Alma Mater Studiorum – Università di Bologna
in cotutela con Université Toulouse 3 - Paul Sabatier

DOTTORATO DI RICERCA IN

Biologia Cellulare e Molecolare

Ciclo XXXII

Settore Concorsuale: 05/E2 BIOLOGIA MOLECOLARE

Settore Scientifico Disciplinare: BIO/11 BIOLOGIA MOLECOLARE

TDP1 deficiency and genomic instability in non-replicating cells

Presentata da: SIMONA SALIMBENI

Coordinatore Dottorato

Prof. Giovanni Capranico

Supervisore

Prof. Giovanni Capranico

Supervisore

Dr. Olivier Sordet

Esame finale anno 2020

ABSTRACT

Spinocerebellar ataxia with axonal neuropathy (SCAN1) is a rare recessive neurodegenerative syndrome associated with cerebellar atrophy and peripheral neuropathy. It is caused by a homozygous missense mutation in the tyrosyl-DNA phosphodiesterase-1 (TDP1) gene (A1478G). This results in a substitution of histidine for arginine-493 (H493R) in the TDP1 catalytic site, leading to reduced TDP1 activity. TDP1 hydrolyzes the bond between a DNA 3'-end and a tyrosyl moiety within a trapped topoisomerase I cleavage complex (TOP1cc). TDP1 not only excises trapped TOP1ccs but also processes other 3'-end-blocking lesions, including 3'-phosphoglycolates that result from oxidation of DNA. However, how TDP1 H493R mutation promotes the SCAN1 phenotype, which is associated with the death of post-mitotic neurons, is unclear. DNA double-strand breaks (DSBs) are infrequent but among the most harmful genomic lesions. Their defective repair can induce cell death, and they have been implicated in the pathogenesis of several human diseases, including neurodegenerative syndromes. Hence, my PhD objective was to investigate whether the SCAN1 phenotype could be related to an accumulation of DSBs in non-replicating cells harboring the H493R mutation of TDP1. The only available models to study the impact of TDP1 H493R mutation were lymphoblastoid cell lines derived from SCAN1 patients compared to those of healthy individuals. Hence, we have generated models of osteosarcoma U2OS cells homozygous for TDP1 H493R or TDP1 KO employing the CRISPR-Cas9 technique. We have also generated primary lung WI38 hTERT fibroblasts TDP1 KO. We found that both TDP1 H493R and TDP1 KO cells accumulate endogenous DSBs, primarily in the G1 phase of the cell cycle compared to the S phase. A similar increase of DSBs was observed in quiescent WI38 hTERT cells following the depletion of TDP1 with siRNA, suggesting the replication-independent nature of DSBs. Treatment of TDP1 H493R and TDP1 KO cells with camptothecin to induced trapped TOP1ccs, further suggests that accumulation of DSBs could be related to the defective removal of TOP1ccs. Next, we asked whether DSB accumulation in those cells could be related to an increase in DSB production and/or a defect in their repair. Notably, R-loop structures that form co-transcriptionally can induce DSBs in non-replicating cells. We found that TDP1 deficiency modulated R-loop levels at some gene loci, raising the possibility of their implication in DSB formation. Analysis of DSB repair following camptothecin treatment revealed that both TDP1 H493R and TDP1 KO cells were defective in the repair of DSBs in G1 but not in S, with TDP1 H493R having the most pronounced effect. These results suggest that DSBs would accumulate specifically in TDP1-deficient cells that do not undergo replication, due to a defective repair of those breaks. Together, our results provide insights into the etiology of the SCAN1 neurodegenerative syndrome.

Table of Contents

1. Introduction

1.1. Topoisomerases

DNA topoisomerases are ubiquitous and essential enzymes that are present in all living systems and some viruses. They solve DNA topological entanglements that result from vital cellular processes, such as replication, transcription, recombination, and chromatin remodeling. These enzymes dissipate the helix supercoils by providing a temporary DNA break. All topoisomerases transiently nick the DNA phosphodiester bone through a transesterification reaction using a tyrosine (Tyr) residue of their active site as a nucleophile. Then, the Tyr residue remains, temporally, covalently linked to a strand end, generating an intermediate named topoisomerase cleavage complex (TOPcc). Topoisomerases are divided into two types depending on the structure and the number of DNA strand cleaved. Type I topoisomerases are all monomeric, except the *Methanopyrus kandleri* reverse gyrase (Krah *et al.*, 1996), and they cleave one strand of the DNA. While type II are multimeric proteins that cleave both strands. Additionally, they are grouped into subfamilies, A and B. Type I-A are covalently attached to the 5' phosphate (5' P-Y) and dissipate only negative supercoils passing the intact strand through the broken one. The type I-B is covalently attached to the 3' phosphate (3' P-Y) and can relax both positive and negative supercoils by rotating the nicked strand around the intact one (Figure 1) (Capranico, Marinello and Chillemi, 2017). The topoisomerases type II, allow the occurrence of the topological change by generating a double-strand break being covalently attached to the 5' ends of the cleavage site. Then, by a mechanism that requires energy, they pass a second double-strand DNA through the cleavage (Liu *et al.*, 1983; Nitiss, 2009) (Figure 2). In the human genome exist two isoforms of topoisomerase type II-A: α and β (Deweese and Osheroff, 2009). *In vitro* analysis showed that type IIα solves positive supercoils 10 times faster than negative ones, while type IIβ has no such preference (McClendon, Rodriguez and Osheroff, 2005). The first DNA topoisomerase I enzyme was discovered in 1971 in *E.Coli* and was initially named ω protein (Wang, 1971). Since then several DNA topoisomerases have been identified. The human genome encodes for six topoisomerases, type I-A (TOP3α and TOP3β), type I-B (nuclear TOP1 and mitochondrial TOP1), type II-A (TOP2α and TOP2β). In add exists a seventh topoisomerase named Spo11 (homolog to type II-B), which is limited to the germinal cells. Spo11 creates programmed double-strand breaks to initiate homologous recombination during meiosis (Bergerat *et al.*, 1997).

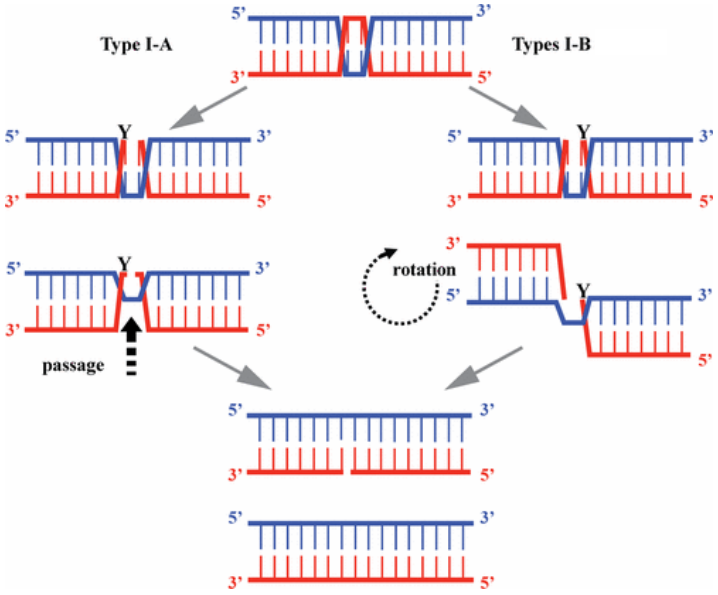

Figure 1: Catalytic mechanisms of type I DNA topoisomerases. On the left, type I-A enzymes catalyze the passage of the intact strand through the strand break; on the right type I-B enzymes remove DNA supercoils by controlling a rotation of the broken 5′ end around the intact strand. Finally, the enzyme reseals the strand cut with a reverse cleavage reaction. Adapted from (Capranico, Marinello and Chillemi, 2017).

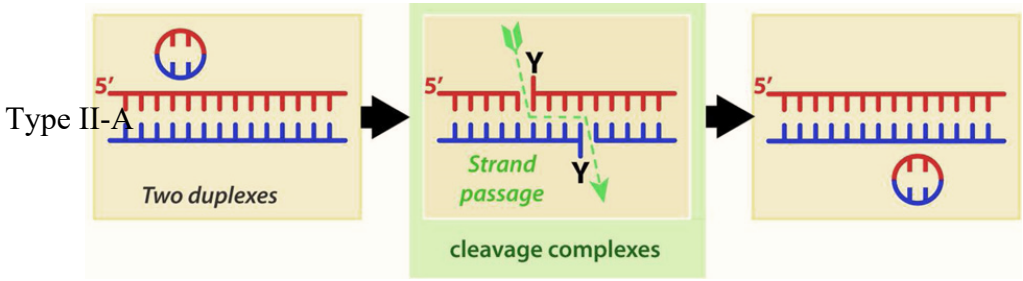

Figure 2: Catalytic mechanisms of type IIA DNA topoisomerases. Type IIA enzymes act by strand passage with double-strand break. Adapted from (Pommier, 2013).

1.1.1. **Topoisomerase Type I-B (TOP1)**

Type I-B topoisomerase is ubiquitous in eukaryotes and is present also in specific viruses, such as poxviruses (vaccinia and variola), and bacteria (Krogh and Shuman, 2002; Pommier *et al.*, 2010; Capranico, Marinello and Chillemi, 2017). The human type I-B topoisomerases are the nuclear topoisomerase 1 (TOP1) and the mitochondrial topoisomerase (mtTOP1). The nuclear TOP1 is essential for vertebrate and *Droshophila* (Lee *et al.*, 1993) but not in yeast (Uemura and Yanagida, 1984; Goto and Wang, 1985) and it is constitutively expressed during the whole

cell cycle (Baker *et al.*, 1995). TOP1 carry out his functions using only the free energy stored in DNA helix, without the need of any cofactors. It is responsible for the dissipation of both positive and negative supercoils associated with DNA replication, transcription, and chromatin condensation (Capranico, Marinello and Chillemi, 2017) (Champoux, 2001). TOP1 acts following a catalytic cycle defined "*controlled rotation*", which is characterized by four steps: *binding of the enzyme to the DNA*; *DNA cleavage*; *single-strand passage by controlled rotation*; *DNA religation* (Pommier *et al.*, 1998; Pourquier and Pommier, 2001). At first, TOP1 binds non-covalently to the DNA. This occurs with a substrate specificity related both to the DNA tertiary structure and to the sequence (Champoux, 2001). Various studies display the binding preference for the double-stranded DNA (Been and Champoux, 1984) and for a combination of nucleotides that extends from positions −4 to −1, 5'-(A/T)(G/C)(A/T)T-3' with the enzyme covalently linked to the −1 thymine residue. Sometimes, at the -1 position, a cytosine can be found (Champoux, 2001). Once bound to the DNA, TOP1 transiently cleaves one strand of the DNA duplex through a covalent 3'- phosphotyrosyl enzyme-DNA intermediate, defined as TOP1 cleavage complex (TOP1cc). This allows the controlled rotation of the broken strand around the intact one (Liu and Wang, 1987; Stewart *et al.*, 1998; Champoux, 2001; Wang, 2002; Koster *et al.*, 2005; Pommier *et al.*, 2010). Finally, once the topological change occurs, TOP1 religates the single-strand break. The latter step occurs by a nucleophilic attack of the 5' hydroxyl group (-OH) of the nicked DNA strand to the enzyme-DNA intermediate phosphotyrosine bond, followed by the rejoining of the cleaved strand (Pommier *et al.*, 1998). Under normal conditions, the religation step is favored over the cleaving step, and hence the TOP1cc is very transient and almost no detectable (Pourquier and Pommier, 2001). However, there are some factors, such as DNA modification and/or drug treatments, which can interfere with the normal relegation reaction. Precisely, these factors cause a misalignment of the 5'OH of the cleaved strand and the phosphotyrosine bond, leading to the stabilization of the TOP1cc (see section 1.1.2). The lack of correct religation can also lead to an alternative ligation, which arises with the 5'-OH of an exogenous strand. This can result in an improper DNA recombination (Pourquier, Pilon, *et al.*, 1997; Pommier *et al.*, 1998). The fundamental function of TOP1 is to resolve torsional strain generated by helix-tracking proteins, such as replication and/or transcription complexes (Pourquier and Pommier, 2001). Because the replication machinery cannot rotate freely around the DNA helix, an accumulation of positive supercoils occur upstream of the replication fork. If these supercoils are not unwound by TOP1 they can slow down, pose, or stall the replication machinery (Wang, 2002; Tuduri *et al.*, 2009). TOP1

activity is also required for transcription by dissipating the DNA supercoiling that results from the progression of the RNA Polymerase (RNAP). Precisely, positive supercoils accumulate in front of the transcriptional complex, and negative supercoils behind it (Liu and Wang, 1987; Baranello *et al.*, 2016). The accumulation of negative supercoiling upstream of the RNAP can favor the formation of R-loop, which are non-B DNA structures (see section 1.3) (Drolet, Bi and Liu, 1994; Drolet *et al.*, 2003). In human cells, the activity of a TOP1 may be involved in many other processes such as chromatin remodeling, transcription activation/repression, DNA repair, and RNA splicing (Rossi *et al.*, 1996; Straub *et al.*, 1998; Leppard and Champoux, 2005). The mtTOP1 is a nuclear-encoded topoisomerase, exclusively localized to mitochondria, that regulates mtDNA topology by relaxing negative supercoils generated by mtDNA replication and transcription. mtTOP1 loss drives to a defective mitochondrial function, an increase of oxidative radicals, and DNA damage. These last effects are features that are involved in the pathogenesis of cancer and a growing number of neurological, muscular and metabolic disorders (Sas *et al.*, 2007; Douarre *et al.*, 2012; Sobek *et al.*, 2013; Zhang *et al.*, 2014; Ghosh *et al.*, 2019).

1.1.2. **Trapping of TOP1**

Under physiological conditions, the catalytic cycle of TOP1 is very fast (~6000 cycles per minute) (Seol *et al.*, 2012) and the TOP1cc reverse rapidly and are nearly undetectable. However, the TOP1cc can be selectively trapped by drugs (TOP1 inhibitors) and under a broad range of DNA lesions (summarized in Table 1), which impair the correct alignment of the DNA ends within the TOP1cc preventing its religation (Pommier *et al.*, 2003). The reference chemical for TOP1 inhibitors is camptothecin (CPT), an alkaloid derived from the Chinese plant *Camptotheca acuminate* (Wall *et al.*, 1966). There is evidence that yeast deleted for TOP1 are immune to CPT (Eng *et al.*, 1988), and cells that present genetic modification for the TOP1 gene are resistant to this drug (Pommier *et al.*, 1999). CPT inhibits specifically the religation step of the TOP1. Once the enzyme has cleaved the DNA, CPT intercalates between the nitrogenous bases at the nick site preventing the ligation and stabilizing the normally transient TOP1cc (Figure 3) (Staker *et al.*, 2002; Pommier and Marchand, 2005). The CPT and its water-soluble derivatives, named camptothecins (CPTs), e.g. topotecan and irinotecan, are the only TOP1 inhibitors approved for clinical use. They induce DNA damage for therapeutic purposes and represent clinically important antitumor drugs (Wall *et al.*, 1966; Pommier *et al.*, 1998, 2016; Capranico, Marinello and Chillemi, 2017). However, despite their effectiveness, CTPs

are active only in lactone form and have the limitation of being readily inactivated, by the opening of the α-hydroxylactone E ring, at physiological pH (Figure 4A). To exceed this limit, non-CPT TOP1 inhibitors were investigated and developed. Among them, indenoisoquinolines and dibenzonaphthyridinones are under clinical development (Figure 4B) (Pommier, 2013; Pommier *et al.*, 2016). TOP1 can be also trapped in presence of DNA lesions such as abasic sites, uracil mismatches, nicks, and gaps when they occur in the proximity of the enzyme cleavage site (Pourquier, Pilon, *et al.*, 1997; Pourquier, Ueng, *et al.*, 1997). The localization of the DNA alteration can influence the TOP1 activity outcome. Single lesions directly downstream the cleavage site, at position +1,+2 or +3, increase the accumulation of TOP1cc, while lesions rightly upstream the cleavage site, at position -1, -2, -3, and to lesser extent abasic sites at positions -4, -5 or -6, abolish their accumulation (Pourquier and Pommier, 2001). The integrity of the DNA is daily threatened by endogenous and exogenous factors (Barnes and Lindahl, 2004). DNA is highly susceptible to reactive oxygen species (ROS) which can derive as a result of metabolic and/or biochemical reactions, such as lipid peroxidation and cellular respiration, and also from near exposure to ultraviolet light (UV) (Blount *et al.*, 1997; De Bont and van Larebeke, 2004). The most frequent oxidative damage that occurs in mammalian cells is the 7,8-dihydro-8-oxoguanine (8-oxoG) (~10^5 lesions per rat cell per day) (Park *et al.*, 1992). If it is not repaired, a miss-incorporation drives to a C:G→T:A transversion (Page *et al.*, 1995), a common mutation in human cancer. Another potential mutagenic is the 5-hydrocytosine (5-ohC), that derives from the oxidation of the pyrimidines (Feig, Sowers and Loeb, 1994; Purmal, Kow and Wallace, 1994). Pourquier and colleagues have shown that the 8-oxoG base present at the scissile +1 or +2 position, and 5-ohC incorporated at +1 position, increase the DNA cleavage mediated by TOP1, enhancing the formation of TOP1cc (Pourquier *et al.*, 1999). Then Lesher et al. in 2002, through the analysis of the crystal structure of TOP1, proposed that 8-OxoG has an impact on this enzyme by stabilizing its inactive, DNA-bound state (Lesher *et al.*, 2002). Even more frequent than oxidative lesions are the ribonucleotide incorporation into the DNA. In yeast, during physiological replication, up to one ribonucleotide is incorporated every thousand bases (McElhinny *et al.*, 2010; Williams, Lujan and Kunkel, 2016). These ribonucleotides can be removed by the RNase H2 enzyme. However, if the RNase H2 acts after that TOP1 binds the ribonucleotides, they can be transformed into nicks after TOP1 cleavage at the ribose. This arises to 2', 3'-cyclic phosphate ends, which are not suitable ends neither for the DNA polymerases extension nor for ligation by ligases, leading to mutagenic consequences as short deletions in repeat sequences (Sekiguchi and Shuman, 1997; Kim *et al.*, 2011; Huang,

Ghosh and Pommier, 2015). Moreover, genetic defects on proteins involved in the excision of the TOP1cc, as in Tyrosyl-DNA phosphodiesterase 1 (TDP1) (see section 1.2.6) (Takashima *et al.*, 2002; Interthal *et al.*, 2005; Hirano *et al.*, 2007), in cooperation with Ataxia Telangiectasia Mutated (ATM) defects (Katyal *et al.*, 2014), boost the TOP1cc long life and lead to a growing number of human pathological conditions, including cancers, neurodegenerative diseases and autoimmune syndromes (Pommier *et al.*, 2016).

Causes	Consequences for TOP1 enzymes
Anticancer drugs acting as interfacial inhibitors	Trapping of TOP1cc by irinotecan, topotecan, indenoisoquinolines* and tumour-targeting camptothecin derivatives
Oxidative DNA lesions (8-oxoguanine, 8-oxoadenosine and 5-hydroxycytosine)	Induction and trapping of TOP1cc
Abasic sites and DNA mismatches	Formation of irreversible TOP1cc
Carcinogenic base adducts (methylated bases, exocyclic adducts, benzo[*a*]pyrene adducts and crotonaldehyde adducts)	Induction and trapping of TOP1cc
Nicks and DNA strand breaks	Formation of irreversible TOP1cc, double-stranded breaks, genomic deletions and recombination
UV lesions (pyrimidine dimers and 6.4-photoproducts)	Induction of TOP1cc
Ribonucleotide incorporation into DNA	Formation of TOP1cc that generate nicks with 2′,3′-cyclic phosphate ends and short deletions in repeat sequences
Natural and food products	Unknown
Genetic defects	Unrepaired TOP1cc due to TDP1 defects in cooperation with ATM defects
Transcription activation	Stabilization of TOP1cc at enhancers

ATM, ataxia telangiectasia mutated; TDP, tyrosyl-DNA phosphodiesterase; TOPcc, topoisomerase cleavage complex.
*Indenoisoquinoline derivatives are in clinical trials.

Table 1: Endogenous and exogenous factors producing TOP1 cleavage complex. Adapted from (Pommier *et al.*, 2016).

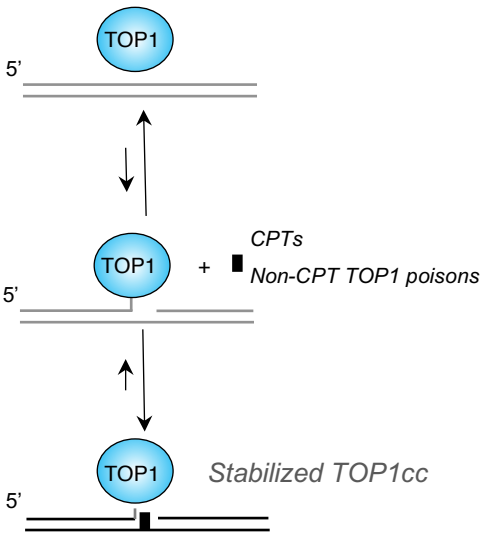

Figure 3: Mechanisms of action of TOP1 inhibitors. CPTs and non-CPTs poisons trap TOP1cc by blocking the DNA relegation. Adapted from (Pourquier and Lansiaux, 2019).

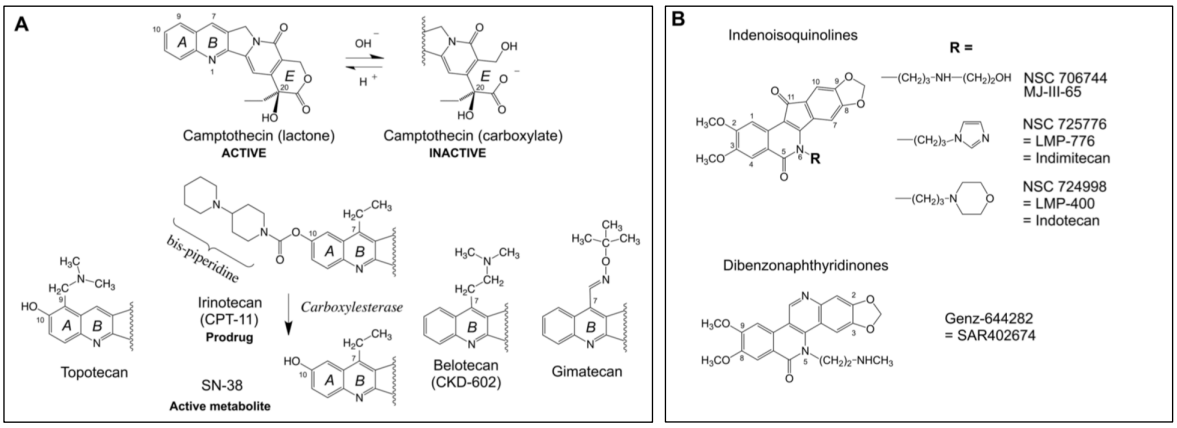

Figure 4: TOP1 Inhibitors. (A) Camptothecin and its clinical derivatives. The facile and reversible opening of the α-hydroxylactone E ring of camptothecin is shown at the top. (B) Non-camptothecins. Adapted from (Pommier, 2013).

1.1.3. **TOP1-mediated DSBs**

Stabilized TOP1cc are only potentially lethal and, if they are not resolved, become irreversible TOP1cc. If a collision between irreversible TOP1cc and the replication forks and/or the transcription machinery occurs, they can be further converted into a lethal lesion, more permanent single-strand breaks (SSBs) or double-strand breaks (DSBs) (reviewed in (Pourquier and Pommier, 2001; Pommier *et al.*, 2006, 2014, 2016)), (Pommier *et al.*, 2003; Ljungman and Lane, 2004; Capranico *et al.*, 2007; Sordet *et al.*, 2009; Cristini *et al.*, 2019). Single-strand breaks (SSBs) are breaks occurring in one strand of the DNA double helix and are commonly associated by loss of a single nucleotide and 5'- or 3'-end damage at the break site (Caldecott, 2008). The CPT itself is relatively non-cytotoxic in short exposures (< 1h) because it generates transient stabilized-TOP1cc, defined reversible TOP1cc. So, the cytotoxicity of CPT is due to the conversion of the reversible TOP1cc into irreversible and DNA damage once arises the collision with the replication/transcription apparatus (Pommier, 2006). The probability to induce lethal lesions because of CPT is correlated to the time and the concentration of treatments and also to the cell type (cell proliferation, DNA replication, and transcription rate) (Holm *et al.*, 1989; Hsiang, Lihou and Liu, 1989; Huang *et al.*, 2010).

In proliferating cells, the main cytotoxic mechanism of TOP1 inhibitors is the production of replication-coupled DSBs (RC-DSBs) (Pommier, 2006). These lasts can arise by two different mechanisms: the "replication run-off" (Figure 5A) (Strumberg *et al.*, 2000) and the Mus81-Eme1 cleavage (Figure 5B) (Regairaz *et al.*, 2011). The first mechanism drives to the formation of 5' phosphorylated blunt-ended DSBs as a result of the extension of the leading strand up to the TOP1cc at the 5'end. Because the reversibility of the TOP1cc is allowed by the 5'OH terminus, the 5' phosphorylation avoids the religation by TOP1 itself (Strumberg *et al.*, 2000). Concerning the second mechanism, the RC-DSBs depend on the 3' flap endonuclease Mu81-Eme1 activity. The latter cleaves at the level of the stalled replication fork, generating DSBs which further let the dissipation of positive supercoils accumulated because of CPT-induced TOP1 inhibition. This pathway helps the replication fork recovery and cell survival (Regairaz *et al.*, 2011). Likewise, to replication, CPT-stabilized TOP1cc becomes irreversible TOP1cc if occur the collision with the transcriptional machinery (Wu and Liu, 1997; Pommier, 2006). This collision further leads to the production of TOP1-linked SSBs (Hsiang *et al.*, 1985), for review see (Ashour, Atteya and El-Khamisy, 2015), and transcription-coupled DSBs (TC-DSBs) (Sordet *et al.*, 2009, 2010; Cristini *et al.*, 2016, 2019). *Sordet et al.* primarily described

the production of TC-DSBs in post-mitotic primary neurons and lymphocyte and cells outside of the S-phase (cells lacking 5-ethynyl-2′-deoxyuridine, EdU, incorporation into the DNA) (Sordet *et al.*, 2009, 2010) and then was confirmed by assorted studies in different cellular models (Das *et al.*, 2009; Huang *et al.*, 2010; Regairaz *et al.*, 2011; Zhang *et al.*, 2011; Sakai *et al.*, 2012; Katyal *et al.*, 2014; Cristini *et al.*, 2016, 2019). Besides, these co-transcriptional DSBs production involves the formation of R-loops (see section 1.3.5) (Aguilera and García-Muse, 2012; Sollier *et al.*, 2014). Trapped TOP1cc are potent transcription-blocking DNA lesions (Pommier, 2006; Capranico *et al.*, 2007), and have also physiological importance, as in neurons, where there are a high oxidative metabolism and high level of transcription (Powell et al., 2013, Huang et al., 2011). Their removal is important to protect cells and their genetic stability, and its relay on two alternative pathways: the TDP1 excision pathway and the endonucleases pathway (see section 1.2.2) (Pommier *et al.*, 2003, 2006; Ashour, Atteya and El-Khamisy, 2015; Xu and Her, 2015).

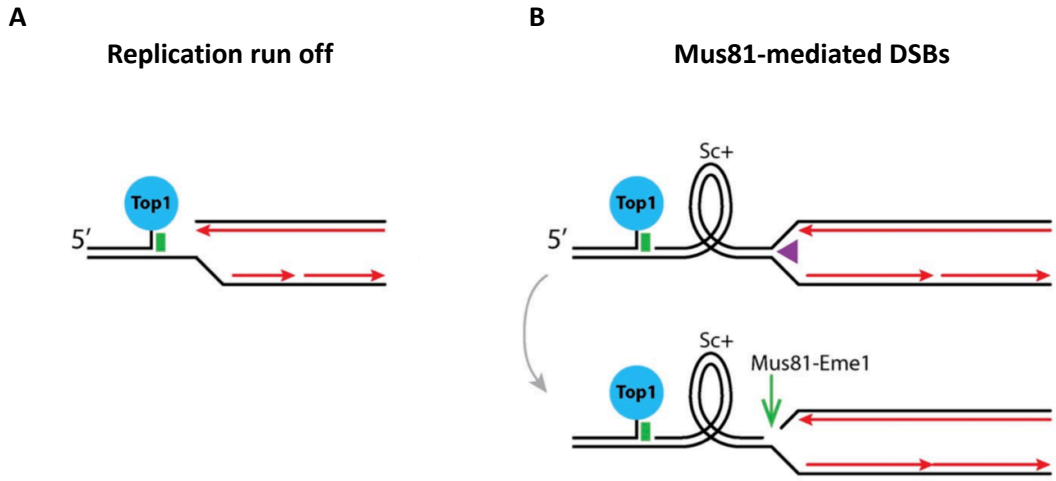

Figure 5: Mechanisms of production of replication coupled DSBs (RC-DSBs) by TOP1cc. (A) Replication run off. (B) Mus81-Eme1 cleavage of the stalled replication fork. Adapted from (Regairaz *et al.*, 2011).

Cells have a molecular mechanism developed to detect and repair DNA damage, named DNA damage response (DDR). A key role in this response is the phosphorylation of the histone H2AX at Ser 139 (γH2AX) and its accumulation in nuclear foci. A single γH2AX focus mirrors hundreds to thousands of γH2AX proteins that are localized around at least one DSB (William M. Bonner *et al.*, 2008). This is a ubiquitous response to DSBs arising in replicating (Furuta *et al.*, 2003) or non-replicating DNA (Redon *et al.*, 2002). γH2AX is a very sensitive marker for DSBs and can be detected by immunofluorescence (Redon *et al.*, 2002) or immunostaining

(Gorgoulis *et al.*, 2005). In addition to γH2AX, the DSBs can be visualized as nuclear foci containing phosphorylated 53BP1 at S1778 (p54BP1) and by neutral comet assays, which gives as readout the increase in the comet tail moment.

The two main mechanisms of DSBs repair, in eukaryotic cells, are the Non-Homologous End Joining (NHEJ) and the Homologous recombination (HR) (Figure 6 A-B).

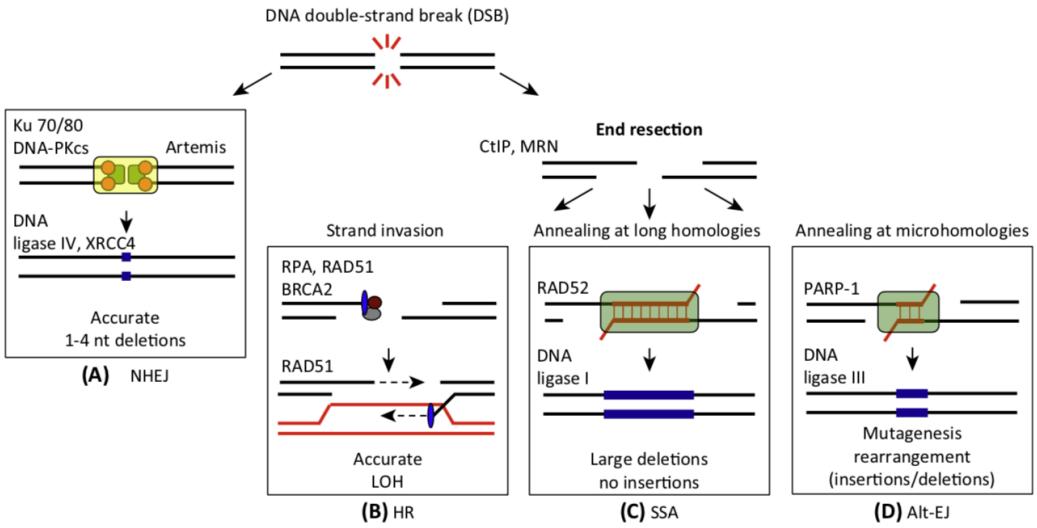

Figure 6: DSBs repair mechanisms. (A) NHEJ; (B) HR; (C) SSA; (D) Alt-EJ. Adapted from (Ceccaldi, Rondinelli and D'Andrea, 2016).

