere, H., Holtgreve-Grez, H., Jauch, A., Pantesco, V., De Vos, J., Thomas, A., Theillet, C., Pommier, Y., Tazi, J., Coquelle, A., Pasero, P., *Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription.* Nat. Cell. Biol., (2009). **11**(11): p. 1315-24.
128. Czubaty, A., Girstun, A., Kowalska-Loth, B., Trzcinska, A. M., Purta, E., Winczura, A., Grajkowski, W., Staron, K., *Proteomic analysis of complexes formed by human topoisomerase I.* Biochim. Biophys. Acta., (2005). **1749**(1): p. 133-41.
129. Straub, T., Knudsen, B. R., Boege, F., *PSF/p54(nrb) stimulates "jumping" of DNA topoisomerase I between separate DNA helices.* Biochemistry., (2000). **39**(25): p. 7552-8.
130. Andersen, F.F., Tange, T. O., Sinnathamby, T., Olesen, J. R., Andersen, K. E., Westergaard, O., Kjems, J., Knudsen, B. R., *The RNA splicing factor ASF/SF2 inhibits human topoisomerase I mediated DNA relaxation.* J. Mol. Biol., (2002). **322**(4): p. 677-86.
131. Katzenberger, R.J., M.S. Marengo, and D.A. Wassarman, *ATM and ATR pathways signal alternative splicing of Drosophila TAF1 pre-mRNA in response to DNA damage.* Mol. Cell. Biol., 2006. **26**(24): p. 9256-67.
132. Duann, P., Sun, M., Lin, C.T., Zhang, H., Liu, L.F., *Plasmid linking number change induced by topoisomerase I-mediated DNA damage.* Nucleic Acids Res., (1999). **27**(14): p. 2905-11.
133. Koster, D.A., Palle, K., Bot, E.S., Bjornsti, M.A., Dekker, N.H., *Antitumour drugs impede DNA uncoiling by topoisomerase I.* Nature., (2007). **448**(7150): p. 213-7.
134. Felsenfeld, G., Clark, D., Studitsky, V., *Transcription through nucleosomes.* Biophys Chem, (2000). **86**(2-3): p. 231-7.
135. Lapidot, M., Pilpel, Y., *Genome-wide natural antisense transcription: coupling its regulation to its different regulatory mechanisms.* EMBO Rep., (2006). **7**(12): p. 1216-22.
136. Rapisarda, A., Zalek, J., Hollingshead, M., Braunschweig, T., Uranchimeg, B., Bonomi, C. A., Borgel, S. D., Carter, J. P., Hewitt, S. M., Shoemaker, R. H., Melillo, G., *Schedule-dependent inhibition of hypoxia-inducible factor-1alpha protein accumulation, angiogenesis, and tumor growth by topotecan in U251-HRE glioblastoma xenografts.* Cancer Res., (2004). **64**(19): p. 6845-8.
137. Liu, J., Kouzine, F., Nie, Z., Chung, H. J., Elisha-Feil, Z., Weber, A., Zhao, K., Levens, D., *The FUSE/FBP/FIR/TFIIH system is a molecular machine programming a pulse of c-myc expression.* EMBO J., (2006). **25**(10): p. 2119-30.

138. Covey, J.M., Jaxel, C., Kohn, K. W., Pommier, Y., *Protein-linked DNA strand breaks induced in mammalian cells by camptothecin, an inhibitor of topoisomerase I.* Cancer Res., (1989). **49**(18): p. 5016-22.
139. Ju, B.G., Lunyak, V. V., Perissi, V., Garcia-Bassets, I., Rose, D. W., Glass, C. K., Rosenfeld, M. G., *A topoisomerase IIbeta-mediated dsDNA break required for regulated transcription.* Science., (2006). **312**(5781): p. 1798-802.