NHEJ can occur during all the cell cycle but is predominant in G0/G1 and G2 (Karanam *et al.*, 2012; Chiruvella, Liang and Wilson, 2013). It is an error-prone mechanism that allows the repair of the DSBs by direct ligation of the blunt broken ends. At first, NHEJ involves the DNA-PK complex (composed by Ku80-Ku70 heterodimer and DNA-PKcs) which recognize the broken ends and Artemis, a structure-specific nuclease, and other end-processing factors needed for the ligation by the ligation complex (Cottarel *et al.*, 2013). The latter includes DNA ligase IV (LigIV), the X-ray cross-complementing group 4 (XRCC4), and XRCC4-like factor (Ochi *et al.*, 2015). Because HR uses a sister or homologous chromatid for the repair, it is generally restricted to cells in S- and G2-phases (Karanam *et al.*, 2012). To take place HR needs a resection of the DNA ends (Wyman and Kanaar, 2006; Hartlerode and Scully, 2010; Grabarz *et al.*, 2012). This resection is enabled by the endonuclease activity of the MRE11–RAD50–NBS1 (MRN) complex and the 5′–3′ strand resection activities of exonuclease 1 (EXO1) and the DNA2–Bloom syndrome protein (BLM) heterodimer, which together converts the blunt DSB end into a 3′ -ssDNA. The MRE11, through its endonuclease activity, nicks the strand to be resected up to 300 nucleotides from the 5'-terminus of the DSB. MRE11 endonuclease

activity requires the interplay with CtBP-interacting protein (CtIP) and is stimulated by protein blocking the DNA end, such as Ku70–Ku80. Then, thanks to its 3′–5′ exonuclease activity, MRE11 extends the nick towards the DNA end. This step allows the displacement of Ku70–Ku80 from the DNA ends. The nick provided by MRE11 enables also the resection in the 5'–3' direction away from the DSB. This latter step is mediated by EXO1, the endonuclease DNA2, and the helicase BLM, which mediate the unwinding and nucleolytic digestion of the 5′ strand of the DNA end to form long 3′-ssDNA overhangs. Then, the resulting ssDNA is immediately coated by the RPA complex. There are other DNA end resection regulators that have been determined. For instance, breast cancer type 1 susceptibility protein (BRCA1), in complex with BRCA1-associated RING domain protein 1 (BARD1), which interacts with CtIP and MRN. The other crucial step in HR is the strand invasion mediated by the recombinase RAD51. This is mediated by the displacement of RPA with the help of mediator proteins (including BRCA1 and BRCA2), searching for homology and allowing an error-free DNA repair. There are also two alternatives error-prone pathway that can play in the DSBs repair: the alternative end-joining (alt-EJ) and the single-strand annealing (SSA) (Figure 5 C, D) (for review see (Van Gent, Hoeijmakers and Kanaar, 2001; Wyman and Kanaar, 2006; Hartlerode and Scully, 2010; Grabarz *et al.*, 2012; Ceccaldi, Rondinelli and D'Andrea, 2016; Piazza and Heyer, 2019)). The object of studies is the choice of DSB repair pathway and the molecular mechanism regulating this choice (reviewed in (Chapman, Taylor and Boulton, 2012)). The major factor is that the repair of DNA DSBs relies on whether DNA end resection occurs. Thus, when the resection is blocked, NHEJ is favored. While, when the DNA resection occurs, three pathways (HR, alt-EJ, and SSA) can compete for the repair of DSBs. The regulation of DNA end resection through the cell cycle in part explains how HR occurs only in S phase and G2 phase. A DSB present during DNA replication (S phase) is more efficiently processed than during G2, and this is most likely related to a raised resection observed in cycling compared with G2 cells (Zierhut and Diffley, 2008). During the S phase cell cycle-dependent kinase (CDK) activity is increased while in the G1 phase is decreased. It provides activating signals to the resection machinery and to proteins that play in HR. Another factor that influences the choice of one pathway over another is the chromatin status (Beucher *et al.*, 2009; Lemaître and Soutoglou, 2014). Moreover, the tumor suppressors p53-binding protein 1 (53BP1) and Breast cancer type 1 susceptibility protein (BRCA1) are two factors that are enriched at DSB sites and are crucial regulators of DSB repair fate by NHEJ and HR, respectively (reviewed in (Symington and Gautier, 2011; Chapman, Taylor and Boulton, 2012; Scully *et al*., 2019)).

Defects in cellular DNA repair mechanisms have been linked to genome instability, carcinogenesis, premature aging syndromes, and neurological disorders (Burma, Chen and Chen, 2006; Rass, Ahel and West, 2007). Several neurodegenerative diseases are caused by mutations in genes involved in nucleic acid homeostasis, summarized in Table 2. They are diseases caused by defects in proteins having critical roles in the DNA-damage response and/or involved in the repair of DSBs. These defects and the non-proliferative nature of neuronal cells may make these cells prone to progressive accumulation of unrepaired DNA lesions which may cause neuronal cell death.

Syndrome	Disease gene	Neurological implications	Other
Nucleotide excision repair			
Xeroderma pigmentosum (XP)	*XPA, XPB, XPD, XPF, XPG, XPV*	microcephaly, progressive neurodegeneration	neoplasm of the skin and eyes
Cockayne syndrome (CS)	*XPB, XPD, XPG, CSA, CSB*	microcephaly, progressive neurodegeneration	dwarfism
Trichothiodystrophy (TTD)	*XPB, XPD, TFB5/TTD-A*	microcephaly	brittle hair
DNA damage response/DSB repair			
Ataxia telangiectasia (A-T)	*ATM*	ataxia, progressive neurodegeneration	immunological implications, lymphoid malignancy
Ataxia telangiectasia-like disorder (ATLD)	*MRE11*	ataxia, progressive neurodegeneration	immunological implications, lymphoid malignancy
Nijmegen breakage syndrome (NBS)	*NBS1*	microcephaly	immunological implications, lymphoid malignancy, short stature
ATR-Seckel syndrome (ATR-Seckel)	*ATR*	microcephaly	dwarfism
Primary microcephaly 1 (MCPH1)	*MCPH1/BRIT1*	microcephaly	
LIG4 syndrome	*LIG4*	microcephaly	immunodeficiency, lymphoid malignancy, developmental and growth delay
Immunodeficiency with microcephaly	*Cernunnos/XLF*	microcephaly	immunodeficiency, lymphoid malignancy
SSB repair			
Ataxia with oculomotor apraxia 1 (AOA1)	*Aprataxin*	ataxia, progressive neurodegeneration, oculomotor apraxia	hypoalbuminemia, hypercholesterolemia
Spinocerebellar ataxia with axonal neuropathy (SCAN1)	*Tyrosyl-DNA phosphodiesterase 1*	ataxia, sensory loss	hypoalbuminemia, hypercholesterolemia

Syndromes are classified according to the defects in nucleotide excision repair, the DNA-damage response/DNA double-strand break repair, and DNA single-strand break repair. Note that defects in single-strand break repair primarily affect the nervous system. The indicated clinical implications are intended to give an overview and are not comprehensive.

Table 2: DNA-repair deficiency and neurodegeneration. Adapted from (Rass, Ahel and West, 2007).

1.2. Tyrosyl-DNA phosphodiesterase 1 (TDP1)

1.2.1. Discovery, structure and catalytic mechanism of reaction of TDP1

The tyrosyl-DNA phosphodiesterase 1 (TDP1) gene was identified in *S. Cerevisiae* by Nash and colleagues in 1996 (Yang *et al.*, 1996). It encodes for an enzyme, functionally conserved from yeast to humans, which catalyzes the hydrolysis of a phosphodiester bond between a tyrosine residue and a DNA 3'- phosphate (Povirk, 1996; Pouliot *et al.*, 1999; Interthal, Pouliot and Champoux, 2001). Sequence comparisons, mutational analyzes, and crystal structure

determination have shown that TDP1 is a member of the phospholipase D (PLD) superfamily. This latter is a highly diverse group of proteins that catalyze phosphoryl transfer reactions (Interthal, Pouliot and Champoux, 2001; Davies *et al.*, 2002b). Despite a relatively low level of sequence conservation between yeast and human TDP1, they share similar overall architectures and analogous 3'-phosphotyrosyl processing activity (Interthal, Pouliot and Champoux, 2001; Cheng *et al.*, 2002; He *et al.*, 2007). Human TDP1 is a 68-kDa protein ubiquitously expressed. It is predominantly nuclear and partially localized in mitochondria, where it removes TOP1mtcc (Fam, Chowdhury, *et al.*, 2013; Huang and Pommier, 2019). It comprises 608 amino acid residues (Figure 7) and is organized in two domains: the N-terminus domain and the C-terminus domain (Interthal, Pouliot and Champoux, 2001; Pommier *et al.*, 2014). The N-terminus domain (amino acids 1-148) regulates TDP1 recruitment to DNA damage sites and its stability as a result of post-translational modifications. The C-terminus domain is important for TDP1 catalytic activity as it contains two catalytic Histidine-Lysine-Asparagine (HKN) motifs, spaced by 210 amino acid residues. The first HKN motif (amino acids 262-289) and the second HKN motif (amino acids 492-522) are the most conserved regions of the protein. The singularity of human TDP1 and its orthologs is the substitution of the aspartates (D) of the characteristic HKD catalytic motifs found in the other members of the PLD superfamily by asparagines (N) (HKN motifs) (Interthal, Pouliot and Champoux, 2001; Pommier *et al.*, 2014). The histidine (H) and lysine (K) residues of these motifs are necessary for TDP1 catalytic activity (Stuckey and Dixon, 1999; Ponting and Kerr, 2008). Mutations of histidine (H263A) or lysine (K265S) in the first HKN motif lead to the complete loss of TDP1 activity, while mutations of histidine (H493A/R/N) or lysine (K495S) in the second HKN motif cause a strong decrease in TDP1 activity (Interthal, Pouliot and Champoux, 2001; Raymond *et al.*, 2004).

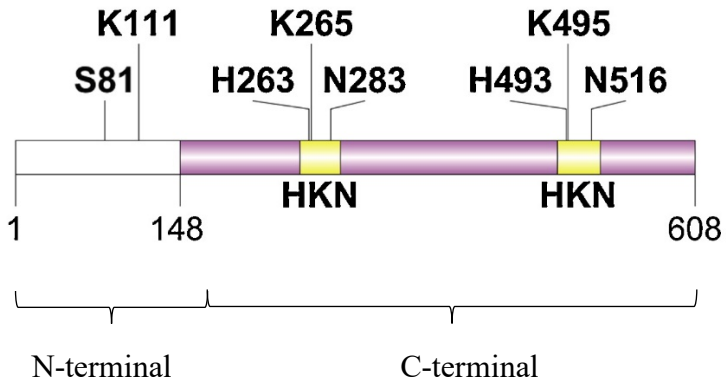

Figure 7: Representative schema of TDP1. The N-terminal is represented in white (residues 1–148) and the site of phosphorylation, at serine 81 (S81), and SUMOylation, at lysine 111 (K111) are indicated. The C-terminal in pink (residues 149-608), present the conserved catalytic HKN motives in yellow. In the first motive (residues 262-284), H263, K265, N283 are indicated; the second motive (residues 492-517) H493, K495, N516 are indicated. Adapted from (Pommier *et al.*, 2014).

X-ray crystallographic analysis showed that human TDP1 is a monomer characterized by an α–β–α–β–α sandwich tertiary structure composed of two α–β–α domains sharing similar topology (A). The active site, constituted by the two closeness HKN motifs, is located along a pseudo-2-fold axis of symmetry between the two domains. It creates an asymmetric "binding-channel" both in shape and surface charge distribution (Figure 8 B). This "binding-channel", above of the active site, is narrow and positively charged to bind the single-stranded DNA substrate. While, below the active site, it is loose to form a bowl-shaped basin negatively charged allowing the binding to a protein substrate (Davies *et al.*, 2002a, 2002b, 2003). Thanks to these features, TDP1 can excise a wide range of 3' - blocking DNA (or RNA) lesions, resolving them into 3′-phosphate end products. This occurs by a two step-reaction with a covalent intermediate. This reaction occurs without the cooperation of nucleotide cofactors or metals (Gottlin *et al.*, 1998; Interthal, Pouliot and Champoux, 2001; Pommier *et al.*, 2014).

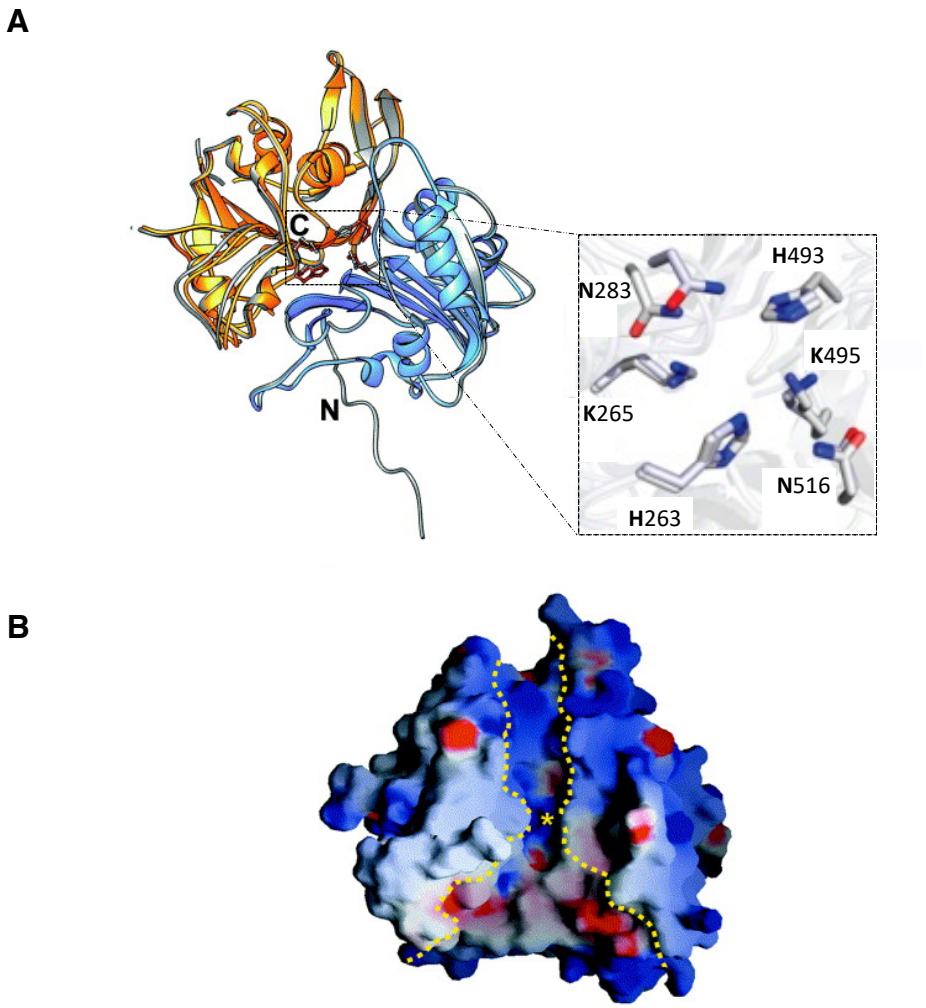

Figure 8: Overall structure of TDP1. (A) N-terminal domain (residues 162–350) colored blue and the C-terminal domain (residues 351–608) colored yellow. The active site residues His263, Lys265, His493, and Lys495 are shown as ball-and-stick structures. (B)The yellow asterisk is located at the center of the active site. Yellow dotted lines represent the sides of the substrate binding cleft. The narrow, positively charged cleft is above the active site in this figure and the wider, mixed charge binding cleft is below the active site. Adapted from (Davies *et al.*, 2002a).

1.2.2. **TDP1 excision pathway**

The different nature of the TOP1cc (transcription-mediated-TOP1cc or replication-mediated TOP1cc) of the DNA lesion and the distinct cellular background (post-mitotic or replicating cells) influence the choice of the pathway for the repair of the trapped TOP1cc. As previously mentioned, the two alternative pathways are the TDP1 excision pathway and the endonucleases pathway (Figure 9) (Pommier *et al.*, 2003, 2006; Ashour, Atteya and El-Khamisy, 2015; Xu and Her, 2015). We focused our attention on the TDP1 excision pathway (Figure 9A), which is

considered the main pathway for the resolution of the transcription-mediated TOP1cc (El-Khamisy *et al.*, 2005; Miao *et al.*, 2006; Hudson *et al.*, 2012). The endonuclease pathway (Figure 9B) is preferentially involved in the repair of replication-mediated TOP1cc. Concerning this latter, it is unclear if a previous TOP1 proteasome degradation is needed. It requires multiple 3'-flap endonucleases complexes that were initially identified in Saccharomyces cerevisiae thanks to CPT-induced DSBs repair studies in absence of TDP1(Vance and Wilson, 2002; Deng *et al.*, 2005; Pommier *et al.*, 2006). These multiples endonucleases excise the DNA on the scissile strand to release the DNA-TOP1 complex. The removal occurs at a few nucleotides away from the 3'-Phosphate (P)-TOP1 end and the resulting DNA lesions are then processed by Homologous Recombination (HR) or by Non-Homologous End Joining (NHEJ) (Pommier *et al.*, 2006; Xu and Her, 2015).

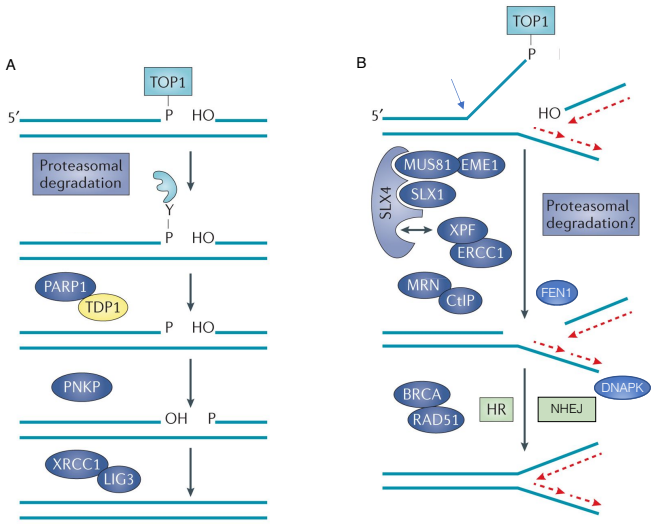

Figure 9: Representation of the main pathways for TOP1cc repair. (A) TDP1 excision pathway. (B) Endonuclease pathway. Adapted from (Ashour, Atteya and El-Khamisy, 2015).

TDP1 function is straightly correlated to the excision of the TOP1 cleavage complexes (Yang *et al.*, 1996; Pouliot *et al.*, 1999; Pommier *et al.*, 2006). The efficiency of TDP1 to hydrolyze covalently bound tyrosine from the 3' end of DNA, without excising it neither of a single base, is critical in the repair of the TOP1cc (Yang *et al.*, 1996; Pouliot *et al.*, 1999; Interthal, Chen and Champoux, 2005). Furthermore, there is evidence that TDP1 enzymatic activity is more efficient for substrates containing long nucleotides and short peptides. Thus, the native full-length TOP1 peptide needs to be previously partially proteolyzed or denatured to be cleaved by TDP1 (Debethune, 2002; Interthal, Chen and Champoux, 2005; Interthal and Champoux,

2011). As previously anticipated, the hydrolysis of the 3'-TP occurs by a two-step phosphoryl transfer reaction through a covalent TDP1-DNA intermediate formation (Gottlin *et al.*, 1998). The first step consists of a nucleophilic attack of the TOP1-DNA phosphotyrosyl bond. It is initiated by the catalytic histidine H263 of the first HKN motif with the consequent release of the tyrosine-containing peptide derived from TOP1 and the formation of a TDP1-DNA covalent intermediate. The generation of this intermediate is allowed by the fact that the H493 of the second HKN motif acts as a general acid and gives a proton to the tyrosine-containing peptide leaving group. The second step is the resolution of the TDP1-DNA intermediate by a hydrolysis reaction. This is driven by H493, which acts as a general base activating a water molecule. This event consequently leads to the disengagement of TDP1 and the generation of a 3'-phosphate end (Gottlin *et al.*, 1998; Pommier *et al.*, 2014; Kawale and Povirk, 2018). Subsequently, the 3'-phosphate end is converted by polynucleotide kinase 3'-phosphatase (PNKP) into 3'-hydroxyl and 5′-phosphate termini, which are suitable for gap-filling and ligation by polymerase β (Pol β) and DNA ligase 3α (Lig3α), respectively. TDP1, PNKP, and Lig3α are part of the X-ray repair cross-complementing 1 (XRCC1) complex (Figure 9) (Pommier *et al.*, 2014; Kawale and Povirk, 2018). As previously cited, stabilized TOP1cc are potent transcription-blocking DNA lesions (reviewed in (Pommier et al., 2006, 2014, 2016)) and their removal depends primarily on TDP1 (Pouliot et al., 1999; Capranico et al., 2007). The repair of TOP1cc by TDP1 and the further recovery of RNA synthesis have been related to TOP1 downregulation (Debethune, 2002; Desai *et al.*, 2003; Lin *et al.*, 2008). Moreover, mutations in PNKP gene lead to the autosomal recessive neurodevelopment disease denoted MCSZ (Shen *et al.*, 2010), and a homozygous single point mutation in TDP1 gene causes the neurodegenerative syndrome, spinocerebellar ataxia with axonal neuropathy (SCAN1), characterized by the death of post-mitotic neurons (see section 1.2.6) (Takashima *et al.*, 2002). Both PNKP defective cells and SCAN1 cells accumulate comparable levels of CPT-induced DNA damage (El-Khamisy *et al.*, 2005). Afterward, Miao *et al.* showed that the transcription inhibitor DRB (5,6-dichloro-1-D-ribofuranosylbenzimidazole), decreases DNA-protein crosslinks (DPC), which represents the DNA part of TOP1cc, in CPT-treated SCAN1 cells. While no decrease was detected in the presence of the replication inhibitor, aphidicolin (APH). This indicates that SCAN1 cells are defective for the repair of transcription-induced TOP1cc (Miao *et al.*, 2006).

1.2.3. **TDP1 post transcriptional modification**

The N-terminal domain of TDP1 is subject to post-translational modifications that regulate the activity and the stability of the protein.

Phosphorylation: TDP1 can be phosphorylated at serine 81 (Figure 7) by ataxia telangiectasia mutated (ATM) and DNA-dependent protein kinase (DNA-PK) (Das *et al.*, 2009). ATM is a fundamental transducer of the DSB response (Shiloh, 2006) and DNA-PK is the main player in the NHEJ pathway of DSB repair (Weterings and Chen, 2007). This phosphorylation does not affect TDP1 enzymatic activity but rather increases TDP1 stability, and regulates its subcellular distribution (Das *et al.*, 2009; Chiang, Carroll and El-Khamisy, 2010). The latter was followed using the pS81-TDP1 antibody in confocal immunofluorescence microscopy. It was detected that TDP1 phosphorylation at S81 is associated with the focal accumulation of TDP1 at DNA damage sites. In response to CPT and IR-induced DNA damage, phosphorylated S81-TDP1 foci co-localize with γH2AX and with XRCC1 foci (Das *et al.*, 2009). And this phosphorylation increases the binding between TDP1 and Lig3α and XRCC1, facilitating the repair of TOP1-induced DSBs, protecting cells against CPT- and IR- induced DNA damage (Das *et al.*, 2009; Chiang, Carroll and El-Khamisy, 2010).

Sumoylation*:* Another post-transcriptional modification is the SUMOylation, which implies that TDP1 interacts with Ubc9 (SUMO E2 conjugation enzyme). It has no influence on the catalytic activity of the enzyme and promotes its accumulation at the DNA damage site, to promptly repair the SSBs (Hudson *et al.*, 2012). This accumulation is partially transcription-dependent, underlining the importance of TDP1 in protecting not cycling cells from transcription-induced TOP1cc damage (Hudson *et al.*, 2012; Das *et al.*, 2014).

Parylation: TDP1 undergoes the PARylation by Poly (ADP-ribose) polymerase-1 (PARP1) at unknown lysine residues, which does not influence TDP1 activity (Li *et al.*, 2013; Das *et al.*, 2014). PARP1 is a ubiquitous chromatin-associated enzyme, which catalyzes the polymerization of ADP-ribose moieties derived from nicotinamide adenine dinucleotide (NAD+) onto several substrates, including itself. This protein modification, named PARylation, has critical roles in the regulation of DNA repair (Ray Chaudhuri and Nussenzweig, 2017). In 2014, Das et al. found that human TDP1 detects TOP1-induced DNA damage through PARP1 (Das *et al.*, 2014). Once PARP1 detect and bound the TOP1cc-DNA lesion (Bowman *et al.*, 2001; Zhang *et al.*, 2011; Patel *et al.*, 2012; Das *et al.*, 2014), PARylates itself and TDP1.

PARP1, through its C-terminus, directly interacts with the N-terminal of TDP1. This interaction and modification not only stabilize TDP1 but also enhances its recruitment to TOP1cc. Then the PARP1-TDP1 complex recruits the X-ray repair cross-complementing protein 1 (XRCC1) and PARP1 is released from the DNA, which is easily repaired (Das *et al.*, 2014). XRCC1 is a protein scaffold that forms a multimeric repair complex, stabilizing, and stimulating different components of the single-strand break repair (SSBR). It interacts directly with PARP1 (Das *et al.*, 2014), PNKP, Pol β and Lig3α (Whitehouse *et al.*, 2001; Caldecott, 2008) and indirectly, through the Lig3α, with TDP1 (Das et al., 2014, El-Khamisy et al., 2005, Plo et al., 2003). DNA ligation is the final step of the single-strand breaks (SSBs) repair. Moreover, experiments carried out on Rhabdomyosarcoma cells knock-down (KD) for TDP1 and treated with PARP1-inhibitor (Fam, Walton, *et al.*, 2013) and on DT40 cells KO for TDP1 co-treated with a PARP1 inhibitor and CPT. These experiments have shown that PARP1 and TDP1 have an epistatic function (Brettrager, Segura and van Waardenburg, 2019). This latter term addresses the molecular interactions between proteins, indicating whether they operate within the same pathway or directly complex with each other (Phillips, 2008).

1.2.4. **Other substrates of TDP1**

TDP1 is able to hydrolyze a wide variety of phosphodiester linked 3'- and 5'- DNA adducts (Figure 10) (Pommier *et al.*, 2014). Nitiss and coworkers have shown that yeast TDP1 is able to resolve 5'-phosphotyrosyl linkage (Nitiss *et al.*, 2006), a feature that has been characterized also *in vitro* for human TDP1 (Murai *et al.*, 2012). However, this ability is less efficient than Tyrosyl-DNA phosphodiesterase 2 (TDP2), suggesting that TDP1 takes action only as in a back-up pathway for TOP2cc repair (Nitiss *et al.*, 2006; Murai *et al.*, 2012). TDP1 excises 3'-DNA adducts with a preference for ssDNA compared to dsDNA and retains activity especially for 3′-overhanging- and blunt-ended dsDNA (Yang *et al.*, 1996; Debethune, 2002; Raymond, Staker and Burgin, 2005). TDP1 can cleave efficiently a phosphoamide bond, which characterizes the covalent TDP1-DNA intermediate linkage. It can also remove an intact TDP1 molecule from the DNA *in vitro* (Interthal, Chen and Champoux, 2005). TDP1 is associated with the SSBR, which is part of base excision repair (BER) (Plo *et al.*, 2003; El-Khamisy, Hartsuiker and Caldecott, 2007). The role of TDP1 in this process was primarily studied by Caldecott and colleagues, who showed that it repairs the SSBs caused by aberrant TOP1 activity or oxidative stress (Caldecott, 2008). TDP1 plays a role in the repair of 3'- DNA lesions caused by base alkylation: 3′-abasic and 3′-deoxyribose phosphate end (Interthal, Chen and Champoux,

2005; Murai *et al.*, 2012; Alagoz, Wells and El-Khamisy, 2014) or oxidation: 3'-phosphoglycolates (PG) (Inamdar *et al.*, 2002; Interthal, Chen and Champoux, 2005; Zhou *et al.*, 2005). The 3' apurinic/apyrimidinic (AP) abasic sites are mutagenic, cytotoxic, and frequent DNA lesions (10,000 AP sites per mammalian cell per day). They are spontaneously produced by the N-glycosylic bond cleavage between the deoxyribose and the base, or during base excision repair (BER) pathway, or because alkylating agents, such as methyl methane sulfonate (MMS) and temozolomide (Lebedeva, Rechkunova and Lavrik, 2011; Murai *et al.*, 2012; Alagoz, Wells and El-Khamisy, 2014). PG are generated by free radical-mediated DNA cleavage (oxidative fragmentation of DNA sugars), radiomimetic drugs or ionizing radiation (Henner, Grunberg and Haseltine, 1983; Povirk, 1996; Inamdar *et al.*, 2002; El-Khamisy, Hartsuiker and Caldecott, 2007). Moreover, it seems that TDP1 participates in the same pathway in mitochondria, where it translocates in response to ROS (Fam *et al.*, 2018) and it repairs mtDNA from TOP1mtcc (Chiang *et al.*, 2017) and damages. To corroborate this, TDP1 was also found to form a complex with Lig3α in this organelle (Prakash and Doublié, 2015; Fam *et al*., 2018). In addition, the role of TDP1 in SSBs repair is also confirmed by the fact that cells deficient for TDP1 show a sensitivity to agents that causes DNA single-strand breaks. For instance, SCAN1 cells are deficient in IR-induced SSBs repair (El-Khamisy, Hartsuiker and Caldecott, 2007) and DT40 chicken cells depleted for TDP1 are sensitive to the alkylating agent, MMS. The same effect has been shown in human cells KD for TDP1 (Murai *et al*., 2012). It has been also shown that TDP1 mutants, in quiescent/G0 *Schizosaccharomyces pombe*, accumulate oxidative DNA damage in a TOP1 independent manner (Ben Hassine and Arcangioli, 2009). The nucleosidase activity of TDP1 can also remove, less efficiently, 3′-terminal deoxyribo- and ribonucleotides when they are not phosphorylated at their 3′-end (Interthal, Chen and Champoux, 2005). In addition, TDP1 is capable to excise 3'-synthetic DNA adducts, as fluorescent substrates and biotin. These features have been used to carry out kinetic analysis of TDP1 enzymatic activity and for the screening of TDP1 inhibitory molecules (Rideout, Raymond and Burgin, 2004; Antony *et al.*, 2007; Dexheimer *et al.*, 2010).

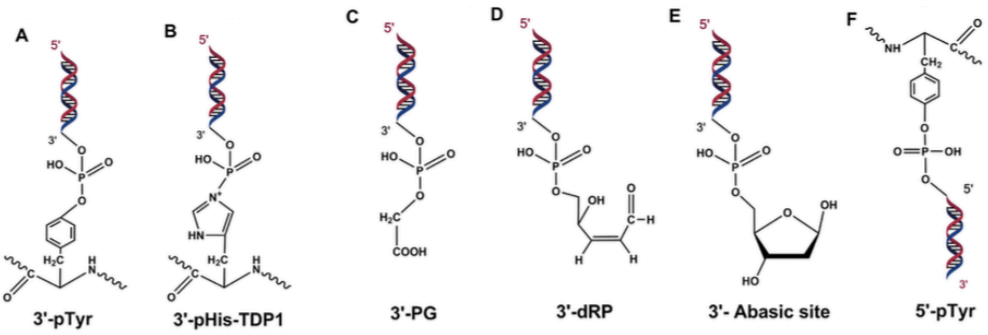

Figure 10: TDP1 substrates. 3′-phosphotyrosyl peptide - canonical substrate of TDP1. (B) 3′-phosphoamide adducts formed by H493R-mutant TDP1 causing covalent linkage of H263-TDP1 to 3′-DNA end. (C) 3′-phosphoglycolates. (D) 3′-deoxyribose phosphate and (E) 3′-abasic sites produced because of oxidative DNA damage. (F) 5′-phosphotyrosyl peptide. Adapted from (Kawale and Povirk, 2018).

1.2.5. **The role of TDP1 in Double-Strand Break (DSB) repair**

DNA double-strand breaks (DSBs) are the most dangerous and severe lesions. They are characterized by the disruption of the Watson-Crick base pairing and the separation of the two DNA ends without any complementary strand which can be used as a template for the repair. For this reason, the two broken extremities might rejoin determining the loss or amplification of a portion of DNA. A correct repair is essential to avoid mutagenesis, carcinogenesis, neurodegeneration, and aging (Lindahl and Barnes, 2000; Khanna and Jackson, 2001; De Bont and van Larebeke, 2004; Jackson and Bartek, 2009). There is evidence that TDP1 is also involved in DSB repair. As previously cited, TDP1 can excise an ample variety of substrates including 3'-PG terminated DNA DSBs (Zhou *et al*., 2009). In add, TDP1 knock out (KO) mice and TDP1 (KO) DT40 chicken cells are hypersensitive to bleomycin (Hirano *et al*., 2007; Murai *et al*., 2012), antitumor antibiotics that cause specific DSBs (Povirk, 1996). Belonging to the same class of antibiotics is the calicheamicin, which produces DSBs with 3′-PG (Povirk, 1996). SCAN1 cells are reported to be sensitive to calicheamicin, which causes chromosomal aberration in these cells, and knockdown of TDP1 in HeLa cells resulted to be modestly sensitive to this antibiotic (Zhou *et al*., 2009). Moreover, Heo and coworkers demonstrated that human TDP1 functionally interacts with players of the initial stages of NHEJ, one of the mechanisms of DSBs repair (Heo *et al*., 2015), and TDP1-deficient yeast displayed, during

NHEJ, a decreased fidelity in repair cohesive ends (Bahmed, Nitiss and Nitiss, 2010). These data give convincing proof concerning the involvement of TDP1 in DSBs repair.

1.2.6. **TDP1 H493R mutation: SCAN1**

A specific single point mutation (Adenine 1478 to Guanine) in the TDP1 gene leads to the substitution of Histidine 493 of the second HKN motif with an Arginine. If this mutation occurs in homozygosis, it causes the autosomal recessive neurodegenerative disorder spinocerebellar ataxia with axonal neuropathy (SCAN1) (Takashima *et al.*, 2002). This is a disease of terminally differentiated post-mitotic neurons whose symptoms appear to be restricted to the nervous system, lack of chromosomal instability, and cancer predisposition (Rass, Ahel and West, 2007). To date, it has been identified only nine people, members of the same family, from Saudi Arabian affected by SCAN1. This latter is clinically characterized by late childhood-onset of cerebellar ataxia, cerebellum atrophy, peripheral neuropathy, and extra-neurological features as hypercholesterolemia and hypoalbuminemia. None of the patients showed problems in cognitive capability and lifespan (Takashima *et al.*, 2002; Walton *et al.*, 2010). The H493R mutation retains partial TDP1 activity and determines a deficiency in the enzyme turnover. This mutation does not affect in carrying out the first step of the reaction, but the second step. This is due to the formation of the TDP1H493R-DNA covalent complexes (Interthal *et al.*, 2005; Katyal *et al.*, 2007; Hawkins *et al.*, 2009). These complexes, compared to the TDP1-DNA complexes, have an increased half-life of ~ 13 min (Interthal *et al.*, 2005). TDP1 protein derived from SCAN1 patients' EBV-immortalized lymphoblastoid cell lines (LCLs) extract, showed an activity reduced of ~100-fold (Interthal *et al.*, 2005) while the recombinant TDP1 H493R mutated of ~25-fold, compared to the wild type (Interthal, Pouliot and Champoux, 2001). Thus, this decreased activity of the H493R-TDP1 and the relatively long half-life of the TDP1H493R-DNA covalent complex determine the persistence of TDP1H493R-DNA complexes. Nowadays, the only enzyme known that can excise the 3′-phosphoamide adducts formed by H493R-TDP1 is the wild-type TDP1. This capability explains the absence of SCAN1 symptoms in heterozygous carriers (Interthal, Chen and Champoux, 2005). Cells derived from patients with SCAN1 are hypersensitive to CPT (Barthelmes *et al.*, 2004; El-Khamisy *et al.*, 2005; Interthal *et al.*, 2005; Miao *et al.*, 2006; Caldecott, 2008; Das *et al.*, 2009; El-Khamisy, 2011) and its clinical derivative irinotecan (Meisenberg *et al.*, 2015). They are defective in repairing the transcriptional CPT-induced TOP1cc (Miao *et al.*, 2006), the CPT-induced DNA SSBs (El-

Khamisy *et al.*, 2005), and in removing 3'-PGs at DSBs (Zhou *et al.*, 2005; Akopiants *et al.*, 2014). To date, all the studies of the role of H493R TDP1 relied principally on the lymphoblastoid cell lines derived from patients with SCAN1 (Barthelmes *et al.*, 2004; El-Khamisy *et al.*, 2005; Interthal *et al.*, 2005; Miao *et al.*, 2006; Caldecott, 2008; Das *et al.*, 2009; El-Khamisy, 2011). To elucidate the role of TDP1 were generated TDP1 KO cell lines (Ben Hassine and Arcangioli, 2009; El-Khamisy *et al.*, 2009; Das *et al.*, 2010; Murai *et al.*, 2012) and TDP1 KO mice (Hirano *et al.*, 2007; Katyal *et al.*, 2007; Hawkins *et al.*, 2009). Three independently developed TDP1 KO mice models revealed an indistinguishable phenotype correlated to human SCAN1 compared to the wild-type mice. Only one out of the three models showed progressive age-linked cerebellar atrophy (Katyal *et al.*, 2007). However, they result in hypersensitive to CPT, to topotecan, and to bleomycin, but not to etoposide (a TOP2 inhibitor) (Pommier *et al.*, 2010) (Hirano *et al.*, 2007; Katyal *et al.*, 2007). Hirano *et al.* reported that only CPT-treated cells expressing TDP1-H493R protein accumulate H493R-DNA complexes and DNA strand breaks, but not the TDP1 KO complemented with the wild-type TDP1. These results suggest that H493R is a neomorphic mutation that generates a TDP1-H493R protein that is qualitatively different from TDP1 because it gives rise to TDP1H493R-DNA complexes. Katyal *et al.* found that TDP1 deficient neurons were defective in repair chromosomal SSBs induced from oxidative stress, CPT, or IR, but not in repairing DSBs. It is clear that TDP1 is required for the neural homeostasis, however, these TDP1 KO mice models reflect a different neuropathology phenotype compared to the human SCAN1. This can be related to the fact that TDP1 KO mice lack of the specific TDP1 H493R protein and/or on the short life-span of the mice (~2 years) compared to the late-childhood onset of the SCAN1. Only very recently, Ghosh and colleagues complemented TDP1 KO mouse embryonic fibroblasts (MEFs) with human SCAN1 H493R-TDP1 using lentivirus constructs. They have shown that trapped H493R TDP1-mitochondrial DNA (mtDNA) complexes increase in presence of mtTOP1 poison (mito-SN38). The trapping of H493R-TDP1 leads to an increase in mtDNA damage and produces potentially toxic nuclear DNA lesions as a consequence of ROS formation. The H493R TDP1-mtDNA complexes promote also mitochondrial fission and dysfunction and activate autophagy and mitophagy. This latter might have an important role in neuroprotection and may be correlated to the late onset of SCAN1 disorder (Ghosh *et al.*, 2019).

1.3. R-loops

1.3.1. How R-loop forms and how they are resolved

R-loops are transient and reversible non-B DNA, three-stranded, structures characterized by a DNA-RNA hybrid, and a displaced single-stranded DNA (ssDNA) (Reaban, Lebowitz and Griffin, 1994; Ginno *et al.*, 2012) (Figure 11). Commonly, R-loops are generated during transcription. They form during the elongating step of the RNA polymerase, when the nascent transcript re-anneals with the complementary DNA template strand, displacing the non-complementary one. This event generates a DNA-RNA hybrid that is thermodynamically more stable than the DNA-DNA helix (Roberts and Crothers, 1992). This is the most accepted model of R-loop formation, supported by the crystallographic structure of the RNA polymerase, which shows that there are two independent channels as a way out for the RNA and the DNA (Westover et al., 2004). Moreover, this model is supported by the fact that events such as the negative supercoiling accumulation behind the elongating Pol II favor R-loop formation (Drolet, Bi and Liu, 1994; Drolet *et al.*, 2003; Roy *et al.*, 2010; Aguilera and García-Muse, 2012).

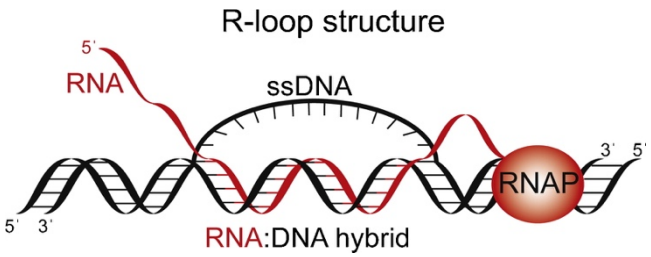

Figure 11: Schematic representation of an R-loop structure. The nascent RNA strand (red) is synthesized by RNA polymerase (RNAP, red oval) and hybridizes with the complementary DNA template strand. The non-template strand is exposed as single-stranded DNA (ssDNA). Adapted from (Hamperl and Cimprich, 2014).

Other traits enhancing R-loop formation are the GC content (GC skew regions) of the DNA sequence and the presence of DNA cleavage. The first is linked to the superior thermodynamic stability of a G-rich RNA bound to a C-rich DNA template (Roberts and Crothers, 1992; Ratmeyer *et al.*, 1994; Ginno *et al.*, 2012). Regarding the presence of DNA cleavage, Roy and colleagues demonstrated in 2010 that a single strand nick in the DNA non-template, occurred downstream of a promoter, efficiently initiate R-loop formation also in absence of G-rich regions (Roy *et al.*, 2010). Probably, the transient displacement of the DNA non-template

caused by the nick help to augment the possibility of the annealing of the transcribed RNA with the DNA template (Roberts and Crothers, 1992). Cells have also developed mechanisms to resolve R-loops. For example, endonuclease enzymes, such as the monomeric RNase H1 and the trimeric RNase H2, conserved from bacteria to humans, hydrolyze the RNA of RNA⁄DNA hybrids (Cerritelli and Crouch, 2009). In addition, cells encode RNA/DNA helicases able to remove R-loops, such as human senataxin, yeast homolog Sen1, (Sen1/SETX) (Skourti-Stathaki et al., 2011, Groh et al., 2017, Mischo et al., 2011), DEAD box protein 23 (Ddx23) (Sridhara et al., 2017), human DEAH box protein 9 (DHX9) (Cristini et al., 2018) and Aquarius (AQR) (Sollier et al., 2014). Another player able to unwind the RNA/DNA hybrid is the nucleopore-associated mRNA export factor Ddx19 (Hodroj *et al.*, 2017). These factors have an important role in maintaining R-loops homeostasis and in preventing genome instability (see section 1.3.5).

1.3.2. **R-loops across the genome**

R-loops span 100-2000 base pairs (Santos-Pereira and Aguilera, 2015). Since the first characterization, through in vitro analysis conducted by Thomas and colleagues in 1976 (Thomas, White and Davis, 1976), followed by in vivo observation of RNA-DNA hybrid at the replication origin of a ColE1-like plasmid in 1980 (Itoh and Tomizawa, 1980) and in cellulo in 1994 (Drolet, Bi and Liu, 1994), they were detected and studied in different organisms from bacteria to higher eukaryotes (Drolet, Bi and Liu, 1994; Li and Manley, 2006; Aguilera and García-Muse, 2012; Groh and Gromak, 2014; Hamperl and Cimprich, 2014; Skourti-Stathaki and Proudfoot, 2014; Xu *et al.*, 2017). During these years the methodologies to identify RNA-DNA hybrids in vivo have been developed (see section 1.3.3). One of the most frequent is based on the use of S9.6 anti-10 nucleotide DNA-RNA hybrid monoclonal antibody (Boguslawski *et al.*, 1986). To date, this is the only available antibody to detect R-loops and it binds DNA-RNA hybrids with a subnanomolar affinity (KD, dissociation constant, of about 0.6 nM). Treatment with RNase H, which degrades specifically the RNA strand in the hybrid, is commonly used to assess the specificity of the S9.6 signal because this antibody also recognizes RNA-RNA duplexes, albeit with a weaker affinity (about five folds less). Nevertheless, S9.6 has a high affinity for AU-rich dsRNA (KD of 2.7 nM) (Phillips *et al.*, 2013). This antibody is commonly used to perform DNA-RNA ImmunoPrecipitation (DRIP), which is a technique that lets the isolation of R-loops from the genomic DNA. Then this DNA-RNA hybrids can be analyzed at specific loci by the following qPCR or at a genome-wide level by sequencing (Ginno *et al.*,

2012). Thanks to this technique, genome wild mapping studies were possible and allowed to identify that R-loops cover about 8% of the yeast genome, 10% of the Arabidopsis genome and nearly 5% of the human genome (Sanz *et al.*, 2016; Wahba *et al.*, 2016; Xu *et al.*, 2017). R-loops formation occurs over tens of thousands of genomic loci (Ginno *et al.*, 2012). They are preferentially present in highly transcribed genes, in promotor and terminator regions, in CpG islands (CGIs) (CpG-rich regions corresponding to unmethylated DNA segments which function as promoters for the majority of human genes), at the Transcription Starting Site (TSS) and termination site, in centromeres and telomeres, in some transposable elements (Ty in yeast), antisense-RNAs or non-coding (nc) RNAs regions (Ginno *et al.*, 2012; Chan, Hieter and Stirling, 2014; El Hage *et al.*, 2014; Wahba *et al.*, 2016; Chen *et al.*, 2017). Additionally, R-loop formation close promoters that present G-rich sequences lead to form G-quadruplex in the non-template-DNA (Figure 12) (Ginno *et al.*, 2012, 2013; Chen *et al.*, 2017; De Magis *et al.*, 2019).

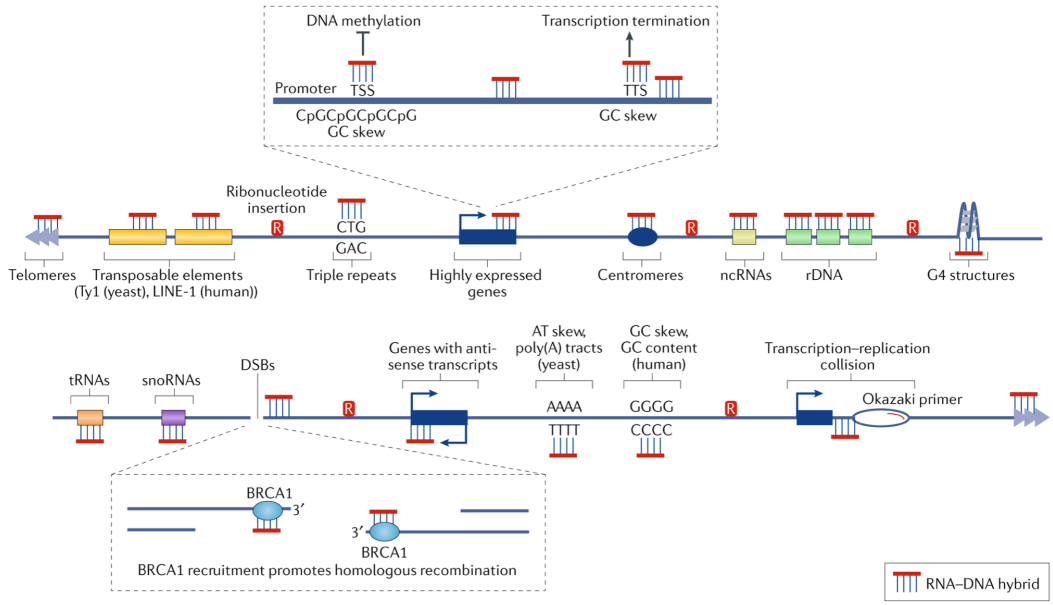

Figure 12: R-loops across the genome. Adapted from (Niehrs and Luke, 2020).

Moreover, by DRIP-qPCR analysis, it was possible to know how frequent do R-loops form. The frequency depends on the locus and oscillates from 1% to 10% of genomic DNA input (Skourti-Stathaki, Kamieniarz-Gdula and Proudfoot, 2014; Sanz *et al.*, 2016; Stork *et al.*, 2016). However, for untranscribed and/or intergenic regions the frequency was from 0.01% to 0.1% of input. It was also investigated the half-life of R-loops generated at the promoters by a kinetic

analysis after treatment with DRB and further DRIP-qPCR. Most DNA-RNA hybrids disappeared within 30 min after transcription block, showing a half-life in an average of 10 min (Sanz *et al.*, 2016). This rapid turnover indicates that R-loops are frequently formed and resolved (see section 1.3.5; Figure 14).

1.3.3. **Methods used for R-loop detection**

Other methods to detect R-loops using the S9.6 antibody include immunofluorescence (IF) microscopy and DNA slot-blot. In the IF the S9.6 antibody is used as a primary antibody that is then recognized by a secondary antibody associated to a fluorophore to detect the signal, it let to identify the localization and the distribution of the RNA-DNA hybrids in cells (Marinello *et al.*, 2013; De Magis *et al.*, 2019). The DNA slot-blot method relies on the immobilization of nucleic acids on solid supports, membranes. Then the membrane is immunoblotted using the S9.6 antibody. This technique allows analyzing quantitatively the overall of R-loops present in genomic DNA extracts (Hegazy, Fernando and Tran, 2020). Another technique used to isolate R-loops is the DNA-RNA In Vitro Enrichment (DRIVE), it is based on the usage of the mutated RNase H, which bind RNA/DNA hybrids without degrading them. However, with this technique, once carried out the sequencing, only 1,224 peaks were identified compared to the 20,862 obtained using the DRIP-seq (Ginno *et al.*, 2012). R-loops can also be detected by an indirect method using sodium bisulfite (Yu *et al.*, 2003). The sodium bisulfite allows detecting unpaired C nucleotide from the double-strand DNA without denaturation of the duplex (Gough, Sullivan and Lilley, 1986). Precisely, the bisulfite deaminates unpaired C converting them into uracil (U). After performing qPCR or sequencing, all the U are incorporated as A by the polymerase, obtaining a transition of C-G to T-A. Thus, C paired with G in R-loop structure or that are methylated do not react with bisulfite, while C belonging to the single-stranded region are converted, letting to identify the single-strand belonging to the RNA-DNA hybrid structure (Yu *et al.*, 2003).

1.3.4. **Physiological roles of R-loops**

R-loops have several physiological roles. There is evidence that R-loops generated in sequences presenting guanine clusters play an important role in the immunoglobulin class switch recombination (CSR) (Yu *et al.*, 2003; Roy and Lieber, 2009). CSR allows the production of different classes of antibody (IgM ad IgG, IgA o IgE) (Kinoshita and Honjo, 2000). The formation of R-loops occurs at the G-rich switch region of the IgH locus during transcription

and let the displaced ssDNA to be accessible to activation-induced cytidine deaminase (AID) (Muramatsu *et al.*, 2000). This latter deaminates the C residues into U, which are then processed into DNA nicks by BER or mismatch repair pathway (Masani, Han and Yu, 2013). The SSBs are then converted into DSBs, which are repaired by NHEJ, this allows DNA end-joining and the "switch" of the Ig class (Stavnezer, Guikema and Schrader, 2008). R-loops play a role in gene expression regulation. Their function in transcription regulation was suggested by the fact that these structures form during transcription and because of their preferential distribution in highly transcribed genes, in promotor and terminator regions (Ginno *et al.*, 2012, 2013). At promoters, R-loops act generally as a transcription activator. They suppress methylation-associated transcription silencing by preventing the binding of the DNA methyltransferases to the DNA (Ginno *et al.*, 2012; Grunseich *et al.*, 2018). At promoters they can also both promote and inhibit the binding of chromatin remodelers, to let the chromatin accessible to the RNA polymerase (Santos-Pereira and Aguilera, 2015). Cloutier and coworkers have shown, in yeast, that R-loops arising from long non-coding RNAs, can bind across promoters, displace repressors, promoting transcription (Cloutier *et al.*, 2016). R-loop function in gene regulation was also found in plants. In *Arabidopsis thaliana*, the antisense transcripts COOLAIR promote the transcriptional silencing of FLOWERING LOCUS C (FLC) in a temperature-dependent manner. There is evidence that the formation of R-loops at the COOLAIR promoter region decreases the expression of these antisense transcripts (Sun *et al.*, 2013). At the transcription termination site, R-loops, formed over of a G-rich region downstream of the poly(A) site, can stall the RNA polymerase. Their further resolution by SETX favors the transcription termination (Figure 13) (Skourti-Stathaki, Proudfoot and Gromak, 2011). Recent studies show that R-loops can also arise after DNA damage, both inhibiting and promoting DNA repair. They could compromise the repair hampering the recruitment of DNA repair factors at DSB sites (Aguilera and Gómez-González, 2017), altering the chromatin structure flanking DSBs (Cohen *et al.*, 2018), and leading to abnormal repair (Amon and Koshland, 2016; Cohen *et al.*, 2018). Examples of RNA–DNA hybrids driving DNA repair are found at the level of DSBs formed in transcribed regions (Michelini *et al.*, 2017; Puget, Miller and Legube, 2019) and of short telomeres (Graf *et al.*, 2017). Regarding the first case, R-loops can generate from the stalling of the transcription machinery, from de novo transcription at the break site, as occur in damage-induced long non-coding RNA (lncRNA), or following annealing in trans of a pre-existing RNA. Once the RNA–DNA hybrid forms, it lures BRCA1. This latter can bind directly to RNA–DNA hybrids in vitro (D'Alessandro *et al.*, 2018). BRCA1 then recruits BRCA2 and

PALB2 (partner and localizer of BRCA2) to the DSB. Once the hybrid is removed, by helicases as SETX or DEAD-box helicase 1 (DDX1), or RNase H1, RNase H2, RAD51 is loaded onto the ssDNA and the HR takes place. At chromosome ends, the telomere repeat-containing RNA (TERRA) tends to generate R-loops. TERRA R-loops endorse DSBs repair by RAD51-mediated homology-directed repair. However, R-loop removal pathway has not been identified and is uncertain if TERRA can generate R-loops in *trans* at short telomeres (reviewed in (Niehrs and Luke, 2020)).

RNA in a regulatory R-loop or hybrid	Hybrid location	Effect
ANRASSF1A (Beckedorff *et al.*, 2013)	*RASSF1A* promoter	H3K27me3 deposition and *RASSF1A* repression
VIM-AS1 (Boque-Sastre *et al.*, 2015)	*VIM1* promoter	NF-κB recruitment and *VIM1* transcription
GATA3-AS1 (Gibbons *et al.*, 2018)	*GATA3* promoter	H3K4me3 deposition and *GATA3* transcription
TARID (Arab *et al.*, 2019)	*TCF21* promoter	DNA demethylation and *TCF21* repression
TERRA (Graf *et al.*, 2017)	Telomeres	Repair of short telomeres by homologous recombination
TIP60–p400-dependent genes (Chen *et al.*, 2015)	Subset of lncRNA in embryonic stem cells	Prevent PRC-mediated gene repression
Select PRC target genes (Skourti-Stathaki *et al.*, 2019)	Promoters	Promote PRC target gene repression
Centromeric RNA (Kabeche *et al.*, 2018)	Centromeres	Activation of Aurora B and centromere function
Damage-induced lncRNA (Michelini *et al.*, 2017) and other RNAs at DSBs (Puget, Miller and Legube, 2019)	Any DSB in a transcribed region	RAD51 recruitment and homologous recombination

ANRASSFIA, RASSF1 antisense RNA 1A; DSB, DNA double-stranded break; H3K27me3, trimethylation of Lys27 on histone H3; H3K4me3, trimethylation of Lys4 on histone H3; lncRNA, long non-coding RNA; NF-κB, nuclear factor-κB; PRC, Polycomb repressive complex; TARID, TCF21 antisense RNA inducing demethylation; *TCF21*, tumour suppressor transcription factor 21; *TERRA*, telomere repeat-containing RNA; *VIM-AS1*, VIM antisense RNA 1.

Table 3: Selected examples of regulatory R-loops and their effects. Adapted from (Niehrs and Luke, 2020).

On the other hand, R-loops may also promote DSB repair (Ohle *et al.*, 2016; Yasuhara *et al.*, 2018) considering that removal of R-loops decreases the efficiency of HR and NHEJ (Lu *et al.*, 2018). Other examples of the regulatory role of R-loops are summarized in Table 3. Also, the

RNA-DNA hybrid is generated during the process of bacteria immune defense by the CRISPR-Cas9 system, a nowadays well-developed tool for genome editing (Ran *et al.*, 2013).

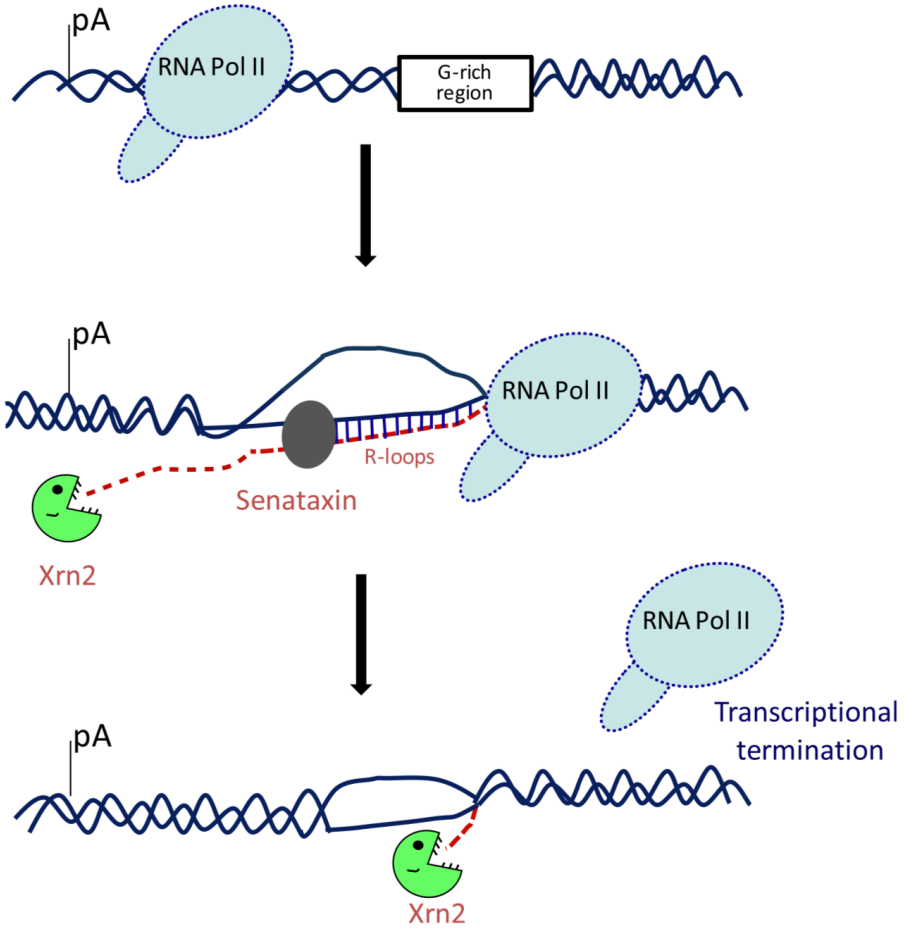

Figure 13: Model for Role of R-Loops and Senataxin in Transcriptional Termination. DNA is shown as solid blue lines and RNA as a dotted red line. Vertical blue lines denote RNA/DNA hybrids. Senataxin is shown as a gray oval over the R-loop region. The 5'–3' exonuclease Xrn2 that degrade the nascent RNA from the site of poly(A) is represented in green. Adapted from (Skourti-Stathaki, Proudfoot and Gromak, 2011).

1.3.5. **Deregulation of R-loop homeostasis and DNA damage**

There are many DNA and RNA metabolism factors that regulate R-loop homeostasis, by maintaining the equilibrium between the prevention of R-loops formation and their removal (Figure 14). For the prevention, the key players are the RNA processing/export factors (Huertas and Aguilera, 2003; Li and Manley, 2005), TOP1 (Drolet *et al.*, 2003; El Hage *et al.*, 2010; Manzo *et al.*, 2018) and chromatin (Sanz *et al.*, 2016; Chen *et al.*, 2017; García-Pichardo *et al.*, 2017; Salas-Armenteros *et al.*, 2017). While, for the removal, the crucial players are the RNase H enzymes (Cerritelli and Crouch, 2009) and the RNA/DNA helicases, as SETX (Mischo *et*

al., 2011; Skourti-Stathaki, Proudfoot and Gromak, 2011; Groh and Gromak, 2014), Ddx23 (Sridhara *et al.*, 2017), DHX9 (Cristini *et al.*, 2018) and AQR (Sollier *et al.*, 2014). Defect of these factors and conflicts between the replication and the transcription machinery (Gan *et al.*, 2011; Aguilera and García-Muse, 2012; Hamperl and Cimprich, 2014; Kotsantis *et al.*, 2016) lead to an accumulation of unscheduled R-loops. These last drive to transcription-associated mutagenesis (TAM), recombination (TAR), DNA damage and genomic instability, events correlated to cancer and genetic diseases (Huertas and Aguilera, 2003; Costantino and Koshland, 2015; Sollier and Cimprich, 2015; García-Muse and Aguilera, 2019).

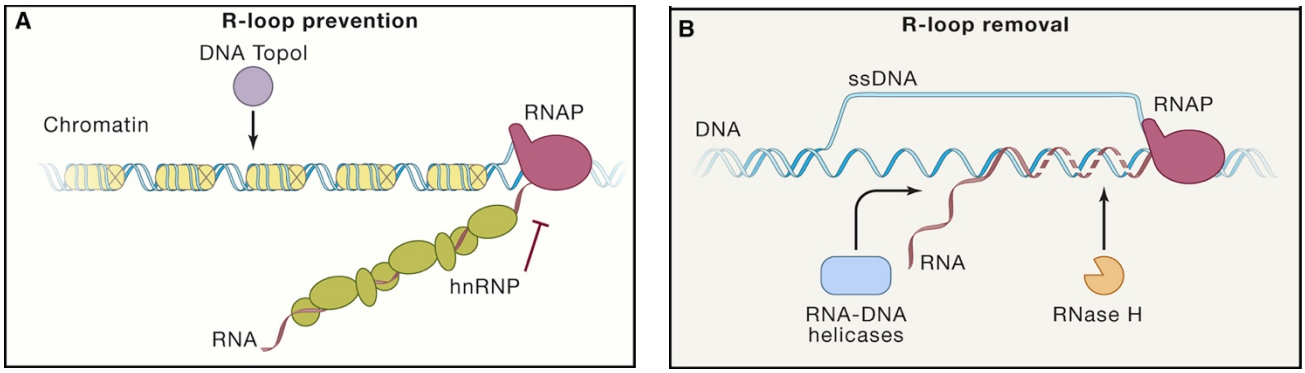

Figure 14: Schematic representation of the factors that regulate R-loops homeostasis. (A) R-loop prevention occurs by specific RNA-binding proteins (represented in green color) that are involved in RNA biogenesis (hnRNP); by topoisomerase 1 (Topo I) (represented in lilac color) that resolves the negative supercoiling behind the elongating RNA polymerase II (RNAP), and by chromatin (in yellow color). (B) R-loop removal is obtained by RNase H enzymes (in orange) that degrade the RNA strand of the hybrid and by DNA-RNA helicases (in light blue) that unwind the hybrid. Adapted from (García-Muse and Aguilera, 2019).

R-loop formation is more mutagenic for the non-transcribed ssDNA compared to the transcribed strand, which generates the RNA-DNA hybrid (Beletskii and Bhagwat, 1996). In fact, the ssDNA is more sensitive to mutations and damage because of its exposition to endogenous enzymes such as nucleases or deaminases, e.g., AID. These last convert cytosine in uracil, changing consequently a C:G base pair into a U:G mismatch (Petersen-Mahrt, Harris and Neuberger, 2002; Chaudhuri, Khuong and Alt, 2004). Because the nascent mRNA has a key role in generating R-loops at genes, deficiencies in messenger ribonucleoprotein particle (mRNP) assembly can stimulate R-loop formation (Santos-Pereira and Aguilera, 2015). The first evidence that R-loops were causing genomic instability was found in *S. Cerevisiae* THO complex (involved in the mRNP biogenesis and transcription) mutants (Huertas and Aguilera, 2003). In these mutants, R-loop formation drives to hyper TAR phenotype, plasmid, and chromosome loss. While, in mammalians, the link between R-loop formation and genomic instability associated to DSB generation was shown in chicken DT40 and Hela cells depleted

for the serine/arginine-rich splicing factor 1 (SRSF1; known as ASF and SF2) (Li and Manley, 2005; Aguilera and Gómez-González, 2008). Alteration on DNA topology regulation leads to negative supercoils accumulation during transcription, this is also a source of unscheduled R-loops. Tuduri and colleagues reported, in mammalian TOP1-depleted cells, an increase of DNA breaks and replication defects in transcription- and RNase H-dependent manner (Tuduri *et al.*, 2009). Further, it has been reported that transcriptional-DSBs are produced by two SSBs on the opposing DNA, one result of the TOP1cc removal performed by the TDP1 pathway, and the other one arising from the cleavage of the R-loops, carried out by specific endonucleases, such as XPF, XPG, and FEN1 (Sollier *et al.*, 2014; Cristini *et al.*, 2019). Another possibility of genomic instability induced by R-loops is through interfering with DNA replication, blocking the progression of the replication forks (Gan *et al.*, 2011). Studies have shown that R-loops are responsible for replication fork stalling when it collides with the transcription elongation apparatus, and these transcription-replication conflicts (TRCs) induce genome stability (Gan *et al.*, 2011; Aguilera and García-Muse, 2012; Kotsantis *et al.*, 2016). The different cell process head-on orientation (HO) or the co-directional (CD) TRCs influence the formation, respectively promoting or reducing R-loops, and induce distinct DNA damage responses in human cells (Figure 15) (Hamperl *et al.*, 2017).

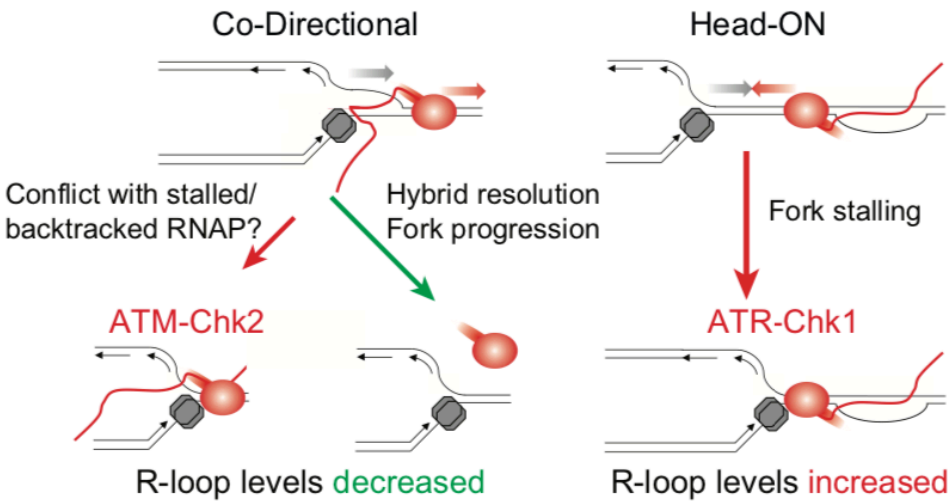

Figure 15: Model for how head-on and co-directional transcription-replication conflicts regulate R-loop homeostasis and induce distinct DNA damage responses in human cells. The gray box represents the replication machinery, while the red oval the transcription one. The DNA and the RNA are represented respectively in gray and red lines. Adapted from (Hamperl *et al.*, 2017).

Therefore, R-loops, which have been associated with DSBs markers, such as phosphorylated γH2AX histone and DNA damage checkpoint activation (Skourti-Stathaki and Proudfoot, 2014; Sollier and Cimprich, 2015), can also lead to chromosomal rearrangements and genome instability afterward DNA cleavage repair by error-prone mechanisms.

Thus, given R-loops' involvement in gene expression regulation and DNA damage, their regulation is crucial as the DNA-RNA hybrid should form at the right place and time to let that certain regulatory processes take place without any deleterious consequences.

1.3.6. **R-loops and Human diseases**

Several diseases, such as cancers and neurodegenerative disorders, are emerging to be related to R-loop accumulation (some summarized in Table 4) (Richard and Manley, 2017; Perego *et al.*, 2019). In many neuronal disorders, repeat expansions have a crucial role. For instance, Friedreich's ataxia patients present GAA repeat expansions and in the cells deriving from patients, it was found that R-loops arise at the level of this trinucleotide repeats, leading to the transcriptional silencing of the genes related to this disease (Groh *et al.*, 2014). Moreover, Li and co-workers were able to restore the gene expression levels in Friedreich's ataxia cells by introducing an antisense oligonucleotide and a dsRNA able to identify the GAA trinucleotide repeats (Li, Matsui and Corey, 2016). The involvement of the R-loops in silencing long non-coding RNAs has been studied also in human cells. For instance, in Prader-Willi syndrome (PWS), belonging to the autism spectrum disorders, they prevent the expression of the Ubiquitin Protein Ligase E3A (Ube3a) (Powell *et al.*, 2013). Moreover, inactivating mutations in RNase H1 and SETX can drive to disastrous neurological disorder (Santos-Pereira and Aguilera, 2015). For example, a mutation on SETX is correlated to ataxia with oculomotor apraxia type 2 (AOA2) and amyotrophic lateral sclerosis type 4 (ALS4). Cells deriving from AOA2 patients present and increase in R-loops levels and modification in neuronal genes (Becherel *et al.*, 2015). Also in ALS, it was found that gene alteration drives an accumulation of R-loops in motor neurons (Perego *et al.*, 2019). The loss of Ddx23 is frequent in adenoid cystic carcinoma (ACC) (Sridhara et al., 2017) while DHX9, which interacting with PARP1 prevents R-loop-DNA damage, is overexpressed in cancer (Cristini *et al.*, 2018). Very recently, it has been shown that Embryonal Tumours with Multi-layered Rosettes (ETMRs) cells present a very upregulated expression of DNA/RNA helicases. In add, ETMRs cells KO for Dicer1, a protein possessing a helicase activity, show a correlated increase of the S9.6 antibody and γH2AX signal. Similar results were obtained in ETMRs treated with TOP1 inhibitor and were even

pronounced when the treatment was performed in combination with PARP inhibitor. These results suggest not only that R-loops could be involved in chromosomal instability in ETMRs but also that miRNA processing and TOP1 defects can boost R-loops levels (Lambo *et al.*, 2019).

Gene	Disease	Cause
SETX	Ataxia-ocular apraxia type 2 (AOA2) and amyotrophic lateral sclerosis type 4 (ALS4)	Mutations in the RNA–DNA helicase *SETX*
FXN	Friedreich ataxia (FRDA)	Expansion of GAA repeats in *FXN* gene promotes R-loop formation, H3K9me2 and decreased *FXN* expression
FMR1	Fragile X syndrome (FXS) and fragile X-associated tremor/ataxia syndrome (FXTAS)	Expansion of CGG repeats in *FMR1* gene promotes R-loop formation, H3K9me2 and decreased *FMR1* expression
C9orf72	Amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD)	Expansion of GGGGCC repeats causes R-loop formation and accumulation of aborted transcripts
BRCA1	Cancer	Genome instability caused by R-loop accumulation in BRCA1-deficient cells
BRCA2	Cancer and Fanconi anaemia (FA)	Genome instability caused by R-loop accumulation in BRCA2-deficient cells
FIP1L1	Cancer	Genome instability caused by R-loop accumulation in FIP1L1-deficient cells inferred by the yeast mutant *fip1Δ*
BRE1	Cancer	Genome instability caused by R-loop accumulation in BRE1-deficient cells
SRSF1	Cancer	Deregulation of cancer-associated genes due to *SRSF1* overexpression
ORF57	Kaposi sarcoma-associated herpesvirus (KSHV)	Sequestration of human TREX complex by ORF57 causes R-loop formation and DNA damage

C9orf72, chromosome 9 open reading frame 72; FIP1L1, factor interacting with PAPOLA and CPSF1; *FMR1*, fragile X mental retardation 1; *FXN*, frataxin; H3K9me2, histone H3 lysine 9 dimethylation; *SETX*, senataxin; *SRSF1*, serine/arginine-rich splicing factor 1.

Table 4: Genes related to R-loop metabolism that can cause human diseases if dysfunctional. Adapted from (Santos-Pereira and Aguilera, 2015).

2. AIMS

The aim of my PhD research project is to characterize the genome instability and DNA damage triggered by the H493R mutation of the TDP1 gene in non-replicating cells.

Thus, I have first investigated whether the H493R TDP1 mutant leads to an increase in endogenous DSBs and genomic instability in not replicating cells (G1 phase) and compare it to TDP1 knockout (KO). To further analyze the reasons why this increase in DSBs is primarily in the G1 phase compared to the S phase. Then, I questioned how these DSBs are produced, investigating the involvement of R-loops in this mechanism.

I used several molecular and cellular techniques to achieve the project aims. I generated homozygous H493R TDP1 U2OS, TDP1 KO U2OS cell lines TDP1 and TDP1 KO WI-38

hTERT cell lines by CRISPR-Cas9 technology. Microscopy analysis of EdU (5-ethynyl-2'-deoxyuridine) incorporation allowed us to discriminate the G1 phase from S phase cells and we took advantage of the Operetta High Content Imaging System technology to measure γH2AX and p53BP1 nuclear foci as a readout of DSBs formation. Moreover, to investigate DSBs getting rid of replication, I performed microscopy analysis and neutral comet assay on quiescent WI-38 hTERT RNAi for TDP1. Then, I carried out DNA:RNA hybrid immunoprecipitation analysis (DRIP) with Ab S9.6 to investigate if TDP1 deficiency promotes R-loops modulation. This would suggest about their potential involvement in the mechanism of these DSBs production. To study if these DSBs derive from an increase in their production and/or a failure in their repair, I then focused on CPT-induced DSBs. I performed microscopy analysis to measure γH2AX nuclear foci in cells treated with CPT and after CPT removal.

3. Results

The main objective of my PhD thesis was to investigate the genomic instability in the presence of the H493R-mutated TDP1 gene and to compare it to TDP1 gene deficiency, focusing our attention mainly to non-replicating cells (G0-G1 phase). To achieve this goal, we employed the CRISPr-Cas9 methodology in two human cell lines: WI38 hTERT and U2OS. The Cas9 nuclease associates with a guide RNA (gRNA) to target DNA sequences complementary to the gRNA and called protospacers. DNA double-strand breaks (DSBs) generated by the Cas9 nuclease are generally repaired by NHEJ or HR. In the absence of a repair template, the DSBs repaired through the NHEJ process can give rise to insertion/deletion (indel). Indel can promote a shift in the reading frame, resulting in a null allele. If a DNA template with homology for the cut region is provided, it can be used for HR repair, thus allowing the insertion of the desired sequence at the Cas9 cutting site (Figure 16). The repair template can either be in the form of double- stranded DNA or single-stranded DNA oligonucleotides (ssODNs). The latter supplies a simple and effective method to make small edits in the genome, such as the introduction of single-nucleotide mutations. However, the repair pathway that is chosen and the editing efficiency can vary widely depending on the cell type and state, as well as on the genomic locus, the chromatin packaging, and repair template (Ran *et al.*, 2013; Jeggo and Downs, 2014; Scully *et al.*, 2019).

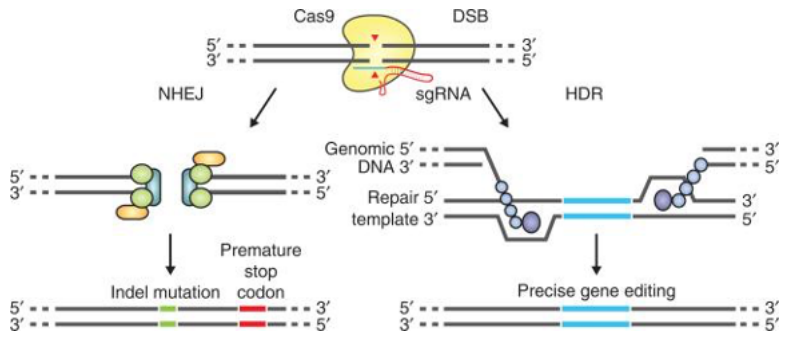

Figure 16. DSB repair promotes gene editing. DSBs induced by Cas9 (yellow) can be repaired in one of two ways. In the error-prone NHEJ pathway, the ends of a DSB are processed by endogenous DNA repair machinery and rejoined, which can result in random indel mutations at the site of the junction. Indel mutations occurring within the coding region of a gene can result in frameshifts and the creation of a premature stop codon, resulting in gene knockout. Alternatively, a repair template in the form of a plasmid or ssODN (single-stranded oligodeoxyribonucleotide) can be supplied to leverage the homology-directed repair (HDR) pathway, which allows high fidelity and precise editing. Single-stranded nicks to the DNA can also induce HDR. Adapted from (Ran *et al.*, 2013).

3.1. Generation of WI38 hTERT TDP1 KO stable cell lines

We used primary human WI38 fibroblasts immortalized with hTERT (Jeanblanc *et al.*, 2012) because non-transformed cells typically have low genomic instability (Löbrich *et al.*, 2010), and they can be induced in quiescence following serum deprivation (Dimri, Hara and Campisi, 1994; Coller, Sang and Roberts, 2006), allowing the analysis of replication-independent damage. The strategy used to generate WI38 hTERT TDP1 KO cells is illustrated in Figure 17. To generate the WI38 hTERT TDP1 KO stable cell lines we targeted the exon 14 of *TDP1* (Figure 17A) (see Methods for the sequences of the gRNAs). WI38 hTERT cells, 24 hours after seeding in 6-well plates, were transduced with lentiviral constructs that deliver hSpCas9 and blasticidin resistance. The cells infected were grown for a week in a medium containing blasticidin (10 µg/ml), to allow for the selection of clones that have integrated the Cas9. Besides, following the protocol's instructions, we cloned the gRNA sequence into the lentiGuide-Puro vector. The WI-38hTERT/Cas9 cells were infected with this latter. Then, cells were kept under puromycin (2 µg /ml) selection for two weeks. For the amplification of stable final clones KO for TDP1, survivors' colonies were isolated into a 24-well plate and then further amplified into larger 6-well plates. Thus, stable clones were grown in duplicate, harvested, and then processed for analysis at protein and DNA levels (Figure 17B).

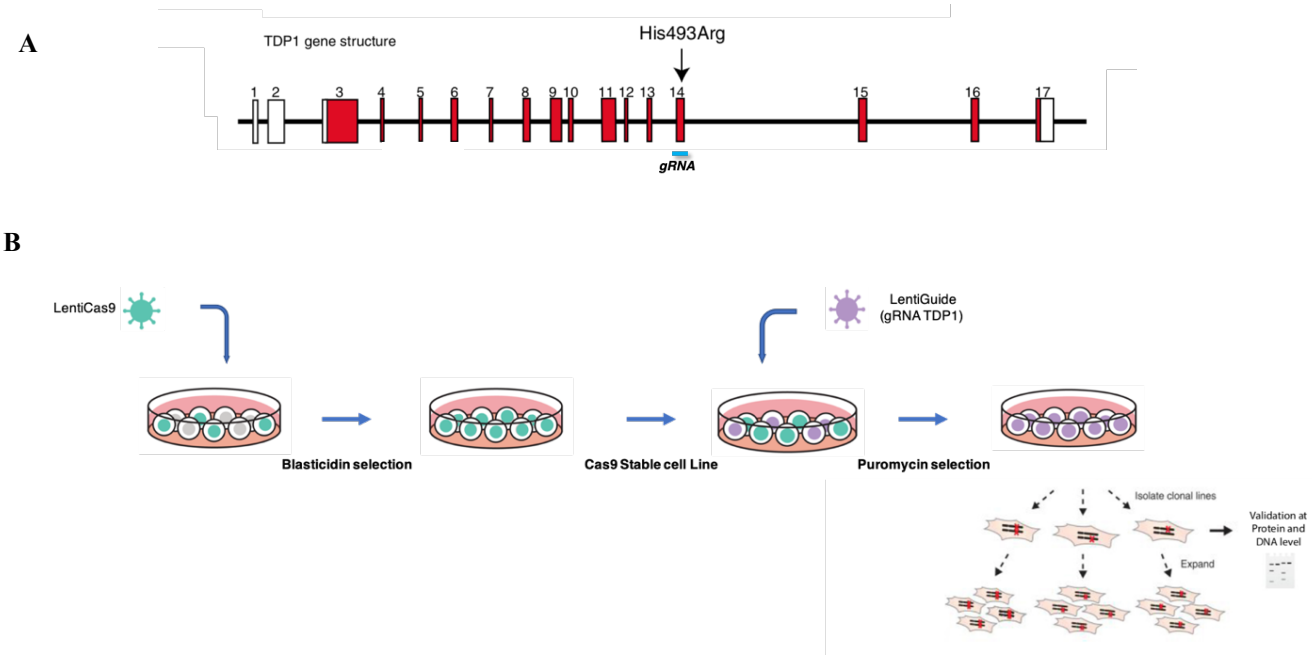

Figure 17. (A-B**) Schematic representation of the experimental design to generate the WI38 hTERT TDP1 KO stable cell lines. (**A**)** Schematic representation of the TDP1 gene structure and the gRNA selected to generate WI38 hTERT TDP1 KO stable cell lines. Exons are represented as boxes and introns as lines, in blue is indicated the gRNA targeting the exon 14, the arrow indicates the H493R mutation**.** Adapted from (Takashima *et al.*, 2002); **(**B**)** Schematic representation of the procedure used.

The choice of western blotting (WB) as a readout for such screening was made possible by the availability of a commercial antibody that recognizes the 1-50 aa at the N-terminal of the Human TDP1 protein. Thus, if there were any truncated proteins in our cell models would be detected. It was an efficient and practical readout for our purpose. WI38 hTERT potential TDP1 KO clones were processed to obtain a whole cell protein extract. These latter were then analyzed by WB to evaluate the levels of TDP1. We obtained that 4 out of the 6 clones screened showed undetectable levels of TDP1 protein (data not shown). We selected 2 clones to proceed with our analysis named WI38 hTERT TDP1 KO#1 and KO#2. In Figure 18A, looking at the expected TDP1 size (68kDa), our data showed that both clones, presented not detected level of protein compared to the WT. No truncated protein products were detected. Then, we validated the TDP1 KO#1 and the TDP1 KO#2 at the DNA level. For this purpose, cells were harvested for DNA extraction to perform PCR amplification of TDP1 exon 14 (Figure 18B). The purified amplicons were sent to sequencing. The analysis of the chromatograms (Figure 18C) confirmed that, compared to the parental WI38 hTERT cell line, both TDP1 KO#1 and TDP1 KO#2 underwent indel events at the expected Cas9 cut site, causing the disruption of the reading frame.

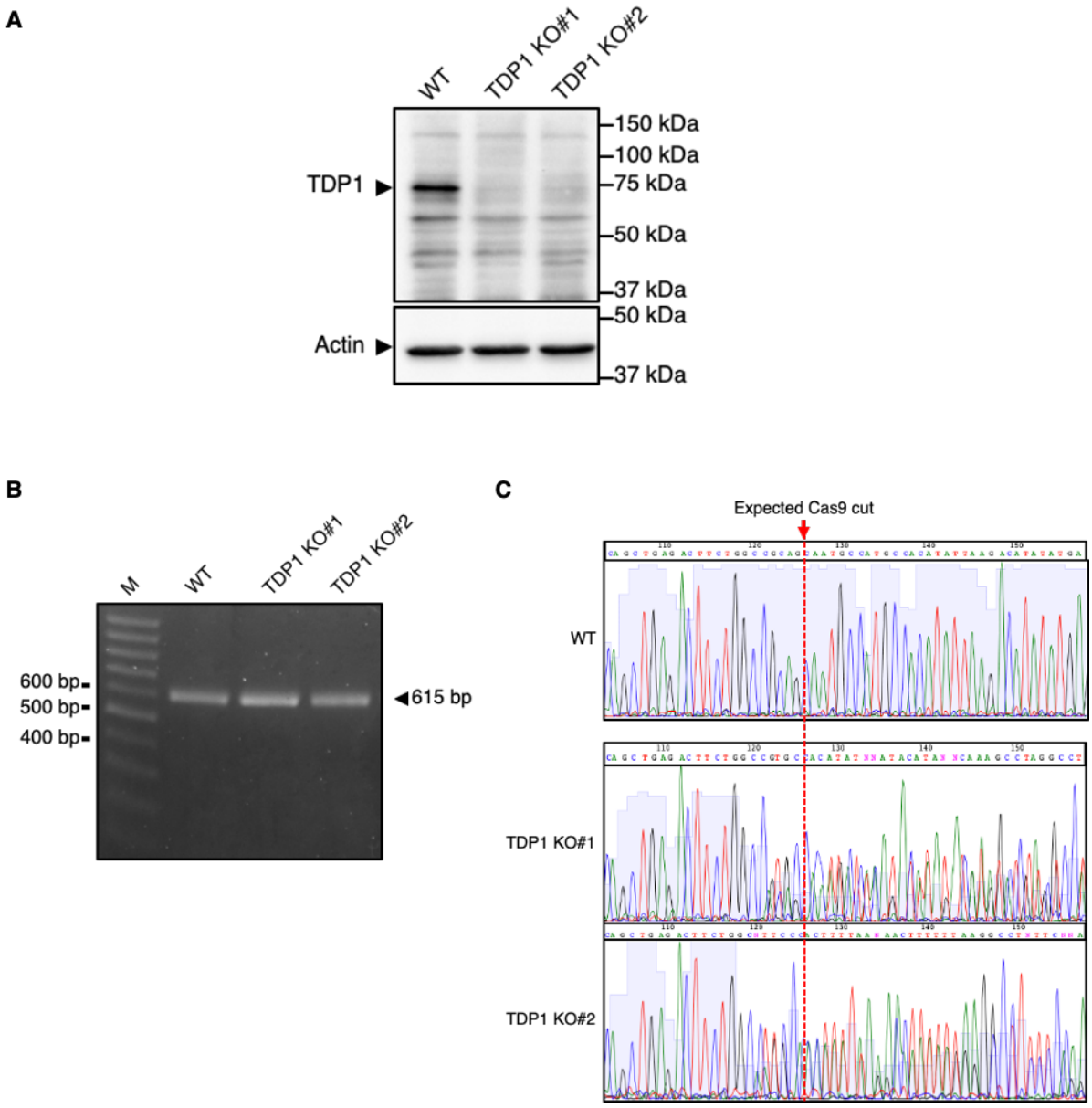

Figure 18. Validation at protein and DNA level of WI38 hTERT TDP1 KO. (A) WB analysis of TDP1 levels. Whole cell protein extracts were immunoblotted against TDP1. Actin was used as loading control; (B) PCR amplicons of TDP1 exon 14 in the WI38 hTERT TDP1 KO clones. Agarose gel showing the TDP1 Exon 14 amplicons (615 bp); (C) Sequencing analysis of the TDP1 locus in the WI38 hTERT clones. Chromatogram showing the genome editing by CRISPR/Cas9 system. Genomic DNA of the parental U2OS cell line and the TDP1 KO clones were used as a template in PCR with primers that amplify at the level of the gRNA targeted TDP1 Exon 14. The purified amplicons were sent to sequencing. The red arrow and the red dashes line indicate the expected Cas9 cut site.

We proceeded to carry out the TIDE (Tracking of Indels by DEcomposition) analysis. The TIDE is a software that allows the identification and the quantification of the editing efficacy at a given locus using the chromatograms. This software requires as input a control sequence data file (e.g. WI38 WT), a sample sequence data file (e.g. WI38 TDP1 KO#1), and a character string representing the gRNA sequence (20 nt). TIDE uses a specific algorithm to first align the

gRNA sequence to the control sequence to determine the position of the expected Cas9 break site. Then, the control sequence region upstream of the break site is aligned to the sample sequence in order to determine any offset between the two sequence reads. The software uses the peak heights for each base to determine the relative abundance of aberrant nucleotides over the length of the whole sequence trace (Brinkman *et al.*, 2014). One attractive feature of TIDE is the possibility to make accurate studies by intuitive graphs, as the Indel Spectrum (indel stands for insertion/deletion). This latter presents two main parameters: *indel identification*, which presents the different nucleotide insertions and deletions that occurred around the Cas9 break site, as well as the percentage of sequences presenting such modification. Nucleotides insertion/deletion multiple of 3 lead to an in-frame insertion/deletion of amino acid(s) in the protein sequence, and thus do not correspond to a null-allele; e*fficiency*, indicates the percentage of sequences present in the PCR products, which are resolved by the *indel* model presented. Values lower than 90% (value fixed by the software developer) indicates that events that occurred around the break site are not resolved by the software and could indicate the presence of undetected wild type sequences.

By TIDE analysis, both clones WI38 hTERT TDP1 KO#1 and TDP1 KO#2 presented a high editing *efficiency* (~93% and 97%, respectively). The TDP1 KO#1 presented a deletion at the Cas9 cut site of 4 and 11 nt and the TDP1 KO#2 of 13 and 14 nt. Not multiple of 3 so not significant to give rise to a truncated protein (Figure 19A, B).

Primary cells are challenging to edit and often refractory to DNA transfection and other gene delivery methods. This is the reason why we used this strategy of co-expressing constitutively both the Cas9 and the guide to obtain TDP1 KO clones. To introduce the A1748G single point mutation on the exon 14 of TDP1 (SCAN1 mutation), we used the osteosarcoma U2OS cell line and changed our strategy of gene editing using CRISPR/Cas9.

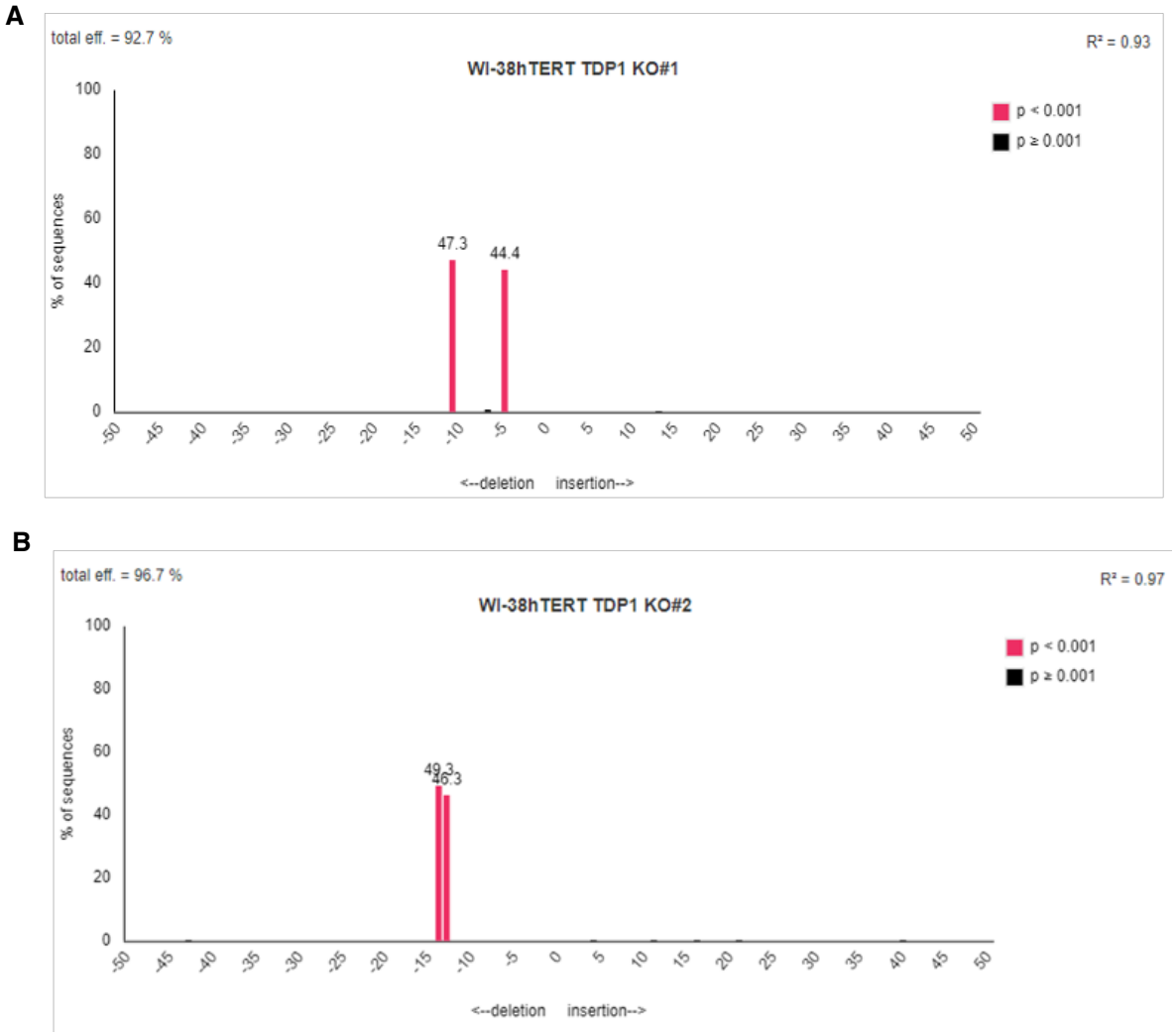

Figure 19. TIDE analysis of WI38hTERT TDP1 KO. Indel spectrum of the selected WI38 hTERT TDP1 KO#1(A) and KO#2(B). Graph representing, on the y-axis, the % of different nucleotide insertion and deletion that occurred at the Cas9 break site. On the x-axis, –number or +number indicate the number of nucleotides deleted or inserted of at the Cas9 break site, respectively. Red bar indicates events that are significantly detected, the P-value threshold, default is p<0.001.

3.2. Generation of U2OS TDP1 KO and H493R cell lines

We used human bone osteosarcoma epithelial U2OS cells because it is a well-established model for genome editing by CRISPR-Cas9. To generate the U2OS TDP1 KO and H493R cell lines, we used the strategy illustrated in Figure 20. To generate the TDP1 KO cell lines was targeted the exon 5 of *TDP1* and to generate the H493R cell lines the exon 14 (Figure 20A). A feature of this strategy used is the co-expression of gRNA targeting the protospacer on the *ATP1A1* gene (Agudelo *et al.*, 2017). Transfection or electroporation performed in the presence of a donor (ATP1A1-RD) with homology arms for the *ATP1A1* targeted site and gain-of-function

mutations, which determine the replacement of the residues Q118R and N129D on the ouabain (plant-derived inhibitor of the Na^+/K^+ ATPase) binding region, allows cells to be resistant to ouabain after that HR repair has occurred (Agudelo *et al.*, 2017). To generate the U2OS TDP1 KO, the eSpCas9 (1.1) No FLAG ATP1A1 G3 Dual sgRNA plasmid was engineered to allow the expression of the Cas9 and the gRNA for *ATP1A1* along with the gRNA targeting the protospacer we selected on the exon 5 of *TDP1* gene. So, we co-transfected the engineered plasmid in the presence of a donor with homology arms for the targeted exon 5 and a single point mutation to insert a STOP codon by HR after the Cas9 cut. Cells were kept under ouabain selection (0.7 uM) for 5 days until all the control cells transfected in the absence of the ATP1A1-RD donor died. Survivors' colonies were isolated into a 24-well plate and then further amplified into larger 6-well plates. Stable clones were then grown in duplicate, harvested, and further processed for analysis at protein and DNA levels.

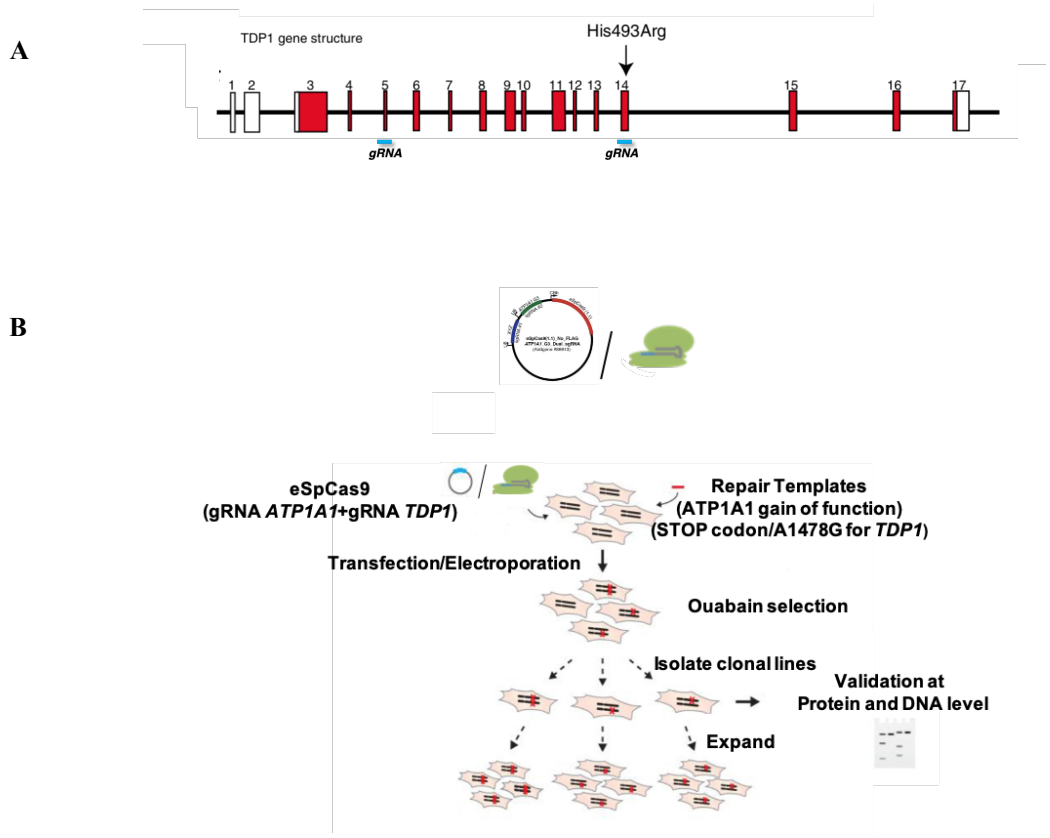

Figure 20 (A-B) Schematic representation of the experimental design to generate U20S TDP1 KO and U2OS H493R cell lines. (A) Schematic representation of the TDP1 gene structure. Exons are represented as boxes and introns as lines, in blue is indicated the gRNA targeting the exon 5, the arrow indicates the H493R mutation. Adapted from (Takashima *et al.*, 2002); (B) Schematic representation of the procedure used. Adapted from (Ran *et al.*, 2013).

This strategy turned out to be very efficient as 5 out of the 6 clones analyzed showed undetectable levels of TDP1 protein by WB. Following the analysis at the DNA level, 4 clones out of 5 resulted to be TDP1 KO because of indels creation that disrupted the open reading frame after NHEJ repair and one resulted to be KO because of the insertion of the stop codon following HR repair (data not shown). So, we selected 2 clones out of the 5 obtained, TDP1 KO#1 and KO#2, to pursue our analysis. Whole cell protein extracts were immunoblotted against TDP1. Both TDP1 KO#1 and KO#2 clones presented an undetectable level of TDP1 protein compared to the wild-type U2OS cell line (Figure 21A). Besides, cells were harvested for DNA extraction to perform PCR amplification of the TDP1 exon 5. The purified amplicons were then sent out for sequencing. The analysis of chromatograms confirmed that, compared to the parental U2OS cell line, both TDP1 KO#1 and TDP1 KO#2 underwent indel events at the expected Cas9 cut site (Figure 21B), causing the disruption of the gene reading frame. By TIDE analysis clone TDP1 KO#1 presented a high total efficiency of ~97% and a deletion of 14 and 29 nucleotides from the Cas9 cut site, not significant to give rise to a truncated protein (Figure 21C). Regarding the clone U2OS TDP1 KO#2 the software was not able to generate an *indel identification* graph. This decomposition error of the software can happen when indels after the Cas9 cut occur too close to the beginning of the sequencing start site. To avoid this problem, it is suggested to amplify a bigger region that includes the targeted site to allow the software to have a bigger window of decomposition to analyze. We used the same strategy to generate the U2OS H493R TDP1 cell lines. At first, we designed the gRNA targeting the exon 14, where the mutation that causes SCAN1 occurs (Takashima *et al.*, 2002), and a ssODNs donor (Paix *et al.*, 2017) with homology arms for the targeted site harboring the A1478G mutation. To increase the editing, we performed two protocols in parallel. Besides the plasmid engineering and transfection described above, we directly deliver in the cells, ribonucleoproteins (RNPs) consisting of the Cas9 proteins in complex with the targeting gRNAs, in the presence of the ssODNs. After selection, we isolated 10 clones ouabain resistant obtained from the transfection and 10 clones ouabain resistant obtained from the electroporation. These were grown in duplicate, harvested, and then processed for analysis at the DNA level.

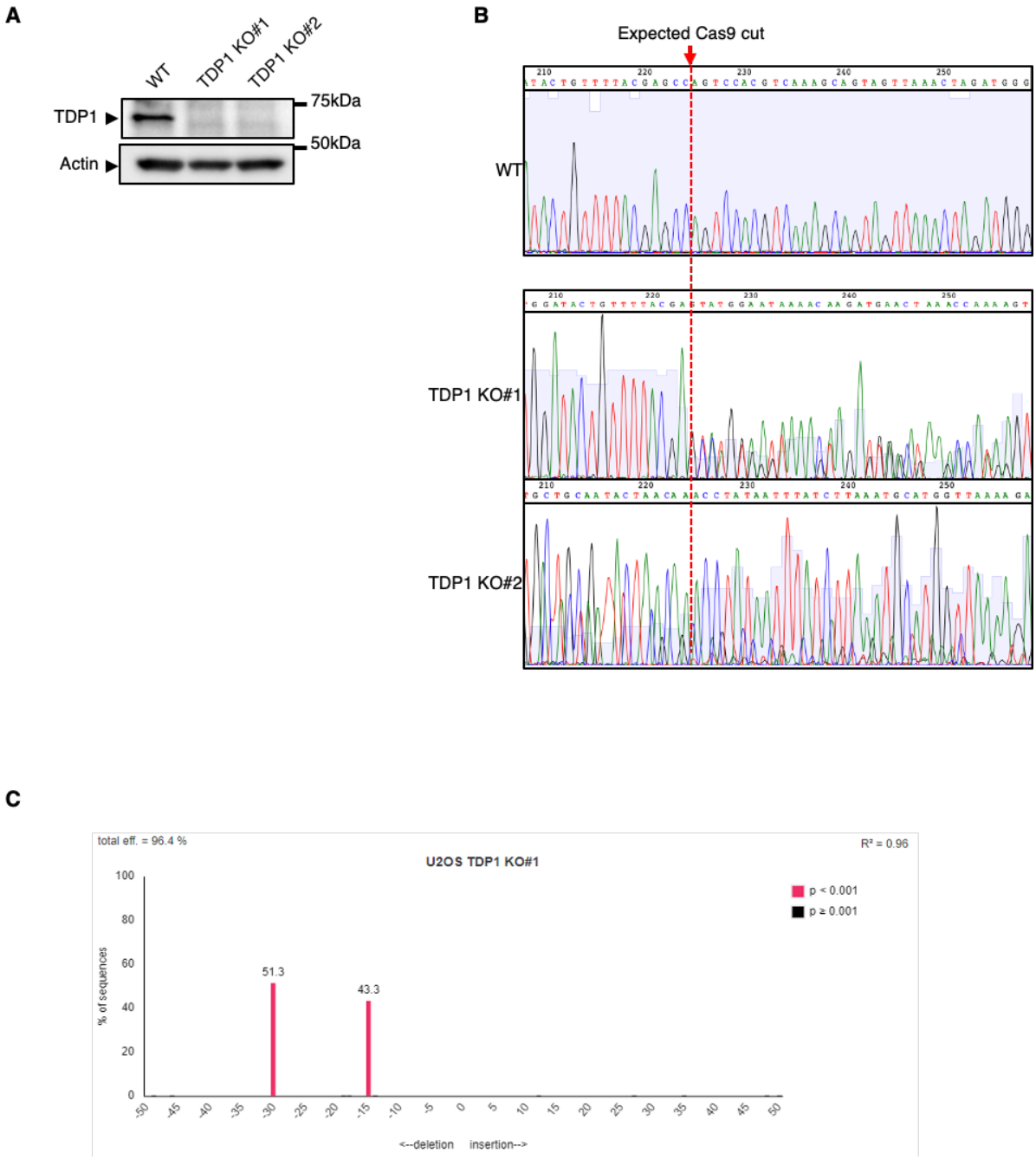

Figure 21. Validation at the protein and DNA level of U2OS TDP1 KO. (A) WB analysis of TDP1 levels. Whole cell protein extracts were immunoblotted against TDP1. Actin was used as loading control; (B) Sequencing analysis of the TDP1 locus in the U20S TDP1 clones. Chromatogram showing the genome editing by CRISPR/Cas9 system. Genomic DNA of the parental U2OS cell line and the TDP1 KO clones were used as a template in PCR with primers that amplify at the level of the gRNA targeted TDP1 Exon 5. The purified amplicons were sent to sequencing. The red arrow and the red dashes line indicate the expected Cas9 cut site; (C) Indel spectrum of the selected U2OS TDP1 KO#1. Graph representing, on the y-axis, the % of different nucleotide insertion and deletion that occurred at the Cas9 break site. On the x-axis, –number or +number indicate the number of nucleotides deleted or inserted at the Cas9 break site, respectively. The red bar indicates events that are significantly detected, the P-value threshold, default is p<0.001.

Analyzing the chromatograms, we detected that 2 out of the 20 clones screened, clone H493R#1 (obtained by electroporation), and clone H493R#2 (obtained by transfection), correctly inserted the ssODSNs through homologous recombination after the Cas9 cut. This gave rise to the A1478G single point mutation insertion and the consequently H493R substitution at the protein level (Figure 22A). To assess whether the A1478G TDP1 mutation occurred in both alleles, we took advantage of the fact that this mutation creates a single BsaA1 restriction site in *TDP1*. Thus, we amplified the *TDP1* exon 14 by PCR and we digested the PCR products with the BsaA1 restriction enzyme. In Figure 22B, we showed that this digestion yielded an uncut PCR product of 615 bp in the WT. PCR products of both H493R#1 and H493R#2 have been completely cut into restriction fragments of the expected size (357 and 229 bp). This result supports that both clones are homozygous for TDP1 H493R gene mutation. To further confirm that the insertion occurred in frame we performed the TIDE analysis. This latter informed us that both clones H493R#1 and H493R#2 did not present any insertion or deletion that occurred at the Cas9 break site (Figure 23A-B). We concluded that they underwent the HR repair after the cut and the insertion occurred in frame. Then, H493R#1 and H493R#2 were harvested and processed for analysis at the protein level. In Figure 24A, our data shows that TDP1 protein in both clones presented approximately a 5-fold reduction in the amount of the expressed protein relative to wild-type (WT) cells. We could exclude that this reduction could be linked to the trapping of H493R TDP1 due to the formation of persistent H493R TDP1–DNA covalent complexes (Interthal *et al.*, 2005) because the WB was performed with sonicated SDS/Tris WCE extracts. This protocol allows the migration of TOP1cc on gels (Cristini *et al.*, 2019). In addition, we carried out the WB without removing the stacking gel before the transfer of the proteins into the membrane and we did not detect TDP1 on the top part of the immunoblot (Figure 24B).

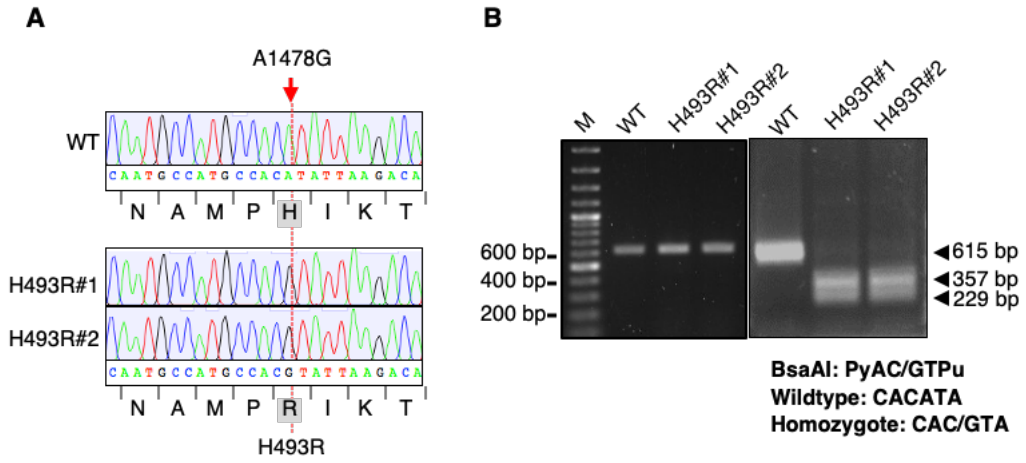

Figure 22. Validation at the DNA level of the U2OS H493R TDP1. (A) Sequencing analysis of the TDP1 exon 14 in the U2OS H493R-TDP1 clones. Chromatogram showing the genome editing by CRISPR/Cas9 system. The red arrow and the red dashes line indicate the expected single point mutation (A1478G) occurred through homologous recombination after the Cas9 cut; (B) BsaA1 restriction analysis of the A1478G nucleotide mutation. On the left, the agarose gel shows the TDP1 Exon 14 amplicons (615 bp). On the right, the agarose gel shows the BsaA1 digestion products (357 bp, 229 bp).

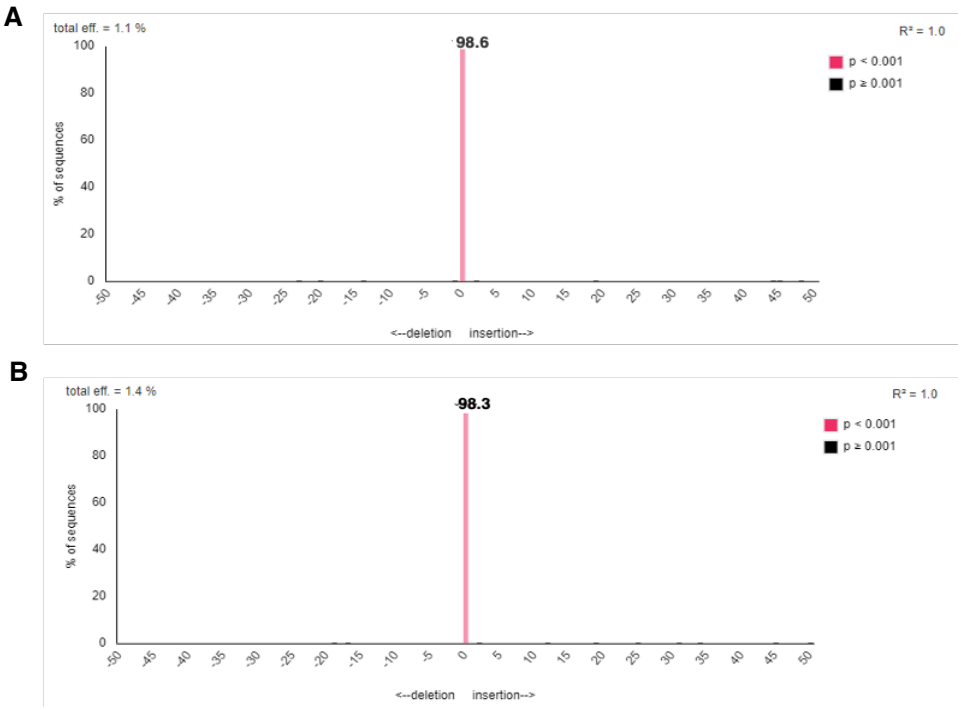

Figure 23. TIDE analysis. Indel spectrum of the selected H493RTDP1 clones (A) H493R#1; (B) H493R#2. Graph representing, on the y-axis, the % of different nucleotide insertion and deletion that occurred at the Cas9 break site. On the x-axis, –number or +number indicate the number of nucleotides deleted or inserted at the Cas9 break site, respectively. The red bar indicates events that are significantly detected, the P-value threshold, default is $p<0.001$.

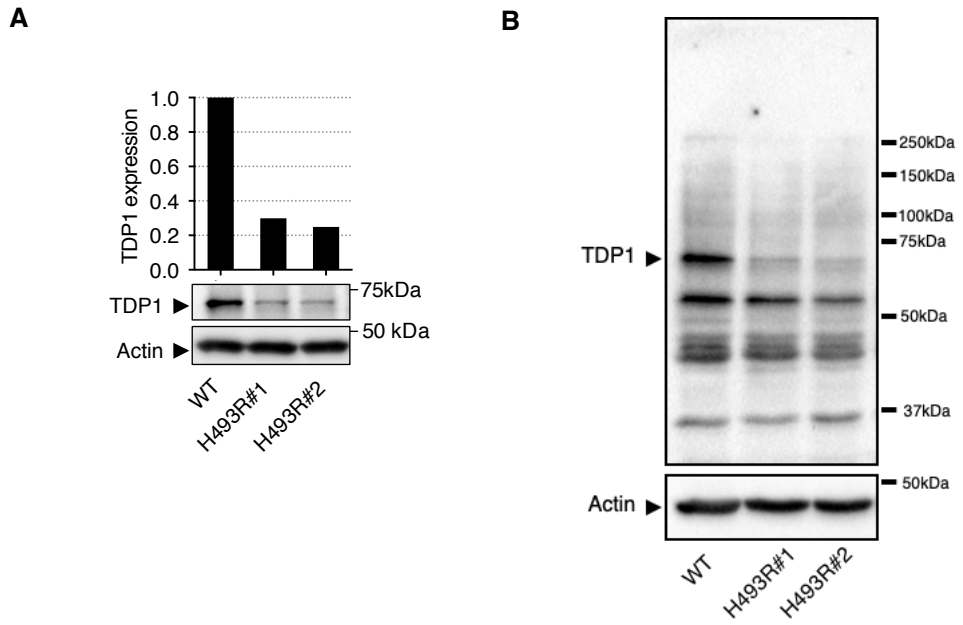

Figure 24. Validation at the protein level of the U2OS H493R TDP1. (A and B) WB analysis of TDP1 levels. Whole cell protein extracts were immunoblotted against TDP1. Actin was used as loading control; (A) Quantification of TDP1 levels relative to actin.

3.3. TDP1 depletion and TDP1 SCAN1 mutation (H493R) increases DSBs outside of the S phase in U2OS cells.

Because DSBs are lethal lesions that can cause neurodegenerative disorder and that the SCAN1 is a neurodegenerative disease that involves post-mitotic neurons (Takashima *et al.*, 2002), we investigated DSBs in H493R TDP1 and TDP1 KO cells we have generated, focusing our attention in non-replicating cells. U2OS cells can't enter quiescence; hence, we analyzed DSBs in G1, prior to DNA replication. To investigate this, U2OS WT, TDP1 KO, and H493R TDP1 cells were incubated with EdU to label newly synthesized DNA before staining for γH2AX (phosphorylated H2AX at S139; a marker of DSBs (William M Bonner *et al.*, 2008)) and Hoechst 33342 (DNA). G1 phase cells were discriminated from the S phase cells by the lack of EdU incorporation into DNA and from G2 cells by low Hoechst 33342 signal (Figure 25). DSBs can be detected by microscopy as nuclear foci containing γH2AX, with a single γH2AX focus reflecting hundreds to thousands of γH2AX proteins that are concentrated around at least one DSB (William M Bonner *et al.*, 2008).

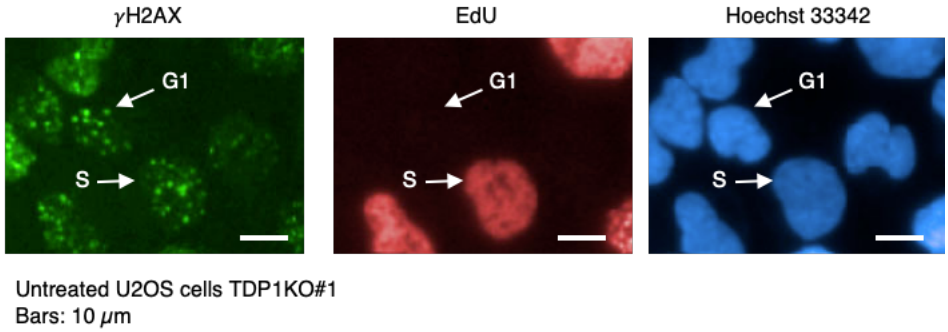

Figure 25. Immunofluorescence microscopy images of untreated cells. U2OS TDP1KO cells untreated incubated with 10 μM EdU for 30 min, co-stained for γH2AX (green), and Hoechst33342 (blue). The G1-phase cells are labeled as G1 and the S-phase cells as S. Bars:10 μm.

Our data showed that both TDP1 KO clones presented a higher number of γH2AX nuclear foci than WT cells in the G1-phase compared to the S-phase cells (Figure 26A, B). In Figure 26C, D our data showed that also TDP1 H493R clones presented the same trend of the TDP1 KO clones. Despite the clone H493R#1 showed a slight increase in γH2AX nuclear foci than WT cells in the G1-phase compared to the H493R#2, we could conclude that results of panel D are similarly consistent with results of panel B.

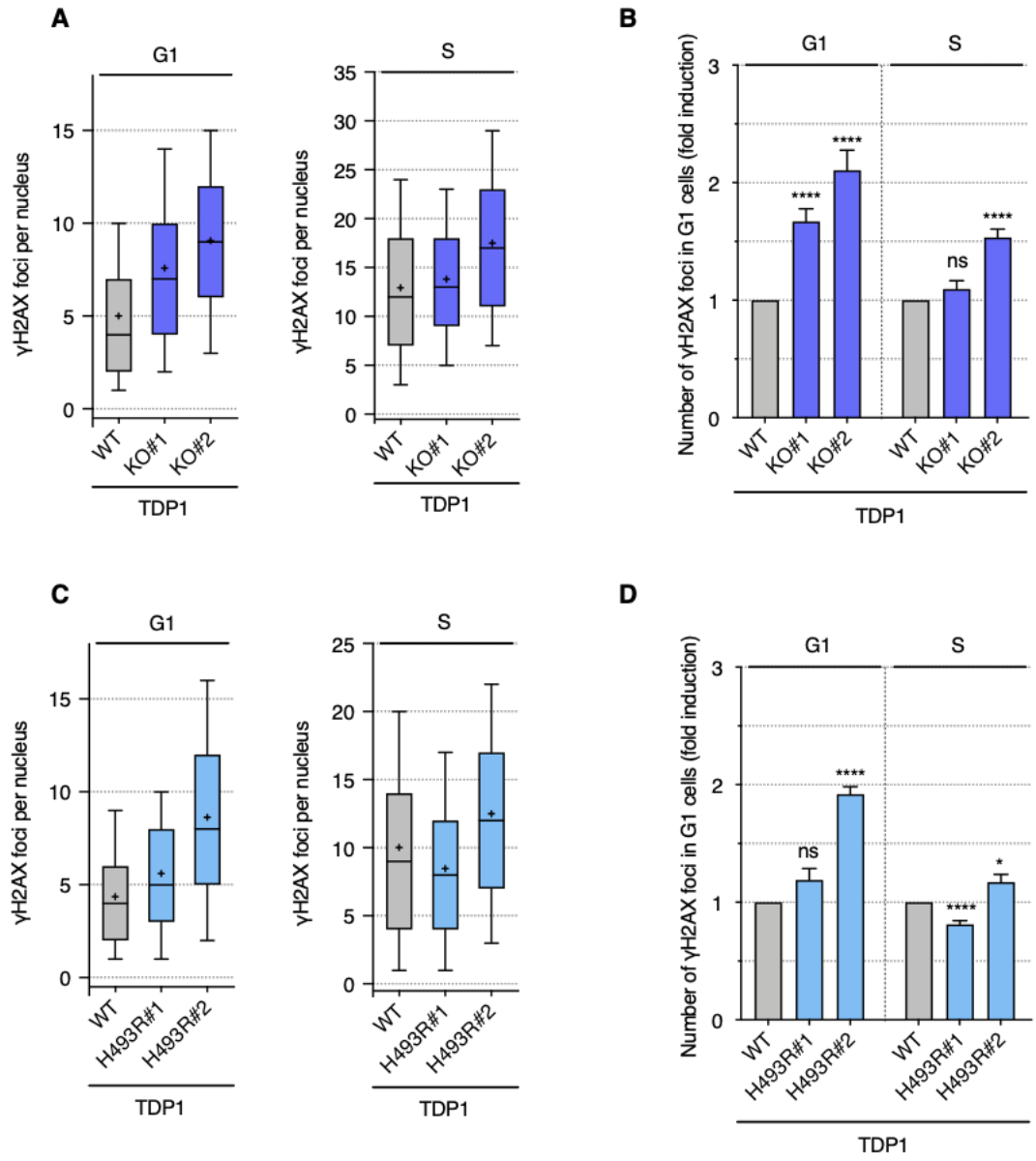

Figure 26. TDP1 depletion and TDP1 SCAN1 mutation (H493R) increases DSBs outside of the S phase in U2OS cells. WT, TDP1 KO, and TDP1 H493R U2OS cells were incubated with 10 μM EdU for 30 min before staining for γH2AX and Hoechst 33342 (DNA). (A and C) The number of γH2AX foci per G1 nucleus (EdU-negative and low Hoechst 33342) and S nucleus (EdU-positive) in a representative experiment out of ≥ 4 (A) or 7 (C) is shown. (B and D) The fold induction of γH2AX was calculated by normalizing to the WT cells. Means ± SEM; n ≥ 4 (B); n = 7 (D). Ns: not significant, *p < 0.05, ****p < 0.0001 (two-tailed unpaired t-test).

3.4. Deficiency of TDP1 increases DSBs outside of S phase and micronuclei in WI38 hTERT cells.

WI38 hTERT can be induced in quiescence following serum deprivation (Dimri, Hara and Campisi, 1994; Coller, Sang and Roberts, 2006), thus allowing the analysis of replication-

independent damage. However, we observed that upon serum deprivation, TDP1 KO clones progressively died making it impossible to continue the experiment. A similar feature was observed in *S. Pombe* where TDP1 KO yeast cells lose viability upon quiescence induction (Ben Hassine and Arcangioli, 2009). It has been proposed that this phenotype could be related to the accumulation of unrepaired 3'-phosphoglycolates resulting from reactive oxygen species (Ben Hassine and Arcangioli, 2009). Thus, we analyzed cells in G1, following the same protocol used for the U2OS cell lines. Cells were incubated with EdU before staining for γH2AX, p53BP1 (phosphorylated 53BP1 at S1778), and Hoechst 33342 (DNA). DSBs were then detected by microscopy as nuclear foci containing γH2AX and p53BP1. In figures 27A, B, C, and D, our data showed that the depletion of TDP1 increased DSBs outside of the S phase in WI38h TERT cells. No analysis of S-phase cells for p53BP1 (Figure 27 C, D) could be performed because when we co-labeled the cells with γH2AX and with p53BP1, we choose a secondary antibody associate with a green fluorophore for γH2AX and a red for p53BP1. The red fluorescence of p53BP1 overlaps with the far-red fluorescence of the EdU (p53BP1 594nm, EdU 647nm). In those experiments, it was only possible to measure p53BP1 foci in cells out of the S-phase. So, we planned to redo these experiments labeling the p53BP1 foci in green to be able to analyze them in the S-phase. We further counted the number of micronuclei (MN) which can be detected by immunofluorescence after cells staining for Hoechst 33342. MN are markers of genomic instability. They are extra-nuclear bodies that contain damaged chromosomes and/or chromosome fragments that were not incorporated into the nucleus after cell division (Luzhna, Kathiria and Kovalchuk, 2013). Our data showed that both WI38 hTERT TDP1KO clones have a higher percentage of cells with micronuclei compared to the WT (Figure 27E). An increase of γH2AX foci has been reported in WI38 hTERT quiescent cells siRNA for TDP1 (Cristini *et al.*, 2016). This suggests that the γH2AX foci observed in G1 are unlikely to simply results from the previous cell cycle. We decided then to use this model to directly look at DSBs by neutral comet assay, which is not possible to perform in cycling cells because this technique does not allow us to discriminate the phases of the cell cycle. Thus, we used quiescent WI38 hTERT cells siRNA for TDP1, resulting in less toxicity compared to our TDP1 KO cells (Cristini *et al.*, 2016, 2019). We then performed the neutral comet assay. The efficiency of the knockdown was controlled by WB (Figure 28A). Our data suggested that TDP1 deficient quiescent cells increase replication-independent DSBs, correlated with increased neutral comet tail moments (Figures 28B and 28C).

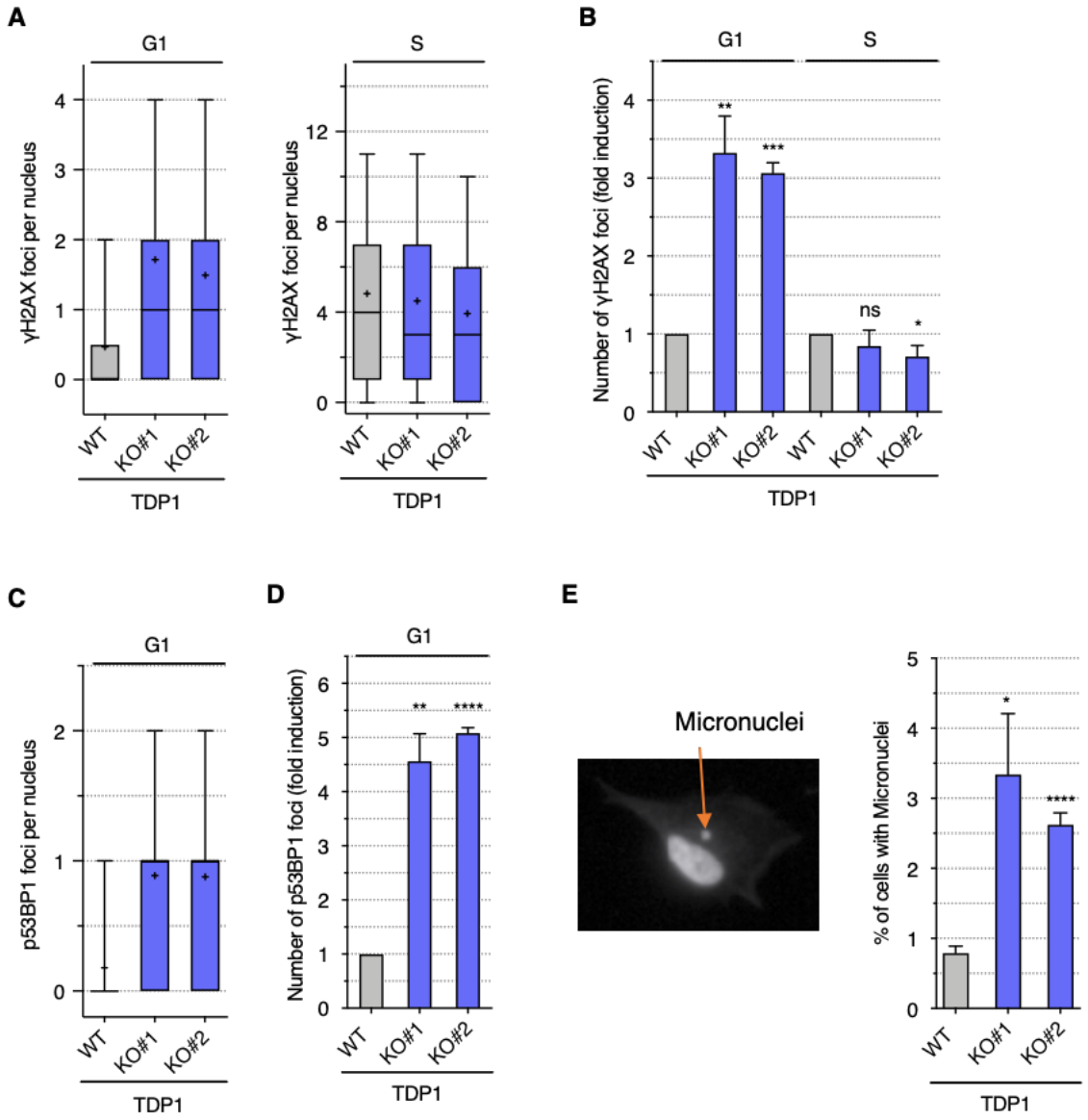

Figure 27. Depletion of TDP1 increases DSBs outside of S phase and micronuclei in WI38 hTERT cells. (A and B) WT and TDP1 KO WI38 hTERT cells were incubated with 10 µM EdU for 30 min before staining for γH2AX and Hoechst 33342 (DNA). (A) The number of γH2AX foci per G1 nucleus (EdU-negative and low Hoechst 33342) and S nucleus (EdU-positive) in a representative experiment out of 3 is shown. (B) The fold induction of γH2AX was calculated by normalizing to the WT cells (means ± SEM; n = 3). (C and D) WT and TDP1 KO WI38 hTERT cells were treated with EdU and analyzed as in (A and B) except that they were stained for p53BP1. (C) Representative experiment out of 3. (D) Results are shown as means ± SEM; n = 3. (E) Detection of micronuclei. Left panel: representative image of a micronuclei. Right panel: quantification of the number of micronuclei in WT and TDP1 KO WI38 hTERT cells. Results are shown as means ± SEM; n ≥ 3. Ns: not significant, *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001 (two-tailed unpaired t-test).

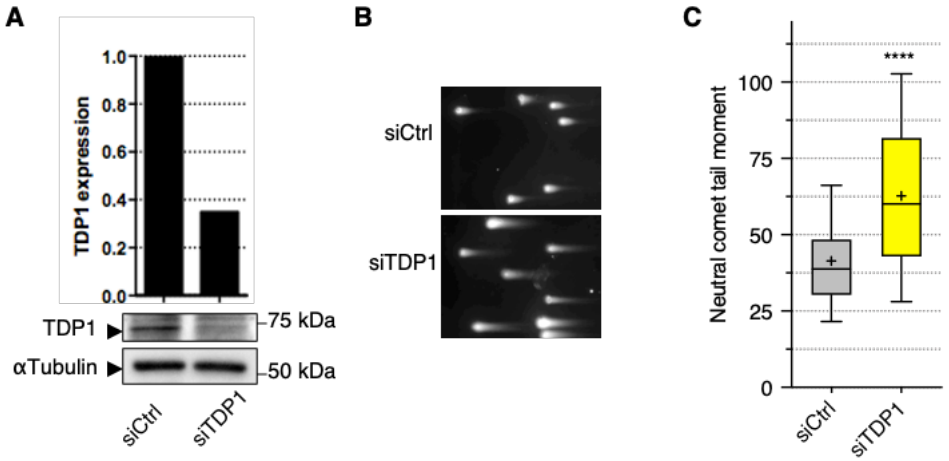

Figure 28. Deficiency of TDP1 increases DSBs in quiescent WI38 hTERT cells. Detection of DSBs by neutral comet assays in quiescent WI38 hTERT cells deficient for TDP1. Cells were transfected with siRNAs against TDP1 (siTDP1) or a control sequence (siCtrl) for 24 h before being cultured in 0.2% serum for 72 h to induce quiescence. (A) Western blot probed with TDP1 antibody. α-tubulin: loading control. Quantification of TDP1 levels relative to α-tubulin. (B) Representative pictures of nuclei. (C) Quantification of neutral comet tail moments in one representative experiment out of 2. ****$p < 0.0001$ (two-tailed unpaired t-test).

3.5. Depletion of TDP1 increases DSBs outside of S phase and micronuclei in HCT116 cells.

In addition to our TDP1 KO cell lines, we used the HCT116 WT and TDP1 KO, generated and kindly provided by Y. Pommier laboratory (Al Abo *et al.*, 2017), as a cell model, to further strengthen our studies. Thus, we analyzed cells in G1, following the same protocol used for the U2OS and the WI38 hTERT cell lines. Cells were incubated with EdU before staining for γH2AX, p53BP1 (phosphorylated 53BP1 at S1778), and Hoechst 33342 (DNA). DSBs were then detected by microscopy as nuclear foci containing γH2AX and p53BP1. In figures 29A, B, C, and D our data showed that HCT116 TDP1 KO increased DSBs outside of the S. In figure 29E we reported that HCT116 TDP1KO clones have also a higher percentage of cells with micronuclei compared to the WT. So, we can conclude that the depletion of TDP1 increases DSBs outside of S phase and micronuclei also in HCT116 cells.

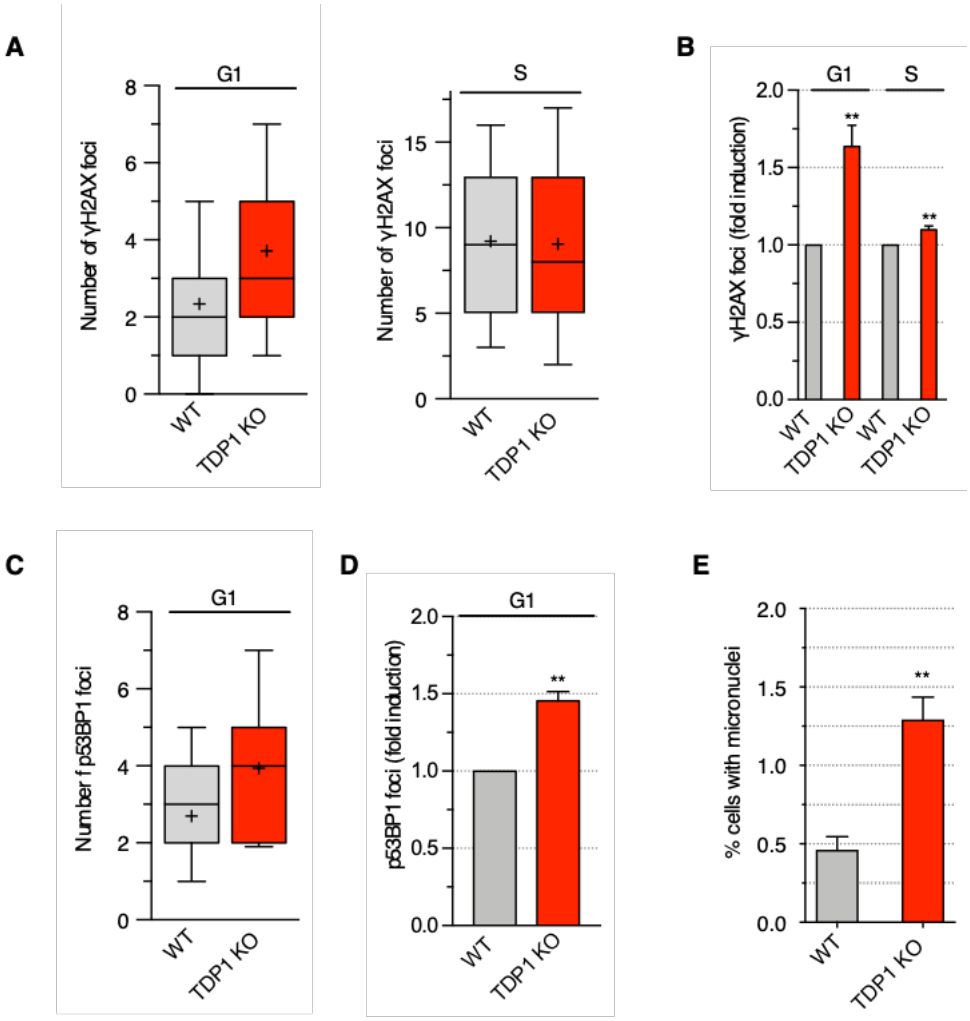

Figure 29. Depletion of TDP1 increases DSBs outside of S phase and micronuclei in HCT116 cells. (A and B) WT and TDP1 KO HCT116 cells were incubated with 10 μM EdU for 30 min before staining for γH2AX and Hoechst 33342 (DNA). (A) The number of γH2AX foci per G1 nucleus (EdU-negative and low Hoechst 33342) and S nucleus (EdU-positive) in a representative experiment out of 3 is shown. (B) The fold induction of γH2AX was calculated by normalizing to the WT cells (means ± SEM; n = 3). (C and D) WT and TDP1 KO HCT116 cells were treated with EdU and analyzed as in (A and B) except that they were stained for p53BP1. (C) Representative experiment out of 3. (D) Results are shown as means ± SEM; n = 3. (E) Quantification of the number of micronuclei in WT and TDP1 KO HCT116 cells. Results are shown as means ± SEM; n = 3. **$p < 0.01$, (two-tailed unpaired t-test).

3.6. TDP1 deficiency modulates the levels of R-loops at some loci in WI38 hTERT cells

Next, we asked whether DSBs accumulation in TDP1-deficient cells could be related to an increase in DSBs production. Notably, R-loop structures that form co-transcriptionally can induce DSBs in non-replicating cells (Cristini *et al.*, 2019). Thus, we tested whether TDP1 deficiency would increase R-loop levels. We examined R-loops in WI38 hTERT WT and TDP1 KO by DNA/RNA immunoprecipitation (DRIP) using S9.6 antibody (Boguslawski *et al.*, 1986) and examined R-loops at some gene loci by qPCR using primers designed at transcription starting site (TSS)-proximal regions of β-ACTIN and γ-ACTIN, at gene body regions of PTB, GEMIN7, β-ACTIN, γ-ACTIN and EGR genes, and at and in an intergenic region as a negative control (region upstream of the TSS of β-ACTIN gene) (see Methods for the sequences of the primers). We decided to investigate these regions because it was shown that in quiescent WI38 hTERT cells, the trapping of TOP1cc modulates the distribution of R-loops at these loci (Cristini et al. 2019), and TDP1-deficient cells are defective in the removal of TOP1cc (Pommier, 2006). Our data (Figure 30) show that TDP1 deficiency induces a drop of DRIP signal at the β-ACTIN and γ-ACTIN TSS and an increase in the gene body of PTB and GEMIN 7 genes. Our result highlight that TDP1 deficiency causes a genomic redistribution of R-loops at some genes. This could suggest a potential involvement of R-loops in the mechanism of DSBs production.

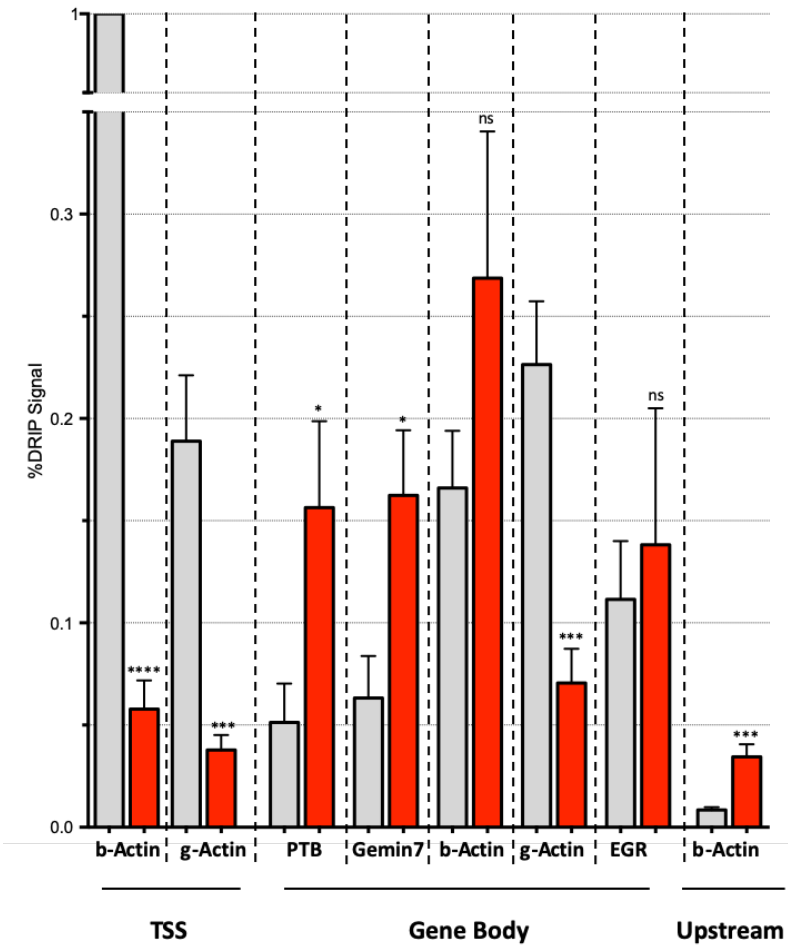

Figure 30. DRIP analysis in WI38 hTERT cells. Cells were harvested for genomic DNA extraction. R-loops were immunoprecipitated by DRIP. And then the level of R-loops was analyzed at TSS-proximal regions of β-ACTIN and γ-ACTIN genes, gene body regions of PTB, GEMIN7, β-ACTIN, γ-ACTIN and EGR genes, and upstream region of β-ACTIN gene by qPCR using the primers indicated in the Methods and material section. Values are normalized to b-actin ‘‘TSS proximal’’ amplicon (means ± SEM; n = 3). *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001 (two-tailed unpaired t-test).

3.7. TDP1 depletion and TDP1 SCAN1 mutation (H493R) increase CPT-induced DSBs primarily in G1 phase of U2OS cells.

TDP1 processes various 3’-end-blocking lesions besides trapped TOP1ccs, such as 3’-phosphoglycolates (Inamdar *et al.*, 2002; Interthal, Chen and Champoux, 2005; Zhou *et al.*, 2005). To examine the potential implication of TOP1cc, we took advantage of the fact that CPT selectively traps TOP1cc (Pommier *et al.*, 2003; Pommier, 2006). Also, there is evidence that TDP1-depleted cells accumulate TOP1 peptide-linked SSB intermediates (El-Khamisy *et al.*, 2005; Interthal *et al.*, 2005) and transcriptional DSBs in response to CPT (Cristini *et al.*, 2016, 2019). Moreover, both TDP1 depletion and H493R mutation confers hypersensitivity to this

drug (El-Khamisy *et al.*, 2005; Interthal *et al.*, 2005; Miao *et al.*, 2006; Gao *et al.*, 2014). Thus, WT, TDP1 KO, and H493R TDP1 U2OS cells were incubated with EdU before treatment with CPT and stained for γH2AX and Hoechst 33342 (DNA). We counted the number of γH2AX foci per G1 nucleus and measured the intensity of γH2AX staining per S nucleus. Concerning the S phase cells, we used the intensity of γH2AX staining as readout of DSBs formation in CPT-treated cells because the number of γH2AX foci is too high and hence the staining is rather diffuse, which does not allow to distinguish individual γH2AX foci and to further reliably count them (Figure 31). In figures 32A, B, we observed that depletion (KO) or mutation (H493R) of TDP1 increased the induction of γH2AX foci in the G1 phase of CPT-treated U2OS cells. TDP1 KO#1 showed a ~2-fold increase in the number of γH2AX foci and KO#2, H493R#1, and H493R#2 ~1.5-fold increase compared to the WT (Figure 32B). Conversely, in S phase, TDP1KO#1 and H493R#1 did not show an increase in γH2AX intensity per nucleus while TDP1 KO#2 and H493R#2 did (Figure 32C, D); hence it was not possible to conclude about the role of TDP1 in S phase in U2OS cells. To assess whether these differences between clones might be related to the high dose effect of CPT, it would be interesting to repeat the experiment in the presence of lower doses of CPT. In any case, our data indicate that TDP1 depletion and TDP1 SCAN1 mutation (H493R) increase CPT-induced DSBs in G1 phase of U2OS cells.

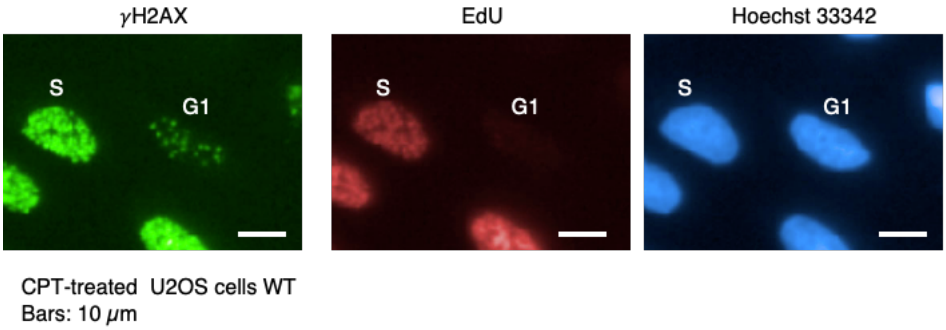

Figure 31. Immunofluorescence microscopy imagines of CPT treated cells. U2OS WT cells treated with CPT (25 mM; 1 h), after 30min of EdU incubation, co-stained for γH2AX (green) and Hoechst33342 (blue). The G1-phase cells are labeled as G1 and the S-phase cells as S. Bars:10 μm.

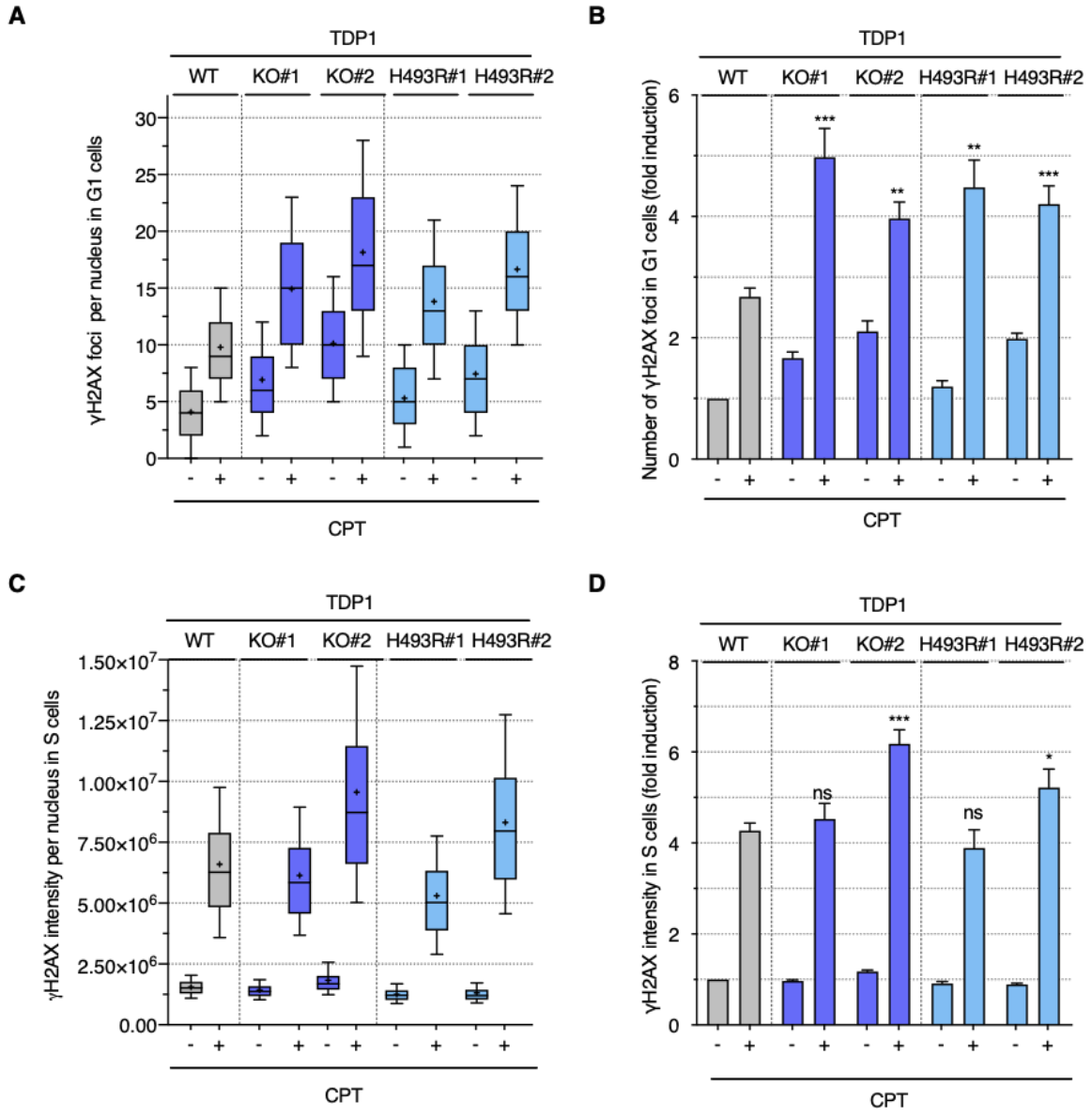

Figure 32. TDP1 depletion and TDP1 SCAN1 mutation (H493R) increase CPT-induced DSBs primarily in G1 phase of U2OS cells. WT, TDP1 KO, and TDP1 H493R U2OS cells were incubated with 10 μM EdU for 30 min before treatment with CPT (25 μM; 1 h) and stained for γH2AX and Hoechst 33342 (DNA). (A and C) The number of γH2AX foci per G1 nucleus (EdU-negative and low Hoechst 33342) (A) and the intensity of γH2AX staining per S nucleus (EdU-positive) (C) in a representative experiment out of ≥ 4 (A) or 7 (C) is shown. (B and D) The fold induction of γH2AX was calculated by subtracting the number of foci (B) or the intensity of staining (D) of WT cells from that of TDP1 KO cells (B) or TDP1 H493R cells (D) and normalized to WT cells. Means ± SEM; n ≥ 4 (B); n = 7 (D). Ns: not significant, *p < 0.05, **p < 0.01, ***p < 0.001 (two-tailed unpaired t-test).

3.8. Depletion of TDP1 increases CPT-induced DSBs in G1 and S phases of WI38 hTERT cells.

To investigate whether we could detect the same increase of CPT-induced DSBs in the WI38 hTERT TDP1 KO cell lines, we performed the same experiment described in section 4.7. Looking at the G1 phase cells (Figure 33A, B) TDP1 deficiency increased significantly the number of γH2AX per nucleus compared to the WT. In figure 33B, TDP1 KO#1 and KO#2, in response to CPT, showed a ~2.5-fold and ~1.5-fold increase in the number of γH2AX foci, respectively, compared to the WT. In figure 33C, D, our data showed that TDP1 deficiency also increased the γH2AX intensity per nucleus in response to CPT in S phase. Thus, we conclude that, in response to CPT, WI38 hTERT TDP1KO cells increase DSBs both in G1 and S phase.

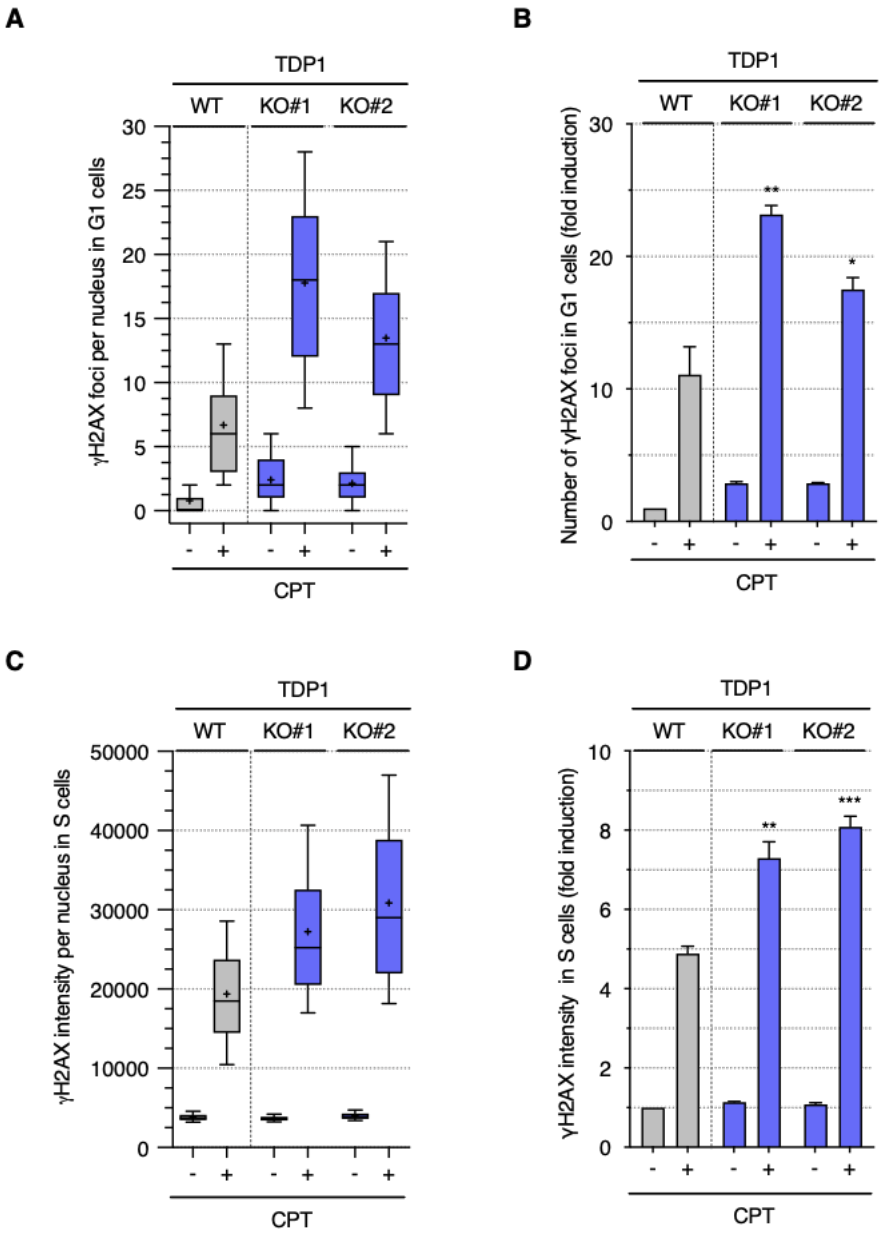

Figure 33. Depletion of TDP1 increases CPT-induced DSBs in G1 and S phases of WI38 hTERT cells. WT and TDP1 KO WI38 hTERT cells were incubated with 10 µM EdU for 30 min before treatment with CPT (25 µM; 1 h) and stained for gH2AX and Hoechst 33342 (DNA). (A and C) The number of γH2AX foci per G1 nucleus (EdU-negative and low Hoechst 33342) (A) and the intensity of γH2AX staining per S nucleus (EdU-positive) (C) in a representative experiment out of 3 is shown. (B and D) The fold induction of γH2AX was calculated by subtracting the number of foci (B) or the intensity of staining (D) of WT cells from that of TDP1 KO cells and normalized to WT cells. Means ± SEM; n = 3. Ns: not significant, *p < 0.05, **p < 0.01, ***p < 0.001 (two-tailed unpaired t-test).

3.9. TDP1 depletion and TDP1 SCAN1 mutation (H493R) prevent the repair of CPT-induced DSBs in G1 phase of U2OS cells

Next, we examined whether the accumulation of DSBs in TDP1 deficient (KO) and mutated (H493R) cells, which occurs primarily in G1 in untreated and CPT-treated cells (see Figure 34), could be related to a defect in the repair of those breaks. To test this, we analyzed the reversal kinetics of γH2AX (see protocol in Figure 34A). Cells were incubated with EdU before treatment with CPT, washed free of the drug (W), and cultured in the presence of EdU for up to 6 h (release; R). Unlike continuous exposure to CPT, this protocol allows us to study DSB repair, as TOP1cc reverse fully within minutes after washing out CPT (Pommier *et al*., 2016), and then DSBs are no longer produced. The use of EdU, which labeled newly synthesized DNA, further allows analyzing the repair kinetics in G1 (EdU negative) and S (EdU positive) phase of the cell cycle. In U2OS TDP1 WT, γH2AX reversed rapidly in both G1 phase (Figure 34B, C) and S phase (Figure 34D, E) after CPT removal. TDP1 depletion (KO) and TDP1 H493R mutation led to persistent γH2AX in G1 phase (Figure 34B, C) but not in S phase (Figure 34D, E), suggesting that TDP1 promotes the repair of CPT-induced DSBs specifically in G1 phase. Notably, TDP1 H493R mutation completely abrogated the reversal γH2AX in G1 up to 6 hours after CPT removal, while TDP1 KO only reduced it (Figure 34D, E). These results further suggest that TDP1 H493R mutation leads to unrepairable DSBs in response to CPT while backup repair pathways may exist in TDP1 depleted cells.

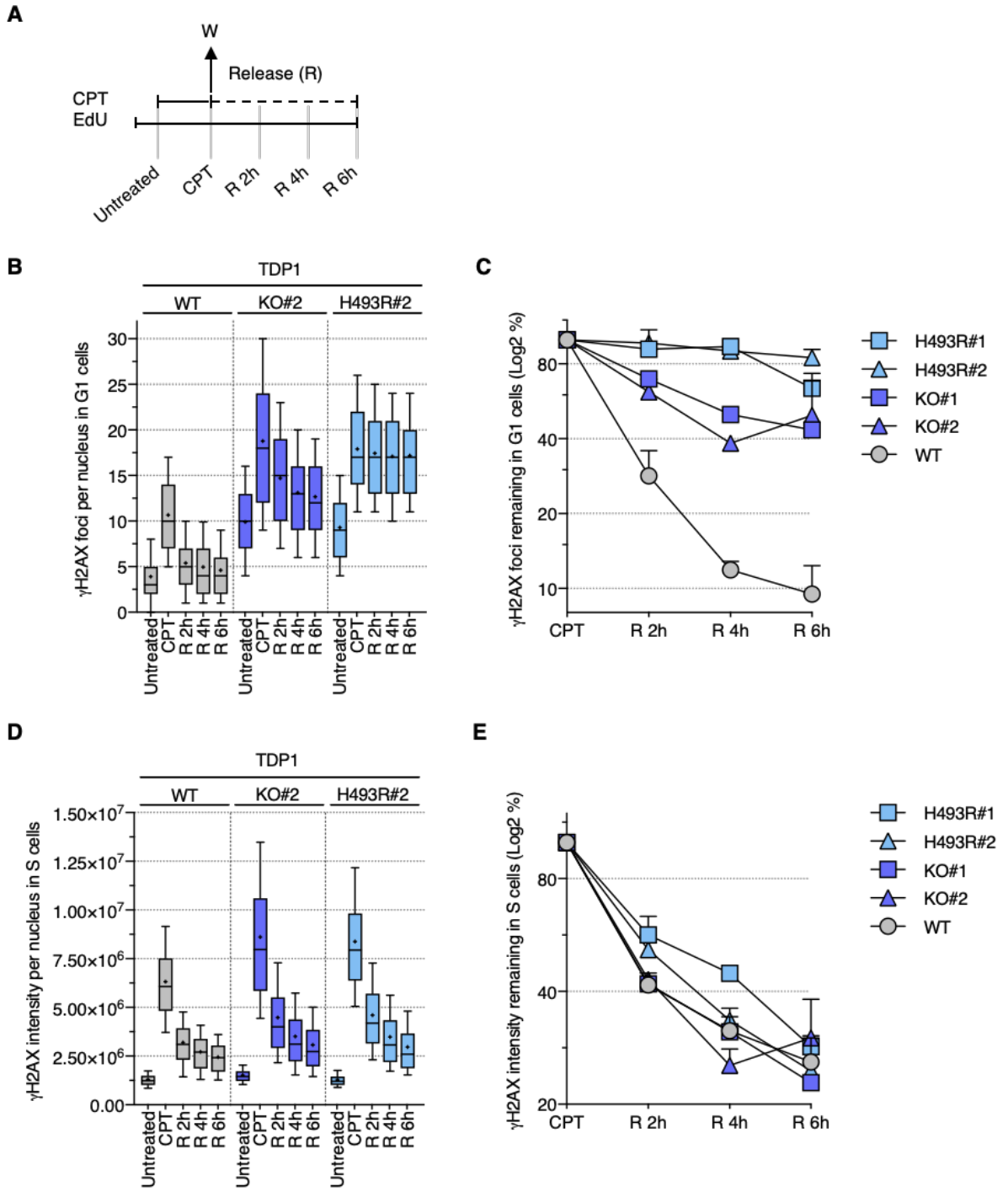

Figure 34. TDP1 depletion and TDP1 SCAN1 mutation (H493R) prevent the repair of CPT-induced DSBs in G1 phase of U2OS cells. (A) Protocol to study γH2AX reversal following CPT removal. Cells were incubated with 10 µM EdU for 30 min before treatment with CPT (25 µM; 1 h), washed (W), and cultured in CPT-free medium in the presence of 10 µM EdU for up to 6 h (release; R). Cells were then stained for γH2AX and Hoechst 33342 (DNA). (B and D) The number of γH2AX foci per G1 nucleus (EdU-negative and low Hoechst 33342) (B) and the intensity of γH2AX staining per S nucleus (EdU-positive) (D) in a representative experiment out of 2 is shown. (C and E) The percentages of γH2AX remaining following CPT removal were calculated by subtracting the number of foci (C) or the intensity of staining (E) of untreated cells from that of treated cells and normalized to cells treated with CPT. Means ± SEM; n = 2.

4. DISCUSSION

My PhD project has focused on the characterization of the genomic instability in non-replicating cells harboring the H493R mutation of TDP1, compared to cells depleted for TDP1. The interest in studying TDP1 mutations is linked to the fact that, even if the role in DNA repair of TDP1 is well known, how TDP1 H493R mutation promotes the SCAN1 phenotype is unclear. The SCAN1 syndrome is likely related to the selective death of post-mitotic neurons leading to the atrophy of the cerebellum.

My work provides new and stringent cellular models that have allowed a comparative study of TDP1 H493R mutation, gene deletion, and WT under the same genetic background. The results provide strong evidence that SCAN1 TDP1 triggers genome instability in human cells and DNA double-strand breaks mainly in G1 cells. The mechanism of such genome instability is likely related to an altered balance of R loop levels in the nuclear genome. The findings thus reveal new aspects of the molecular etiology of SCAN1 neurodegenerative syndrome in human patients. In the following section, I will discuss the findings in the context of literature and future prospective.

4.1. Why do we observe a decrease in the amount of H493R TDP1 protein?

A reduction in the amount of TDP1 H493R mutant protein was visible in the immunoblot analysis of our H493R cell extracts compared to WT (Figure 24A, B). This is consistent with similar observations were made in TDP1 H493R mutant cell extracts derived from SCAN1 patients (Interthal *et al.*, 2005), even though we observed a ~5-fold reduction whereas the reduction was less pronounced in SCAN1 patients cell extracts. We wondered if the reduction in WB could be due to persistent covalent crosslinks between the mutated TDP1 and genomic DNA, (Hirano *et al.*, 2007; Interthal *et al.*, 2005). Thus, we processed TDP1 H493R#1 and #2 cells to obtain whole cell extracts and performed the WB without removing the stacking gel before the transfer of the proteins into the membrane. Since DNA cannot migrate into the gel, if TDP1 H493R mutant protein was covalently linked to the DNA we should have detected TDP1 on the upper part of the immunoblot (Figure 24B). However, we did not detect any signal from the upper part of the gel. Thus, the protein reduction in WB of cell extracts is likely related to a real decrease in the cellular amount of H493R TDP1 protein. Thus, I propose that the mutant TDP1 may trigger a feedback mechanism that regulates the expression of its own gene. As

DNA damage levels and genome instability are increased in the presence of the SCAN1 TDP1, the reduction of SCAN1 TDP1 protein may be a consequence of the DNA damage response.

However, other mechanisms are possible. As we know that the lack of phosphorylation by PARP interferes with TDP1 stabilization (Das *et al.*, 2014), the single point mutation A1478G of the SCAN1 gene might also influence the stability of the protein. Many diseases result from single point mutations that influence various aspects of mRNA metabolism, including processing, mRNA splicing, translational control, stability, and export (Mendell and Dietz, 2001; Nicholson *et al.*, 2010). In addition, the mutation could destabilize the protein by reducing the free energy gap between folded and unfolded states (Quan, Lv and Zhang, 2016). Interestingly, a prediction model of the folding of WT and R119G mutant proteins suggests that the H493R mutation could significantly reduce protein stability (Li *et al.*, 2017). Similarly, R119G mutation on Pyrroline-5-carboxylate reductase (P5CR1) reduces its enzymatic activity and is reported to be linked to Autosomal recessive cutis laxa (ARCL). Therefore, the effects of the SCAN1 mutation on TDP1 stability needs to be determined in further studies.

4.2. Why do TDP1 depletion and TDP1 SCAN1 mutation (H493R) prevent the repair of CPT-induced DSBs in G1 phase of U2OS cells?

As already described in the literature (Barthelmes *et al.*, 2004; El-Khamisy *et al.*, 2005; Interthal *et al.*, 2005; Miao *et al.*, 2006; Caldecott, 2008; Hirano *et al.*, 2007; Katyal *et al.*, 2007; Das *et al.*, 2009; El-Khamisy, 2011), we confirmed that TDP1 H493R and TDP1 KO cells were sensitive to CPT. The treatment of TDP1 H493R and TDP1 KO cells with CPT to induce the trapping of TOP1ccs suggested that accumulation of DSBs could be related to the defective removal of TOP1ccs (Figure 33). Moreover, analysis of DSB repair following CPT treatment revealed that both TDP1 H493R and TDP1 KO cells were defective in the repair of DSBs in G1 but not in S phases (Figure 34). This could be related to a different nature/location of the DSBs and/or from the different repair mechanisms operative in proliferating vs. non-replicating cells.

In proliferating cells, the main cytotoxic mechanism of CPT is the production of replication-coupled DSBs (RC-DSBs) (Pommier, 2006), which can arise by two mechanisms. The first is the “replication run-off” (see Introduction section 1.1.3, Figure 5A) (Strumberg *et al.*, 2000) that drives to the formation of 5’ phosphorylated blunt-ended DSBs as result of the extension of the leading strand up to the TOP1cc at the template 5’end (Pouliot *et al.*, 1999; Interthal,

Pouliot and Champoux, 2001; Debethune, 2002). In this case, the TDP1 pathway needs to repair the TOP1ccs at the 3' end of the break site. The second alternative mechanism involves the 3′-flap endonuclease, as MUS81–EME1 which does not directly remove the TOP1-cc but rather cleaves stalled forks to convert collapsed replication forks into DSB, which are further repaired predominantly by HR (Regairaz *et al.*, 2011). To take place, HR needs sister chromatids, the reason why this pathway is restricted to S and G2 phase, and resection of DNA ends (Wyman and Kanaar, 2006; Hartlerode and Scully, 2010; Grabarz *et al.*, 2012).

In non-replicating cells, DSBs can arise from the collision of TOP1ccs with the elongation RNA Pol II complex (Wu and Liu, 1997; Pommier, 2006). The collisions lead to the production of TOP1-linked SSBs (Hsiang *et al.*, 1985; Ashour, Atteya and El-Khamisy, 2015) and transcription-coupled DSBs (TC-DSBs) (Sordet *et al.*, 2009, 2010; Cristini *et al.*, 2016, 2019). It is more likely that transcription-mediated CPT-induced lesions are repaired by SSB repair (SSBR) involving TDP1. CPT-sensitive SCAN1 lymphoblastoid cells (Plo *et al.*, 2003; El-Khamisy *et al.*, 2005; Interthal *et al.*, 2005; Interthal, Chen and Champoux, 2005; Zhou *et al.*, 2005; Miao *et al.*, 2006) and post-mitotic mouse TDP1-KO neurons cells (Katyal *et al.*, 2007; El-Khamisy *et al.*, 2009) show a marked defect in the repair of CPT-induced SSB.

Non-dividing cells, have reduced DSBs repair capability compared to dividing cells, as they can rely only on NHEJ. There is evidence that TDP1 is an accessory component of the NHEJ (Bahmed, Nitiss and Nitiss, 2010; Heo *et al.*, 2015), and it is required for efficient NHEJ. According to the literature, CPT-induced DSBs in TDP1-KO MEFs are due to reduced efficiency of DSB repair (Das *et al.*, 2009). In addition, HEK293 cells deficient in TDP1 showed an increase in insertions at I-SceI-induced DSB repair joints (J. Li et al., 2017), and TDP1 depletion did not lead to a DSB rejoining defect but caused DSB mis-joining partially via NHEJ (Kawale *et al.*, 2018). In addition, recently, it has been suggested that TOP1-associated DSBs are joined via NHEJ, which results in the deletion of the intervening sequence (Cho and Jinks-Robertson, 2019). Thus, the observed increase of DSB levels and reduced DSB repair in TDP1-deficient and H493R-TDP1 cells is likely due to a less efficient DSB repair mechanism in G1 phase, but not in S phase as TDP1 is likely involved in the specific DSB repair mechanism operative in non-proliferating cells.

4.3. How a difference in DSB repair may give rise to the SCAN1 phenotype?

Our results suggest that DSBs would accumulate specifically in TDP1-deficient cells that do not undergo replication, due to a defective repair of the lesion, and that H493R TDP1 causes a phenotype that is even more pronounced than TDP1 KO cells (Figure 34B). This could be explained by the fact that the mutant H493R TDP1 enzyme forms a persistent H493R TDP1-DNA adduct with a half-life of ~13min (Interthal *et al.*, 2005; Hirano *et al.*, 2007), which can be excised only by a wild-type TDP1 enzyme. In fact, TDP1 protein is the only known enzyme that can excise the 3′-phosphoamide adducts formed by H493R-TDP1 (Interthal, Chen and Champoux, 2005). This could also explain why heterozygous A1748G mutation does not rise to SCAN1 disease (Takashima *et al.*, 2002). Interestingly, TDP2 can promote the repair of TOP1-mediated DNA damage in the absence of TDP1 (Zeng *et al.*, 2012; Pommier *et al.*, 2014), therefore explaining the less severe cellular phenotype of TDP1 deficiency.

4.4. TDP1 deficiency induces modulation of R-loop as the CPT.

Our data (Figure 30) show that TDP1 deficiency induces a drop of DRIP signal at the β-ACTIN and γ-ACTIN TSS and an increase in the body of PTB and GEMIN 7 genes. Similar modulation at these genes is also observed in WI-38hTERT quiescent cells treated with CPT (Cristini et al., 2019). It is well known that TDP1 not only excises trapped TOP1ccs but also processes other 3'-end-blocking lesions, including 3'-phosphoglycolates that derive from oxidative DNA damage (Inamdar *et al.*, 2002; Interthal, Chen and Champoux, 2005; Zhou *et al.*, 2005). Moreover, under physiological condition, TOP1cc can be selectively trapped under a broad range of DNA lesions (see Introduction section 1.1.2, Table 1), including oxidative DNA lesions (8-oxoguanine, 8-oxoadenosine, and 5-hydroxycytosine) (Pourquier *et al.*, 1999; Lesher *et al.*, 2002). Thus, all these could elucidate the possibility that in TDP1 deficiency condition, cells accumulate TOP1cc, as CPT-induced TOP1cc, and this can favor alteration of the cellular balance of R-loops levels (Bendixen *et al.*, 1990; Wu and Liu, 1997; Sordet *et al.*, 2008; Tresini *et al.*, 2015; Cristini *et al.*, 2019). However, it is important to assess whether the redistribution of R-loops by TDP1 depletion is S-phase independent. To do so we planned to perform DRIP using the WI38 quiescent cells siRNA for TDP1, following the protocol used to investigate the R-loops in quiescent WI38 cells siRNA for SETX (Cristini et al., 2019), and in sorted living G1 U2OS TDP1 KO and H493R cells. In addition, we programmed to culture the WI-38 TDP1

KO cells in lower oxygen conditions, to reduce the oxidative stress due to the rates of ROS production, and analyze if they survive in quiescence conditions.

4.5. Perspective

My studies show that TDP1 depletion and TDP1 SCAN1 mutation leads to the accumulation of DNA damage in non-proliferating cells, thus leading to genome instability and cell death, in particular in conditions of increased levels of TOP1ccs (such as in CPT-treated cells) (summarized on Table 4). Other molecular aspects of the mechanisms need however to be determined to fully understand the molecular etiology of SCAN1 syndrome. I can foresee several experimental lines: a) in order to investigate the nature of the reduction of H493R protein, we planned to compare expression levels and turnover of the WT and mutant TDP1 genes in our cell models along with kinetic analyzes of the stability of WT and mutant enzymes; b) to investigate if the R-loops redistribution is strictly restricted to non-replicating cells, we plan to perform DRIP using the WI38 quiescent cells siRNA for TDP1 and in sorted living G1 U2OS TDP1 KO and H493R cells; c) to address how TDP1 deficiency and H493R TDP1 mutation can modulate R-loops, we plan to perform a genome-wide mapping of R loops by DRIP-seq approaches along with similar genome-wide mapping of γH2AX/p53BP1 in TDP1 KO and H493R TDP1 cells; d) to investigate if there are any links between the R-loops and DSB formation in our models, we planned to perform microscopy analysis to measure γH2AX/p53BP1 nuclear foci under R-loops modulating conditions (RNase H overexpression or siRNA of SETX); e) to assess if there are any differences in term of TOP1cc trapping we plan to investigate in WT, TDP1 KO and H493R TDP1 cells the number of TOP1cc by immunoblotting immunocomplex of enzyme (ICE) bioassay (isolation of DNA–protein complexes by CsCl gradient centrifugation and immunoblotting with Top1 monoclonal antibody) and the alkaline elution to detect the DNA part of TOP1cc; f) the role of transcription and TOP1 in the accumulation of DSB will be assessed by inhibiting RNA pol II transcription with specific chemical inhibitors and by silencing TOP1. Moreover, there is evidence that aphidicolin has a minimal effect in protecting TDP1-deficient cells from hypersensitivity to CPT compared to normal cells. And, the defective repair of TOP1cc in SCAN1 is independent of aphidicolin (Miao et al., 2006). It would be interesting then to use aphidicolin to study the replication-independent effects of the TDP1 deficiency in our models. The study will determine the main aspects of the mechanism leading to DSB accumulation, genome instability, and cell death in post-mitotic human cells caused by the impairment of TDP1 functions.

		Untreated conditions	CPT treated conditions
TDP1 $^{-/-}$ & H493R TDP1 mutant	U2OS	↑ DSBs outside of the S phase	↑ DSBs primarly in G1 phase ⊘ CPT-induced DSBs' repair in G1 phase
TDP1 deficiency	WI-38	↑ DSBs outside of the S phase and the % of cells with micronuclei Modulate the levels of R-loops at some loci	↑ DSBs primarly in G1 and S phase
	HCT116	↑ DSBs outside of the S phase and the % of cells with micronuclei	

Table 3. Schematics summarizing of the main observations. In the first column is indicated the TDP1 deficiencies; in the second one the model cell line; in the third and fourth columns the results observed. The upward arrow indicates the increase of DSBs, and the avoid signal that TDP1 deficiency prevent the repair of the CPT induced-DSBs.

5. METHODS AND MATERIALS

5.1. Cell culture

Primary human lung embryonic WI38 fibroblasts immortalized with hTERT were obtained from Carl Mann (CEA, Gif-sur-Yvette, France) (Jeanblanc *et al.*, 2012). Cells were cultured at 37° C with 5% CO2 in modified Eagle's medium (MEM) supplemented with 10% (v/v) fetal bovine serum, 1 mM sodium pyruvate, 2 mM glutamine and 0.1 mM non-essential amino acids. Human osteosarcoma U2OS cells were cultured at 37°C with 5% CO2 in DMEM High Glucose (SIGMA) supplemented with 10% (v/v) fetal bovine serum (PAN Biotech).

5.2. Generation of stable cell lines

5.2.1. Generation of U2OS TDP1 KO cells

TDP1 knockout in U2OS cells was generated by CRISPr genome editing strategy, targeting the exon 5 of *TDP1* (Al Abo *et al.*, 2017).

eSpCas9 (1.1)_No_FLAG_ATP1A1_G3_Dual_sgRNA plasmid (Addgene plasmid # 86613 ; http://n2t.net/addgene:86613 ; RRID:Addgene_86613), with the cloned-in target site sequence (GTTTAACTACTGCTTTGACG) were co-transfected in presence of the ATP1A1_plasmid_donor_RD (Addgene plasmid # 86551 ; http://n2t.net/addgene:86551 ; RRID:Addgene_86551) and the TDP1_stop_codon single-stranded oligodeoxyribonucleotides

(ssODNs) donor. Both plasmids were a gift of Yannick Doyon (Agudelo *et al.*, 2017). The ssODNs, were designed using A plasmid Editor (ApE) by M. Wayne Davis, UCSC Genome Browser by UC Santa Cruz and the ExPASy translate by Swiss Institute of Bioinformatic (SIB) scientific databases and software tools. Then, they were synthesized by Eurongentec (HPLC-RP purified and reconstituted in water RNase-Free water, Dharmacon, GE Healthcare at the concentration of 100uM). The correct cloning of the gRNA inside the vector was confirmed by DNA sequencing using G3seqFW primer: GGG AAA CGC CTG GTA TCT TT. The U2OS cell line, seeded in 24-well dishes, all 60-70% confluent, were transfected with jetPEI (Polyplus) according to the manufacturer's instructions.
72 hours post initial transfection, cells were cultured for a week under selection 0,7 uM/ml of Ouabain (SIGMA-Aldrich). Established clones from single-cell selection were subsequently screened and validated at a protein and genomic level.

5.2.2. **Generation of U2OS cell line harboring the H493R mutation**

To generate the U2OS cell line harboring the H493R mutation was targeted the exon14 of *TDP1*, using CRISPR/Cas9 strategy. The gRNA (TATGTGGCATGGCATTGCTG) was design using the MIT CRISPR and Crispor design algorithm.

Two different strategies were used in parallel:

1. cells were transfected, using the jetPEI (Polyplus) according to the manufacturer's instructions, with the eSpCas9 (1.1)_No_FLAG_ATP1A1_G3_Dual_sgRNA plasmid (Addgene plasmid # 86613 ; http://n2t.net/addgene:86613 ; RRID:Addgene_86613), with the cloned-in target site sequence (TATGTGGCATGGCATTGCTG) in presence of the ATP1A1_plasmid_donor_RD (Addgene plasmid # 86551 ; http://n2t.net/addgene:86551 ; RRID:Addgene_86551) and the TDP1_H493R_mut single-stranded oligodeoxyribonucleotides (ssODNs) donor;
2. cells were electroporated with the Amaxa Nucleofector system (Lonza) to let the codelivery of Ultramer single-stranded oligodeoxynucleotides (ssODN), ATP1A1_RD and TDP1_H493R_mut, and the CRISPR/Cas9 ribonucleoprotein (RNP) complexes (IDT Alt-R S.p. Cas9 Nuclease V3 (500 µg) and Alt-R CRISPR-Cas9 sgRNA 10 nmol (ATP1A1) and/or Alt-R CRISPR-Cas9 sgRNA, 2 nmol (TDP1). The cells were

cultured, 4hrs before and until 24hrs after electroporation, in medium supplemented by Nu7441 (Tocris) at the final concentration of 2uM, to inhibit the NHEJ.

72 hours post initial transfection/electroporation, cells were kept under positive selection by supplementing culture medium with the Ouabain (SIGMA-Aldrich) (0,7 uM) towards which parental cells all died and stable clones show resistance. Established clones from single-cell selection were subsequently screened at a protein and genomic level. Moreover, we took advantage of the new site CAT/CGT generated by the A1478G mutation insertion, which is recognized by the BSA A1 restriction enzyme (NEB), to analyzes if the single point mutation that leads to the SCAN1 disease was present in homozygosis. The reaction was performed following the protocol instructions.

5.2.3. **Cloning gRNA into the vector backbone**

The following protocol was used to clone the gRNA into the eSpCas9 (1.1) No_FLAG_ATP1A1_G3_Dual_sgRNA plasmid:

1. **Anneal each pair of the gRNA oligos**:

Oligo 1: 5'-CACCGNNNNNNNNNNNNNNNNNNN-3'
Oligo2: 3' CNNNNNNNNNNNNNNNNNNNCAAA-5'

- 1µL Oligo 1 (100µM)
- 1µL Oligo 2 (100µM)
- 1µL 10X T4 Ligation Buffer NEB with ATP (NEB, M0202S)
- 6.5µL H2O
- 0.5µL T4 PNK NEB
- Qsp 10µL total
- Incubate 37°C 30 min
- Incubate at 95°C 5min and then leave on the bench at RT to cool down for 1hr
- Dilute the annealed oligo 1:250 (250-fold)

2. **Set up digestion ligation reaction:**

- 1µL backbone vector (100ng/µL)
- 2µL Oligos phosphorylated annealed diluted (1/250)

- 2µL Tango Buffer 10X (ThermoFisher Scientific, BY5)
- 1µL DTT (0.1M)
- 1µL ATP (10mM)
- 1µL BbSI (NEB #R3539)
- 0.5µL T7 ligase (NEB, M0318S)
- H2O to a final volume of 20µL
- Incubate the ligation in a thermocycler
- 37°C 5min
- 23°C 5 min: cycle the previous two steps for 6 cycles (total running time 1h)
- Hold until proceeding 4°C

3. **Treatment of exonuclease:**
    - 11µL ligation from step 2
    - 1.5µL NEB 4 buffer 10X
    - 1.5µL ATP (10mM)
    - 1µL exo V (RecBCD, NEB M0345S)
    - Incubate the reaction at 37°C for 30min
    - Inactivate reaction at 70°C for 30min
4. **Transformation with 1-2µL of the final product into the competent cell**
5. **Pick colony and sequence verify**

5.2.4. Generation of WI-38hTERT TDP1 KO cells

The WI-38hTERT TDP1 KO was generated targeting the exon 14 (TATGTGGCATGGCATTGCTG) by lentiCas9-Blast and LentiGuide-Puro plasmid (Addgene plasmid #52963; http://n2t.net/addgene:52963; RRID: Addgene_52963), the gift of Feng Zhang (Sanjana, Shalem and Zhang, 2014). Cloning and transduction were carried out following the protocol instructions. The correct insertion of the target guide sequence was then checked by sequencing (3′ sequencing primer hGata4-rev: 5'-ATTGTGGATGAATACTGCC-3').

Established clones from single-cell selection were subsequently screened at a protein and genomic level.

5.3. siRNA Transfection

Cells were transfected with siRNA duplexes using Dharmafect 4 transfection reagent (Dharmacon) for 24 h before inducing quiescence for 72 h. siRNAs used is a pool of 4 siRNAs from Dharmacon (TDP1: M-016112-01).

5.4. Cell Extracts and Immunoblotting

Whole cell lysate was prepared by lysing the cells in a buffer containing 1% SDS (SIGMA), 10mM Tris pH 7,4 (SIGMA) in presence of protease, and phosphatase inhibitors (Halt™ Protease and Phosphatase Inhibitor Cocktail (100X), Thermo Scientific™).

Antibodies used for western blot and dilutions used are: TDP1 (abcam ab224822) 1:1000, beta-actin (Millipore Cat# MAB1501; RRID: AB2223041) 1:50000.

5.5. Genomic DNA extraction and CRISPr/Cas9 editing detection

Genomic DNA was extracted from cells using the GenElute Mammalian Genomic DNA miniprep kit (SIGMA).

The targeted regions were amplified using the following primers:

- Primer-TDP1-exon 5-Fw: TAATAACACCCTGCGGACAAG;
- Primer-TDP1-exon 5-Rv: CCTGAAAGAGCACAGAGGAAA;
- Primer-TDP1-exon 14-Fw: CCAAATGTCAATTTGCATAACC;
- Primer-TDP1-exon 14-Rv: GGTTGCAGTGCTTCTTGTGA.

Phusion High-Fidelity DNA Polymerase (M0530) by Thermo Scientific™ was used following the protocol instructions.

The CRISPR editing was investigated by chromatogram analysis (ApE software) and TIDE analysis.

5.6. Immunofluorescence Microscopy

Immunofluorescence microscopy was performed as described in (Cristini et al., 2019). Cells were seeded in 96-well plates (CellCarrier; PerkinElmer). After treatment with CPT (SIGMA-Aldrich) and 30 min EdU (10 mM) incorporation, to discriminate G1 phase cells from S phase

cells, cells were washed for 5 min with PBS and fixed with 3.7% paraformaldehyde for 15 min. After two washes with PBS, cells were permeabilized with 0.5% Triton X-100 for 15 min and washed twice with PBS. Cells were incubated with 8% bovine serum albumin (BSA) in PBS for 1 h before incubation with a mouse anti-γH2AX antibody (05-636; Millipore) and/or a rabbit anti-p53BP1 antibody (#2675; Cell Signaling Technology) diluted at 1/500 in 1% BSA in PBS for 2 h. Cells were washed three times with PBS and incubated with the appropriate secondary antibody coupled to Alexa Fluor 488, 594, or 647 (Thermo Scientific™) diluted at 1/500 in 1% BSA in PBS for 1 h. After three washes with PBS, nuclei were stained with 1 mg/ml Hoechst 33342 for 15 min, washed twice with PBS, and stored at 4°C until analysis. The incorporated EdU into DNA was detected using the Click-iT EdU Alexa Fluor 647 Imaging Kit (ThermoFisher Scientific) according to the manufacturer's protocol. After the Click-iT reaction, cells were processed for immunolabelling as described above starting at the 8% BSA step. 96-well plates were scanned with a 20X objective using an Operetta High-Content Imaging System or an Operetta CLS High-Content Imaging System (PerkinElmer) with Harmony software (version 4.1 or 4.8). After data acquisition, subsequent analyzes were performed with Columbus software (version 2.5.0 or 2.8.2). γH2AX and p53BP1 foci were detected with the ''C'' method. The number of foci was automatically counted with Columbus software that depends on the ratio between the signal intensity of foci and background noise. For a graphical representation of foci distribution, we used box-and-whisker plots with GraphPad Prism 6 software with the following settings: boxes: 25-75 percentile range; whiskers: 10-90 percentile range; horizontal bars: median number of foci; ''+'': mean number of foci.

5.7. RNA/DNA Immunoprecipitation (DRIP)

RNA/DNA hybrid co-IP was performed as described (Cristini *et al.*, 2018) using the S9.6 antibody (Boguslawski *et al.*, 1986). Briefly, non-crosslinked nuclei were incubated in RSB buffer (10 mM Tris-HCl pH 7.5, 200 mM NaCl, 2.5 mM MgCl2) with 0.2% sodium deoxycholate, 0.1% SDS, 0.05% sodium lauroyl sarcosinate and 0.5% Triton X-100, and sonicated for 10 min (Diagenode Bioruptor). Samples were then diluted 4 times in RSB with 0.5% Triton X-100 (RSB + T) before IP with the S9.6 antibody, bound to protein A dynabeads

(Invitrogen), and preblocked with 0.5% BSA/PBS for 2 h. RNA/DNA hybrids were prepared as described in (Phillips *et al.*, 2013; Cristini *et al.*, 2018).

Primer used for the qPCR:

Gene region	**Primer Fw**	**Primer Rw**
bAct		
TSS*	CGGGGTCTTTGTCTGAGC	CAGTTAGCGCCCAAAGGAC
GB**	GGAGCTGTCACATCCAGGGTC	TGCTGATCCACATCTGCTGG
Upstream	ACCCAGCACCCCCTAATACC	AGCCGGACATGCTTCCAGAG
gAct		
TSS	CCGCAGTGCAGACTTCCGAG	CGGGCGCGTCTGTAACACGG
GB	GTGACACAGCATCACTAAGG	ACAGCACCGTGTTGGCGT
PTB		
GB	GCC GTT GGT ACA AAG GTA GG	GCC CCT TAG GAA TGG AAA AG
Gemin 7		
GB	TCTTCTTCCACCTGGACCAC	GGGACAGAGAGAGTGCCTTG

*Transcription starting site (TSS); **Gene body.

5.8. Comet Assays

Neutral comet assays were performed as described in(Cristini *et al.*, 2019), according to the manufacturer's instructions (Trevigen), except that electrophoresis was performed at 4° C. Slides were scanned by using an AxioObserver Z1 fluorescence microscope (ZEISS) with the objective EC Plan-Neofluar 10X / 0.3 Ph1. Comet tail moments were measured with ImageJ software (version 1.51n) with the plugin OpenComet (http://www.cometbio.org).

Acknowledgements

This work was supported by a PhD fellowship under the French-Italian University VINCI Program.

I thank Dr. Laurent Corcos, and Dr. Nicoletta Landsberger and Dr. Philippe Pourquier for accepting to evaluate my thesis work and for the advices to improve my manuscript. Thank you to Dr. Annalisa Pession and Prof. Gilles Favre for agreeing to be the examiner of my defense. I thank also Prof. Gilles Favre for having welcomed me in your team, for the stimulating, positive and constructive environment in which I worked. A big thanks to Olivier for having trusted me and for having involved me to actively participate in the discussions from the beginning. Thank you especially for the constant scientific discussion we had during my thesis, it really brought me a lot and it was extremely formative. Thank you for directing me to Dr. Yvan Canitrot to discuss about the CRISPr, it was determinant for the success of my project. I would like to thank you also for the passion, patience, calm, and rationality with which you face things, it has been also a source of enrichment for me in facing the daily challenges that research proposes. I have loved the way you mentored me. Thank you to Prof. Capranico, for being the first to believe in me, choosing me as your doctoral candidate for this project. For making me passionate about DNA damage and R-loops and to have been present, despite the physical distance, during my time in Toulouse. Moreover, thanks for directing me to Dr. Ciarrocchi's laboratory. It was a good choice for my personal growth and for the project. Thanks also to all the members of Professor Gille Favre's group, in particular Nesibe, Mathéa, Agnese, Stephanie, Chaty, Betty, Delphine, Seb, Aurélien, Isabelle, Sarah, Sandra, Remie and Claire and Professor Capranico's group in particular Jessica, Stefano, Marco and le Giulie for their constructive discussions, comparisons and daily support. Thank you to Dr. Yvan Canitrot and to Jérémie, to Dr. Ciarrocchi and Dr. Sancisi for the discussion and the advice about the CRISPr, your time and availability.

Thank you to all the people, family and friends, that supported me during these years.

Bibliography

Al Abo, M. *et al.* (2017) 'TDP1 is critical for the repair of DNA breaks induced by sapacitabine, a nucleoside also targeting ATM- and BRCA-deficient tumors', *Molecular Cancer Therapeutics*. doi: 10.1158/1535-7163.MCT-17-0110.

Agudelo, D. *et al.* (2017) 'Marker-free coselection for CRISPR-driven genome editing in human cells', *Nature Methods*. doi: 10.1038/nmeth.4265.

Aguilera, A. and García-Muse, T. (2012) 'R Loops: From Transcription Byproducts to Threats to Genome Stability', *Molecular Cell*. doi: 10.1016/j.molcel.2012.04.009.

Aguilera, A. and Gómez-González, B. (2008) 'Genome instability: A mechanistic view of its causes and consequences', *Nature Reviews Genetics*. doi: 10.1038/nrg2268.

Aguilera, A. and Gómez-González, B. (2017) 'DNA-RNA hybrids: The risks of DNA breakage during transcription', *Nature Structural and Molecular Biology*. doi: 10.1038/nsmb.3395.

Akopiants, K. *et al.* (2014) 'Tracking the processing of damaged DNA double-strand break ends by ligation-mediated PCR: Increased persistence of 3'-phosphoglycolate termini in SCAN1 cells', *Nucleic Acids Research*. doi: 10.1093/nar/gkt1347.

Alagoz, M., Wells, O. S. and El-Khamisy, S. F. (2014) 'TDP1 deficiency sensitizes human cells to base damage via distinct topoisomerase I and PARP mechanisms with potential applications for cancer therapy', *Nucleic Acids Research*. doi: 10.1093/nar/gkt1260.

Amon, J. D. and Koshland, D. (2016) 'RNase H enables efficient repair of R-loop induced DNA damage', *eLife*. doi: 10.7554/eLife.20533.

Antony, S. *et al.* (2007) 'Novel high-throughput electrochemiluminescent assay for identification of human tyrosyl-DNA phosphodiesterase (Tdp1) inhibitors and characterization of furamidine (NSC 305831) as an inhibitor of Tdp1', *Nucleic Acids Research*. doi: 10.1093/nar/gkm463.

Ashour, M. E., Atteya, R. and El-Khamisy, S. F. (2015a) 'Topoisomerase-mediated chromosomal break repair: An emerging player in many games', *Nature Reviews Cancer*. doi: 10.1038/nrc3892.

Ashour, M. E., Atteya, R. and El-Khamisy, S. F. (2015b) 'Topoisomerase-mediated chromosomal break repair: An emerging player in many games', *Nature Reviews Cancer*, 15(3), pp. 137–151. doi: 10.1038/nrc3892.

Bahmed, K., Nitiss, K. C. and Nitiss, J. L. (2010) 'Yeast Tdp1 regulates the fidelity of nonhomologous end joining', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.0909917107.

Baker, S. D. *et al.* (1995) 'Cell cycle analysis of amount and distribution of nuclear DNA topoisomerase I as determined by fluorescence digital imaging microscopy', *Cytometry*. doi: 10.1002/cyto.990190208.

Baranello, L. *et al.* (2016) 'RNA Polymerase II Regulates Topoisomerase 1 Activity to Favor Efficient Transcription', *Cell*. doi: 10.1016/j.cell.2016.02.036.

Barnes, D. E. and Lindahl, T. (2004) 'Repair and Genetic Consequences of Endogenous DNA Base Damage in Mammalian Cells', *Annual Review of Genetics*. doi: 10.1146/annurev.genet.38.072902.092448.

Barthelmes, H. U. *et al.* (2004) 'TDP1 overexpression in human cells counteracts DNA damage mediated by topoisomerases I and II', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M405042200.

Becherel, O. J. *et al.* (2015) 'A new model to study neurodegeneration in ataxia oculomotor apraxia type 2', *Human Molecular Genetics*. doi: 10.1093/hmg/ddv296.

Been, M. D. and Champoux, J. J. (1984) 'Breakage of single-stranded DNA by eukaryotic type 1 topoisomerase occurs only at regions with the potential for base-pairing', *Journal of Molecular Biology*. doi: 10.1016/0022-2836(84)90025-1.

Beletskii, A. and Bhagwat, A. S. (1996) 'Transcription-induced mutations: Increase in C to T mutations in the nontranscribed strand during transcription in Escherichia coli', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.93.24.13919.

Bendixen, C. *et al.* (1990) 'Camptothecin-Stabilized Topoisomerase I-DNA Adducts Cause Premature Termination of Transcription', *Biochemistry*. doi: 10.1021/bi00475a028.

Bergerat, A. *et al.* (1997) 'An atypical topoisomerase II from archaea with implications for meiotic recombination', *Nature*. doi: 10.1038/386414a0.

Beucher, A. *et al.* (2009) 'ATM and Artemis promote homologous recombination of radiation-induced DNA double-strand breaks in G2', *EMBO Journal*. doi: 10.1038/emboj.2009.276.

Blount, B. C. *et al.* (1997) 'Medical Sciences Folate deficiency causes uracil misincorporation into human DNA and chromosome breakage: Implications for cancer and neuronal damage', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.94.7.3290.

Boguslawski, S. J. *et al.* (1986) 'Characterization of monoclonal antibody to DNA · RNA and its application to immunodetection of hybrids', *Journal of Immunological Methods*. doi: 10.1016/0022-1759(86)90040-2.

Bonner, William M *et al.* (2008) 'gammaH2AX and cancer', *Nat Rev Cancer*.

Bonner, William M. *et al.* (2008) 'γH2AX and cancer', *Nature Reviews Cancer*. doi: 10.1038/nrc2523.

De Bont, R. and van Larebeke, N. (2004) 'Endogenous DNA damage in humans: A review of quantitative data', *Mutagenesis*. doi: 10.1093/mutage/geh025.

Bowman, K. J. *et al.* (2001) 'Differential effects of the poly (ADP-ribose) polymerase (PARP) inhibitor NU1025 on topoisomerase I and II inhibitor cytotoxicity in L1210 cells in vitro', *British Journal of Cancer*. doi: 10.1054/bjoc.2000.1555.

Brettrager, E. J., Segura, I. A. and van Waardenburg, R. C. A. M. (2019) 'Tyrosyl-DNA phosphodiesterase I N-Terminal domain modifications and interactions regulate cellular function', *Genes*. doi: 10.3390/genes10110897.

Brinkman, E. K. *et al.* (2014) 'Easy quantitative assessment of genome editing by sequence trace decomposition', *Nucleic Acids Research*. doi: 10.1093/nar/gku936.

Burma, S., Chen, B. P. C. and Chen, D. J. (2006) 'Role of non-homologous end joining (NHEJ) in maintaining genomic integrity', *DNA Repair*. doi: 10.1016/j.dnarep.2006.05.026.

Caldecott, K. W. (2008) 'Single-strand break repair and genetic disease', *Nature Reviews Genetics*. doi: 10.1038/nrg2380.

Capranico, G. *et al.* (2007) 'The effects of camptothecin on RNA polymerase II transcription: Roles of DNA topoisomerase I', *Biochimie*. doi: 10.1016/j.biochi.2007.01.001.

Capranico, G., Marinello, J. and Chillemi, G. (2017) 'Type I DNA Topoisomerases', *Journal of Medicinal Chemistry*. doi: 10.1021/acs.jmedchem.6b00966.

Ceccaldi, R., Rondinelli, B. and D'Andrea, A. D. (2016) 'Repair Pathway Choices and Consequences at the Double-Strand Break', *Trends in Cell Biology*. doi: 10.1016/j.tcb.2015.07.009.

Cerritelli, S. M. and Crouch, R. J. (2009) 'Ribonuclease H: The enzymes in eukaryotes', *FEBS Journal*. doi: 10.1111/j.1742-4658.2009.06908.x.

Champoux, J. J. (2001) 'DNA Topoisomerases: Structure, Function, and Mechanism', *Annual Review of Biochemistry*. doi: 10.1146/annurev.biochem.70.1.369.

Chan, Y. A., Hieter, P. and Stirling, P. C. (2014) 'Mechanisms of genome instability induced by RNA-processing defects', *Trends in Genetics*. doi: 10.1016/j.tig.2014.03.005.

Chapman, J. R., Taylor, M. R. G. and Boulton, S. J. (2012) 'Playing the End Game: DNA Double-Strand Break Repair Pathway Choice', *Molecular Cell*. doi: 10.1016/j.molcel.2012.07.029.

Chaudhuri, J., Khuong, C. and Alt, F. W. (2004) 'Replication protein A interacts with AID to promote deamination of somatic hypermutation targets', *Nature*. doi: 10.1038/nature02821.

Chen, L. *et al.* (2017) 'R-ChIP Using Inactive RNase H Reveals Dynamic Coupling of R-loops with Transcriptional Pausing at Gene Promoters', *Molecular Cell*. doi: 10.1016/j.molcel.2017.10.008.

Cheng, T. J. *et al.* (2002) 'Kinetic studies of human tyrosyl-DNA phosphodiesterase, an enzyme in the topoisomerase I DNA repair pathway', *European Journal of Biochemistry*. doi: 10.1046/j.1432-1033.2002.03059.x.

Chiang, S. C. *et al.* (2017) 'Mitochondrial protein-linked DNA breaks perturb mitochondrial gene transcription and trigger free radical–induced DNA damage', *Science Advances*. doi: 10.1126/sciadv.1602506.

Chiang, S. C., Carroll, J. and El-Khamisy, S. F. (2010) 'TDP1 serine 81 promotes interaction with DNA ligase IIIα and facilitates cell survival following DNA damage', *Cell Cycle*. doi: 10.4161/cc.9.3.10598.

Chiruvella, K. K., Liang, Z. and Wilson, T. E. (2013) 'Repair of double-strand breaks by end joining', *Cold Spring Harbor Perspectives in Biology*. doi: 10.1101/cshperspect.a012757.

Cho, J. E. and Jinks-Robertson, S. (2019) 'Deletions associated with stabilization of the Top1 cleavage complex in yeast are products of the nonhomologous end-joining pathway', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.1914081116.

Cloutier, S. C. *et al.* (2016) 'Regulated Formation of lncRNA-DNA Hybrids Enables Faster Transcriptional Induction and Environmental Adaptation', *Molecular Cell*. doi: 10.1016/j.molcel.2015.12.024.

Cohen, S. *et al.* (2018) 'Senataxin resolves RNA:DNA hybrids forming at DNA double-strand breaks to prevent translocations', *Nature Communications*. doi: 10.1038/s41467-018-02894-w.

Coller, H. A., Sang, L. and Roberts, J. M. (2006) 'A new description of cellular quiescence',

PLoS Biology. doi: 10.1371/journal.pbio.0040083.

Costantino, L. and Koshland, D. (2015) 'The Yin and Yang of R-loop biology', *Current Opinion in Cell Biology*. doi: 10.1016/j.ceb.2015.04.008.

Cottarel, J. *et al.* (2013) 'A noncatalytic function of the ligation complex during nonhomologous end joining', *Journal of Cell Biology*. doi: 10.1083/jcb.201203128.

Cristini, A. *et al.* (2016) 'DNA-PK triggers histone ubiquitination and signaling in response to DNA double-strand breaks produced during the repair of transcription-blocking topoisomerase I lesions', *Nucleic Acids Research*. doi: 10.1093/nar/gkv1196.

Cristini, A. *et al.* (2018) 'RNA/DNA Hybrid Interactome Identifies DXH9 as a Molecular Player in Transcriptional Termination and R-Loop-Associated DNA Damage', *Cell Reports*. doi: 10.1016/j.celrep.2018.04.025.

Cristini, A. *et al.* (2019) 'Dual Processing of R-Loops and Topoisomerase I Induces Transcription-Dependent DNA Double-Strand Breaks', *Cell Reports*. doi: 10.1016/j.celrep.2019.08.041.

D'Alessandro, G. *et al.* (2018) 'BRCA2 controls DNA:RNA hybrid level at DSBs by mediating RNase H2 recruitment', *Nature Communications*. doi: 10.1038/s41467-018-07799-2.

Das, B. B. *et al.* (2009) 'Optimal function of the DNA repair enzyme TDP1 requires its phosphorylation by ATM and/or DNA-PK', *EMBO Journal*. doi: 10.1038/emboj.2009.302.

Das, B. B. *et al.* (2010) 'Role of tyrosyl-DNA phosphodiesterase (TDP1) in mitochondria.', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.1009814107.

Das, B. B. *et al.* (2014) 'PARP1-TDP1 coupling for the repair of topoisomerase I-induced DNA damage', *Nucleic Acids Research*. doi: 10.1093/nar/gku088.

Davies, D. R. *et al.* (2002a) 'Insights into substrate binding and catalytic mechanism of human tyrosyl-DNA phosphodiesterase (Tdp1) from vanadate and tungstate-inhibited structures', *Journal of Molecular Biology*. doi: 10.1016/S0022-2836(02)01154-3.

Davies, D. R. *et al.* (2002b) 'The crystal structure of human tyrosyl-DNA phosphodiesterase, Tdp1', *Structure*. doi: 10.1016/S0969-2126(02)00707-4.

Davies, D. R. *et al.* (2003) 'Crystal structure of a transition state mimic for Tdp1 assembled from vanadate, DNA, and a topoisomerase I-derived peptide', *Chemistry and Biology*. doi: 10.1016/S1074-5521(03)00021-8.

Debethune, L. (2002) 'Processing of nucleopeptides mimicking the topoisomerase I-DNA covalent complex by tyrosyl-DNA phosphodiesterase', *Nucleic Acids Research*. doi: 10.1093/nar/30.5.1198.

Deng, C. *et al.* (2005) 'Multiple endonucleases function to repair covalent topoisomerase I complexes in Saccharomyces cerevisiae', *Genetics*. doi: 10.1534/genetics.104.028795.

Desai, S. D. *et al.* (2003) 'Transcription-Dependent Degradation of Topoisomerase I-DNA Covalent Complexes', *Molecular and Cellular Biology*. doi: 10.1128/mcb.23.7.2341-2350.2003.

Deweese, J. E. and Osheroff, N. (2009) 'The DNA cleavage reaction of topoisomerase II: Wolf in sheep's clothing', *Nucleic Acids Research*. doi: 10.1093/nar/gkn937.

Dexheimer, T. S. *et al.* (2010) 'The DNA binding and 3′-end preferential activity of human

tyrosyl-DNA phosphodiesterase', *Nucleic Acids Research*. doi: 10.1093/nar/gkp1206.

Dimri, G. P., Hara, E. and Campisi, J. (1994) 'Regulation of two E2F-related genes in presenescent and senescent human fibroblasts', *Journal of Biological Chemistry*.

Douarre, C. *et al.* (2012) 'Mitochondrial topoisomerase I is critical for mitochondrial integrity and cellular energy metabolism', *PLoS ONE*. doi: 10.1371/journal.pone.0041094.

Drolet, M. *et al.* (2003) 'The problem of hypernegative supercoiling and R-loop formation in transcription', *Frontiers in Bioscience*. doi: 10.2741/970.

Drolet, M., Bi, X. and Liu, L. F. (1994) 'Hypernegative supercoiling of the DNA template during transcription elongation in vitro', *Journal of Biological Chemistry*.

El-Khamisy, S. F. *et al.* (2005) 'Defective DNA single-strand break repair in spinocerebellar ataxia with axonal neuropathy-1', *Nature*. doi: 10.1038/nature03314.

El-Khamisy, S. F. *et al.* (2009) 'Synergistic decrease of DNA single-strand break repair rates in mouse neural cells lacking both Tdp1 and aprataxin', *DNA Repair*. doi: 10.1016/j.dnarep.2009.02.002.

El-Khamisy, S. F. (2011) 'To live or to die: A matter of processing damaged DNA termini in neurons', *EMBO Molecular Medicine*. doi: 10.1002/emmm.201000114.

El-Khamisy, S. F., Hartsuiker, E. and Caldecott, K. W. (2007) 'TDP1 facilitates repair of ionizing radiation-induced DNA single-strand breaks', *DNA Repair*. doi: 10.1016/j.dnarep.2007.04.015.

Eng, W. K. *et al.* (1988) 'Evidence that DNA topoisomerase I is necessary for the cytotoxic effects of camptothecin', *Molecular Pharmacology*.

Fam, H. K., Chowdhury, M. K., *et al.* (2013) 'Expression profile and mitochondrial colocalization of Tdp1 in peripheral human tissues', *Journal of Molecular Histology*. doi: 10.1007/s10735-013-9496-5.

Fam, H. K., Walton, C., *et al.* (2013) 'TDP1 and PARP1 deficiency are cytotoxic to rhabdomyosarcoma cells', *Molecular Cancer Research*. doi: 10.1158/1541-7786.MCR-12-0575.

Fam, H. K. *et al.* (2018) 'Reactive oxygen species stress increases accumulation of tyrosyl-DNA phsosphodiesterase 1 within mitochondria', *Scientific Reports*. doi: 10.1038/s41598-018-22547-8.

Feig, D. I., Sowers, L. C. and Loeb, L. A. (1994) 'Reverse chemical mutagenesis: Identification of the mutagenic lesions resulting from reactive oxygen species-mediated damage to DNA', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.91.14.6609.

Furuta, T. *et al.* (2003) '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', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M300198200.

Gan, W. *et al.* (2011) 'R-loop-mediated genomic instability is caused by impairment of replication fork progression', *Genes and Development*. doi: 10.1101/gad.17010011.

Gao, R. *et al.* (2014) 'Epigenetic and genetic inactivation of tyrosyl-DNA-phosphodiesterase 1 (TDP1) in human lung cancer cells from the NCI-60 panel', *DNA Repair*. doi:

10.1016/j.dnarep.2013.09.001.

García-Muse, T. and Aguilera, A. (2019) 'R Loops: From Physiological to Pathological Roles', *Cell*. doi: 10.1016/j.cell.2019.08.055.

García-Pichardo, D. *et al.* (2017) 'Histone Mutants Separate R Loop Formation from Genome Instability Induction', *Molecular Cell*. doi: 10.1016/j.molcel.2017.05.014.

Van Gent, D. C., Hoeijmakers, J. H. J. and Kanaar, R. (2001) 'Chromosomal stability and the DNA double-stranded break connection', *Nature Reviews Genetics*. doi: 10.1038/35056049.

Ghosh, A. *et al.* (2019) 'SCAN1-TDP1 trapping on mitochondrial DNA promotes mitochondrial dysfunction and mitophagy', *Science Advances*. doi: 10.1126/sciadv.aax9778.

Ginno, P. A. *et al.* (2012) 'R-Loop Formation Is a Distinctive Characteristic of Unmethylated Human CpG Island Promoters', *Molecular Cell*. doi: 10.1016/j.molcel.2012.01.017.

Ginno, P. A. *et al.* (2013) 'GC skew at the 59 and 39 ends of human genes links R-loop formation to epigenetic regulation and transcription termination', *Genome Research*. doi: 10.1101/gr.158436.113.

Gorgoulis, V. G. *et al.* (2005) 'Activation of the DNA damage checkpoint and genomic instability in human precancerous lesions', *Nature*. doi: 10.1038/nature03485.

Goto, T. and Wang, J. C. (1985) 'Cloning of yeast TOP1, the gene encoding DNA topoisomerase I, and construction of mutants defective in both DNA topoisomerase I and DNA topoisomerase II', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.82.21.7178.

Gottlin, E. B. *et al.* (1998) 'Catalytic mechanism of the phospholipase D superfamily proceeds via a covalent phosphohistidine intermediate', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.95.16.9202.

Gough, G. W., Sullivan, K. M. and Lilley, D. M. (1986) 'The structure of cruciforms in supercoiled DNA: probing the single-stranded character of nucleotide bases with bisulphite.', *The EMBO Journal*. doi: 10.1002/j.1460-2075.1986.tb04195.x.

Grabarz, A. *et al.* (2012) 'Initiation of DNA double strand break repair: signaling and single-stranded resection dictate the choice between homologous recombination, non-homologous end-joining and alternative end-joining.', *American journal of cancer research*.

Graf, M. *et al.* (2017) 'Telomere Length Determines TERRA and R-Loop Regulation through the Cell Cycle', *Cell*. doi: 10.1016/j.cell.2017.06.006.

Groh, M. *et al.* (2014) 'R-loops Associated with Triplet Repeat Expansions Promote Gene Silencing in Friedreich Ataxia and Fragile X Syndrome', *PLoS Genetics*. doi: 10.1371/journal.pgen.1004318.

Groh, M. and Gromak, N. (2014) 'Out of Balance: R-loops in Human Disease', *PLoS Genetics*. doi: 10.1371/journal.pgen.1004630.

Grunseich, C. *et al.* (2018) 'Senataxin Mutation Reveals How R-Loops Promote Transcription by Blocking DNA Methylation at Gene Promoters', *Molecular Cell*. doi: 10.1016/j.molcel.2017.12.030.

El Hage, A. *et al.* (2010) 'Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis', *Genes and Development*. doi: 10.1101/gad.573310.

El Hage, A. *et al.* (2014) 'Genome-Wide Distribution of RNA-DNA Hybrids Identifies RNase

H Targets in tRNA Genes, Retrotransposons and Mitochondria', *PLoS Genetics*. doi: 10.1371/journal.pgen.1004716.

Hamperl, S. *et al.* (2017) 'Transcription-Replication Conflict Orientation Modulates R-Loop Levels and Activates Distinct DNA Damage Responses', *Cell*. doi: 10.1016/j.cell.2017.07.043.

Hamperl, S. and Cimprich, K. A. (2014) 'The contribution of co-transcriptional RNA: DNA hybrid structures to DNA damage and genome instability', *DNA Repair*. doi: 10.1016/j.dnarep.2014.03.023.

Hartlerode, A. J. and Scully, R. (2010) 'Mechanisms of double-strand break repair in somatic mammalian cells (Biochemical Journal (2009) 423, (157-168))', *Biochemical Journal*. doi: 10.1042/BJ4260389.

Ben Hassine, S. and Arcangioli, B. (2009) 'Tdp1 protects against oxidative DNA damage in non-dividing fission yeast', *EMBO Journal*. doi: 10.1038/emboj.2009.9.

Hawkins, A. J. *et al.* (2009) 'In vitro complementation of Tdp1 deficiency indicates a stabilized enzyme-DNA adduct from tyrosyl but not glycolate lesions as a consequence of the SCAN1 mutation', *DNA Repair*. doi: 10.1016/j.dnarep.2008.12.012.

He, X. *et al.* (2007) 'Mutation of a Conserved Active Site Residue Converts Tyrosyl-DNA Phosphodiesterase I into a DNA Topoisomerase I-dependent Poison', *Journal of Molecular Biology*. doi: 10.1016/j.jmb.2007.07.055.

Hegazy, Y. A., Fernando, C. M. and Tran, E. J. (2020) 'The balancing act of R-loop biology: The good, the bad, and the ugly', *Journal of Biological Chemistry*. doi: 10.1074/jbc.REV119.011353.

Henner, W. D., Grunberg, S. M. and Haseltine, W. A. (1983) 'Enzyme action at 3' termini of ionizing radiation-induced DNA strand breaks', *Journal of Biological Chemistry*.

Heo, J. *et al.* (2015) 'TDP1 promotes assembly of non-homologous end joining protein complexes on DNA', *DNA Repair*. doi: 10.1016/j.dnarep.2015.03.003.

Hirano, R. *et al.* (2007) 'Spinocerebellar ataxia with axonal neuropathy: Consequence of a Tdp1 recessive neomorphic mutation?', *EMBO Journal*. doi: 10.1038/sj.emboj.7601885.

Hodroj, D. *et al.* (2017) ' An ATR -dependent function for the Ddx19 RNA helicase in nuclear R-loop metabolism ', *The EMBO Journal*. doi: 10.15252/embj.201695131.

Holm, C. *et al.* (1989) 'Differential Requirement of DNA Replication for the Cytotoxicity of DNA Topoisomerase I and II Inhibitors in Chinese Hamster DC3F Cells', *Cancer Research*.

Hsiang, Y. H. *et al.* (1985) 'Camptothecin induces protein-linked DNA breaks via mammalian DNA topoisomerase I', *Journal of Biological Chemistry*.

Hsiang, Y. H., Lihou, M. G. and Liu, L. F. (1989) 'Arrest of Replication Forks by Drug-stabilized Topoisomerase I-DNA Cleavable Complexes as a Mechanism of Cell Killing by Camptothecin', *Cancer Research*.

Huang, S. Y. N., Ghosh, S. and Pommier, Y. (2015) 'Topoisomerase I alone is sufficient to produce short DNA deletions and can also reverse nicks at ribonucleotide sites', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M115.653345.

Huang, S. Y. N. and Pommier, Y. (2019) 'Mammalian tyrosyl-dna phosphodiesterases in the context of mitochondrial dna repair', *International Journal of Molecular Sciences*. doi: 10.3390/ijms20123015.

Huang, T. H. *et al.* (2010) 'Cellular processing determinants for the activation of damage signals in response to topoisomerase I-linked DNA breakage', *Cell Research*. doi: 10.1038/cr.2010.95.

Hudson, J. J. R. *et al.* (2012) 'SUMO modification of the neuroprotective protein TDP1 facilitates chromosomal single-strand break repair', *Nature Communications*. doi: 10.1038/ncomms1739.

Huertas, P. and Aguilera, A. (2003) 'Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination', *Molecular Cell*. doi: 10.1016/j.molcel.2003.08.010.

Inamdar, K. V. *et al.* (2002) 'Conversion of phosphoglycolate to phosphate termini on 3′ overhangs of DNA double strand breaks by the human tyrosyl-DNA phosphodiesterase hTdp1', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M204688200.

Interthal, H. *et al.* (2005) 'SCAN1 mutant Tdp1 accumulates the enzyme-DNA intermediate and causes camptothecin hypersensitivity', *EMBO Journal*. doi: 10.1038/sj.emboj.7600694.

Interthal, H. and Champoux, J. J. (2011) 'Effects of DNA and protein size on substrate cleavage by human tyrosyl-DNA phosphodiesterase 1', *Biochemical Journal*. doi: 10.1042/BJ20101841.

Interthal, H., Chen, H. J. and Champoux, J. J. (2005) 'Human Tdp1 cleaves a broad spectrum of substrates, including phosphoamide linkages', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M508898200.

Interthal, H., Pouliot, J. J. and Champoux, J. J. (2001) 'The tyrosyl-DNA phosphodiesterase Tdp1 is a member of the phospholipase D superfamily', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.211429198.

Itoh, T. and Tomizawa, J. (1980) 'Formation of an RNA primer for initiation of replication of ColE1 DNA by ribonuclease H', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.77.5.2450.

Jackson, S. P. and Bartek, J. (2009) 'The DNA-damage response in human biology and disease', *Nature*. doi: 10.1038/nature08467.

Jeanblanc, M. *et al.* (2012) 'Parallel pathways in RAF-induced senescence and conditions for its reversion', *Oncogene*. doi: 10.1038/onc.2011.481.

Jeggo, P. A. and Downs, J. A. (2014) 'Roles of chromatin remodellers in DNA double strand break repair', *Experimental Cell Research*. doi: 10.1016/j.yexcr.2014.09.023.

Karanam, K. *et al.* (2012) 'Quantitative Live Cell Imaging Reveals a Gradual Shift between DNA Repair Mechanisms and a Maximal Use of HR in Mid S Phase', *Molecular Cell*. doi: 10.1016/j.molcel.2012.05.052.

Katyal, S. *et al.* (2007) 'TDP1 facilitates chromosomal single-strand break repair in neurons and is neuroprotective in vivo', *EMBO Journal*. doi: 10.1038/sj.emboj.7601869.

Katyal, S. *et al.* (2014) 'Aberrant topoisomerase-1 DNA lesions are pathogenic in neurodegenerative genome instability syndromes', *Nature Neuroscience*. doi: 10.1038/nn.3715.

Kawale, A. S. *et al.* (2018) 'TDP1 suppresses mis-joining of radiomimetic DNA double-strand breaks and cooperates with Artemis to promote optimal nonhomologous end joining', *Nucleic Acids Research*. doi: 10.1093/nar/gky694.

Kawale, A. S. and Povirk, L. F. (2018) 'Tyrosyl-DNA phosphodiesterases: Rescuing the genome from the risks of relaxation', *Nucleic Acids Research*. doi: 10.1093/nar/gkx1219.

Khanna, K. K. and Jackson, S. P. (2001) 'DNA double-strand breaks: Signaling, repair and the cancer connection', *Nature Genetics*. doi: 10.1038/85798.

Kim, N. *et al.* (2011) 'Mutagenic processing of ribonucleotides in DNA by yeast topoisomerase I', *Science*. doi: 10.1126/science.1205016.

Kinoshita, K. and Honjo, T. (2000) 'Unique and unprecedented recombination mechanisms in class switching', *Current Opinion in Immunology*. doi: 10.1016/S0952-7915(99)00072-2.

Koster, D. A. *et al.* (2005) 'Friction and torque govern the relaxation of DNA supercoils by eukaryotic topoisomerase IB', *Nature*. doi: 10.1038/nature03395.

Kotsantis, P. *et al.* (2016) 'Increased global transcription activity as a mechanism of replication stress in cancer', *Nature Communications*. doi: 10.1038/ncomms13087.

Krah, R. *et al.* (1996) 'A two-subunit type I DNA topoisomerase (reverse gyrase) from an extreme hyperthermophile', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.93.1.106.

Krogh, B. O. and Shuman, S. (2002) 'A poxvirus-like type IB topoisomerase family in bacteria', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.032613199.

Lambo, S. *et al.* (2019) 'The molecular landscape of ETMR at diagnosis and relapse', *Nature*. doi: 10.1038/s41586-019-1815-x.

Lebedeva, N. A., Rechkunova, N. I. and Lavrik, O. I. (2011) 'AP-site cleavage activity of tyrosyl-DNA phosphodiesterase 1', *FEBS Letters*. doi: 10.1016/j.febslet.2011.01.032.

Lee, M. P. *et al.* (1993) 'DNA topoisomerase I is essential in Drosophila melanogaster', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.90.14.6656.

Lemaître, C. and Soutoglou, E. (2014) 'Double strand break (DSB) repair in heterochromatin and heterochromatin proteins in DSB repair', *DNA Repair*. doi: 10.1016/j.dnarep.2014.03.015.

Leppard, J. B. and Champoux, J. J. (2005) 'Human DNA topoisomerase I: Relaxation, roles, and damage control', *Chromosoma*. doi: 10.1007/s00412-005-0345-5.

Lesher, D. T. T. *et al.* (2002) '8-Oxoguanine rearranges the active site of human topoisomerase I', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.192282699.

Li, L., Matsui, M. and Corey, D. R. (2016) 'Activating frataxin expression by repeat-targeted nucleic acids', *Nature Communications*. doi: 10.1038/ncomms10606.

Li, Linhua *et al.* (2017) 'Effect of R119G Mutation on Human P5CR1 Dynamic Property and Enzymatic Activity', *BioMed Research International*. doi: 10.1155/2017/4184106.

Li, M. *et al.* (2013) 'The FHA and BRCT domains recognize ADP-ribosylation during DNA damage response', *Genes and Development*. doi: 10.1101/gad.226357.113.

Li, X. and Manley, J. L. (2005) 'Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability', *Cell*. doi: 10.1016/j.cell.2005.06.008.

Li, X. and Manley, J. L. (2006) 'Cotranscriptional processes and their influence on genome

stability', *Genes and Development*. doi: 10.1101/gad.1438306.

Lin, C. P. *et al.* (2008) 'A ubiquitin-proteasome pathway for the repair of topoisomerase I-DNA covalent complexes', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M803493200.

Lindahl, T. and Barnes, D. E. (2000) 'Repair of endogenous DNA damage', in *Cold Spring Harbor Symposia on Quantitative Biology*. doi: 10.1101/sqb.2000.65.127.

Liu, L. F. *et al.* (1983) 'Cleavage of DNA by mammalian DNA topoisomerase II.', *Journal of Biological Chemistry*.

Liu, L. F. and Wang, J. C. (1987) 'Supercoiling of the DNA template during transcription.', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.84.20.7024.

Ljungman, M. and Lane, D. P. (2004) 'Transcription - Guarding the genome by sensing DNA damage', *Nature Reviews Cancer*. doi: 10.1038/nrc1435.

Löbrich, M. *et al.* (2010) 'gammaH2AX foci analysis for monitoring DNA double-strand break repair', *Cell Cycle*. doi: 10.4161/cc.9.4.10764.

Lu, W. T. *et al.* (2018) 'Drosha drives the formation of DNA:RNA hybrids around DNA break sites to facilitate DNA repair', *Nature Communications*. doi: 10.1038/s41467-018-02893-x.

Luzhna, L., Kathiria, P. and Kovalchuk, O. (2013) 'Micronuclei in genotoxicity assessment: From genetics to epigenetics and beyond', *Frontiers in Genetics*. doi: 10.3389/fgene.2013.00131.

De Magis, A. *et al.* (2019) 'DNA damage and genome instability by G-quadruplex ligands are mediated by R loops in human cancer cells', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.1810409116.

Manzo, S. G. *et al.* (2018) 'DNA Topoisomerase I differentially modulates R-loops across the human genome', *Genome Biology*. doi: 10.1186/s13059-018-1478-1.

Marinello, J. *et al.* (2013) 'Antisense transcripts enhanced by camptothecin at divergent CpG-island promoters associated with bursts of topoisomerase I-DNA cleavage complex and R-loop formation', *Nucleic Acids Research*. doi: 10.1093/nar/gkt778.

Masani, S., Han, L. and Yu, K. (2013) 'Apurinic/Apyrimidinic Endonuclease 1 Is the Essential Nuclease during Immunoglobulin Class Switch Recombination', *Molecular and Cellular Biology*. doi: 10.1128/mcb.00026-13.

McClendon, A. K., Rodriguez, A. C. and Osheroff, N. (2005) 'Human Topoisomerase IIα Rapidly Relaxes Positively Supercoiled DNA', *Journal of Biological Chemistry*. doi: 10.1074/jbc.m503320200.

McElhinny, S. A. N. *et al.* (2010) 'Genome instability due to ribonucleotide incorporation into DNA', *Nature Chemical Biology*. doi: 10.1038/nchembio.424.

Meisenberg, C. *et al.* (2015) 'Clinical and cellular roles for TDP1 and TOP1 in modulating colorectal cancer response to irinotecan', *Molecular Cancer Therapeutics*. doi: 10.1158/1535-7163.MCT-14-0762.

Mendell, J. T. and Dietz, H. C. (2001) 'When the message goes awry: Disease-producing mutations that influence mRNA content and performance', *Cell*. doi: 10.1016/S0092-8674(01)00583-9.

Miao, Z. H. *et al.* (2006) 'Hereditary ataxia SCAN1 cells are defective for the repair of

transcription-dependent topoisomerase I cleavage complexes', *DNA Repair*. doi: 10.1016/j.dnarep.2006.07.004.

Michelini, F. *et al.* (2017) 'Damage-induced lncRNAs control the DNA damage response through interaction with DDRNAs at individual double-strand breaks', *Nature Cell Biology*. doi: 10.1038/ncb3643.

Mischo, H. E. *et al.* (2011) 'Yeast Sen1 helicase protects the genome from transcription-associated instability', *Molecular Cell*. doi: 10.1016/j.molcel.2010.12.007.

Murai, J. *et al.* (2012) 'Tyrosyl-DNA phosphodiesterase 1 (TDP1) repairs DNA damage induced by topoisomerases I and II and base alkylation in vertebrate cells', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M111.333963.

Muramatsu, M. *et al.* (2000) 'Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme', *Cell*. doi: 10.1016/S0092-8674(00)00078-7.

Nicholson, P. *et al.* (2010) 'Nonsense-mediated mRNA decay in human cells: Mechanistic insights, functions beyond quality control and the double-life of NMD factors', *Cellular and Molecular Life Sciences*. doi: 10.1007/s00018-009-0177-1.

Niehrs, C. and Luke, B. (2020) 'Regulatory R-loops as facilitators of gene expression and genome stability', *Nature Reviews Molecular Cell Biology*. doi: 10.1038/s41580-019-0206-3.

Nitiss, J. L. (2009) 'Targeting DNA topoisomerase II in cancer chemotherapy', *Nature Reviews Cancer*. doi: 10.1038/nrc2607.

Nitiss, K. C. *et al.* (2006) 'Tyrosyl-DNA phosphodiesterase (Tdp1) participates in the repair of Top2-mediated DNA damage', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.0603455103.

Ochi, T. *et al.* (2015) 'PAXX, a paralog of XRCC4 and XLF, interacts with Ku to promote DNA double-strand break repair', *Science*. doi: 10.1126/science.1261971.

Ohle, C. *et al.* (2016) 'Transient RNA-DNA Hybrids Are Required for Efficient Double-Strand Break Repair', *Cell*. doi: 10.1016/j.cell.2016.10.001.

Page, F. Le *et al.* (1995) 'Mutagenicity of a unique 8-oxoguanine in a human ha-ras sequence in mammalian cells', *Carcinogenesis*. doi: 10.1093/carcin/16.11.2779.

Paix, A. *et al.* (2017) 'Precision genome editing using synthesis-dependent repair of Cas9-induced DNA breaks', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.1711979114.

Park, E. M. *et al.* (1992) 'Assay of excised oxidative DNA lesions: Isolation of 8-oxoguanine and its nucleoside derivatives from biological fluids with a monoclonal antibody column', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.89.8.3375.

Patel, A. G. *et al.* (2012) 'Enhanced killing of cancer cells by poly(ADP-ribose) polymerase inhibitors and topoisomerase I inhibitors reflects poisoning of both enzymes', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M111.296475.

Perego, M. G. L. *et al.* (2019) 'R-Loops in Motor Neuron Diseases', *Molecular Neurobiology*. doi: 10.1007/s12035-018-1246-y.

Petersen-Mahrt, S. K., Harris, R. S. and Neuberger, M. S. (2002) 'AID mutates E. coli

suggesting a DNA deamination mechanism for antibody diversification', *Nature*. doi: 10.1038/nature00862.

Phillips, D. D. *et al.* (2013) 'The sub-nanomolar binding of DNA-RNA hybrids by the single-chain Fv fragment of antibody S9.6', *Journal of Molecular Recognition*. doi: 10.1002/jmr.2284.

Phillips, P. C. (2008) 'Epistasis - The essential role of gene interactions in the structure and evolution of genetic systems', *Nature Reviews Genetics*. doi: 10.1038/nrg2452.

Piazza, A. and Heyer, W. D. (2019) 'Homologous Recombination and the Formation of Complex Genomic Rearrangements', *Trends in Cell Biology*. doi: 10.1016/j.tcb.2018.10.006.

Plo, I. *et al.* (2003) 'Association of XRCC1 and tyrosyl DNA phosphodiesterase (Tdp1) for the repair of topoisomerase I-mediated DNA lesions', *DNA Repair*. doi: 10.1016/S1568-7864(03)00116-2.

Pommier, Y. *et al.* (1998) 'Mechanism of action of eukaryotic DNA topoisomerase I and drugs targeted to the enzyme', *Biochimica et Biophysica Acta - Gene Structure and Expression*. doi: 10.1016/S0167-4781(98)00129-8.

Pommier, Y. *et al.* (1999) 'Topoisomerase I inhibitors: Selectivity and cellular resistance', *Drug Resistance Updates*. doi: 10.1054/drup.1999.0102.

Pommier, Y. *et al.* (2003) 'Repair of and checkpoint response to topoisomerase I-mediated DNA damage', *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis*. doi: 10.1016/j.mrfmmm.2003.08.016.

Pommier, Y. *et al.* (2006) 'Repair of Topoisomerase I-Mediated DNA Damage', *Progress in Nucleic Acid Research and Molecular Biology*. doi: 10.1016/S0079-6603(06)81005-6.

Pommier, Y. (2006) 'Topoisomerase I inhibitors: Camptothecins and beyond', in *Nature Reviews Cancer*. doi: 10.1038/nrc1977.

Pommier, Y. *et al.* (2010) 'DNA topoisomerases and their poisoning by anticancer and antibacterial drugs', *Chemistry and Biology*. doi: 10.1016/j.chembiol.2010.04.012.

Pommier, Y. (2013) 'Drugging topoisomerases: Lessons and Challenges', *ACS Chemical Biology*. doi: 10.1021/cb300648v.

Pommier, Y. *et al.* (2014) 'Tyrosyl-DNA-phosphodiesterases (TDP1 and TDP2)', *DNA Repair*. doi: 10.1016/j.dnarep.2014.03.020.

Pommier, Y. *et al.* (2016) 'Roles of eukaryotic topoisomerases in transcription, replication and genomic stability', *Nature Reviews Molecular Cell Biology*. doi: 10.1038/nrm.2016.111.

Pommier, Y. and Marchand, C. (2005) 'Interfacial inhibitors of protein-nucleic acid interactions', *Current Medicinal Chemistry - Anti-Cancer Agents*. doi: 10.2174/1568011054222337.

Ponting, C. P. and Kerr, I. D. (2008) 'A novel family of phospholipase D homologues that includes phospholipid synthases and putative endonucleases: Identification of duplicated repeats and potential active site residues', *Protein Science*. doi: 10.1002/pro.5560050513.

Pouliot, J. J. *et al.* (1999) 'Yeast gene for a Tyr-DNA phosphodiesterase that repairs topoisomerase I complexes', *Science*. doi: 10.1126/science.286.5439.552.

Pourquier, P., Ueng, L. M., *et al.* (1997) 'Effects of uracil incorporation, DNA mismatches, and abasic sites on cleavage and religation activities of mammalian topoisomerase I', *Journal of Biological Chemistry*. doi: 10.1074/jbc.272.12.7792.

Pourquier, P., Pilon, A. A., *et al.* (1997) 'Trapping of mammalian topoisomerase I and recombinations induced by damaged DNA containing nicks or gaps. Importance of DNA end phosphorylation and camptothecin effects', *Journal of Biological Chemistry*. doi: 10.1074/jbc.272.42.26441.

Pourquier, P. *et al.* (1999) 'Induction of reversible complexes between eukaryotic DNA topoisomerase I and DNA-containing oxidative base damages: 7,8-dihydro-8-oxoguanine and 5- hydroxycytosine', *Journal of Biological Chemistry*. doi: 10.1074/jbc.274.13.8516.

Pourquier, P. and Lansiaux, A. (2019) 'Molecular determinants of response to topoisomerase I inhibitors', *Bulletin du Cancer*. doi: 10.1684/bdc.2011.1474.

Pourquier, P. and Pommier, Y. (2001) 'Topoisomerase I-mediated DNA damage', *Advances in Cancer Research*. doi: 10.1016/s0065-230x(01)80016-6.

Povirk, L. F. (1996) 'DNA damage and mutagenesis by radiomimetic DNA-cleaving agents: Bleomycin, neocarzinostatin and other enediynes', *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis*. doi: 10.1016/0027-5107(96)00023-1.

Powell, W. T. *et al.* (2013) 'R-loop formation at Snord116 mediates topotecan inhibition of Ube3a-antisense and allele-specific chromatin decondensation', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.1305426110.

Prakash, A. and Doublié, S. (2015) 'Base Excision Repair in the Mitochondria', *Journal of Cellular Biochemistry*. doi: 10.1002/jcb.25103.

Puget, N., Miller, K. M. and Legube, G. (2019) 'Non-canonical DNA/RNA structures during Transcription-Coupled Double-Strand Break Repair: Roadblocks or Bona fide repair intermediates?', *DNA Repair*. doi: 10.1016/j.dnarep.2019.102661.

Purmal, A. A., Kow, Y. W. and Wallace, S. S. (1994) 'Major oxidative products of cytosine, 5-hydroxycytosine and 5-hydroxyuracil, exhibit sequence context-dependent mispairing in vitro', *Nucleic Acids Research*. doi: 10.1093/nar/22.1.72.

Quan, L., Lv, Q. and Zhang, Y. (2016) 'STRUM: Structure-based prediction of protein stability changes upon single-point mutation', *Bioinformatics*. doi: 10.1093/bioinformatics/btw361.

Ran, F. A. *et al.* (2013) 'Genome engineering using the CRISPR-Cas9 system', *Nature Protocols*. doi: 10.1038/nprot.2013.143.

Rass, U., Ahel, I. and West, S. C. (2007) 'Defective DNA Repair and Neurodegenerative Disease', *Cell*. doi: 10.1016/j.cell.2007.08.043.

Ratmeyer, L. *et al.* (1994) 'Sequence Specific Thermodynamic and Structural Properties for DNA·RNA Duplexes', *Biochemistry*. doi: 10.1021/bi00183a037.

Ray Chaudhuri, A. and Nussenzweig, A. (2017) 'The multifaceted roles of PARP1 in DNA repair and chromatin remodelling', *Nature Reviews Molecular Cell Biology*. doi: 10.1038/nrm.2017.53.

Raymond, A. C. *et al.* (2004) 'Analysis of human tyrosyl-DNA phosphodiesterase I catalytic residues', *Journal of Molecular Biology*. doi: 10.1016/j.jmb.2004.03.013.

Raymond, A. C., Staker, B. L. and Burgin, A. B. (2005) 'Substrate specificity of tyrosyl-DNA phosphodiesterase I (Tdp1)', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M502148200.

Reaban, M. E., Lebowitz, J. and Griffin, J. A. (1994) 'Transcription induces the formation of a stable RNA·DNA hybrid in the immunoglobulin α switch region', *Journal of Biological*

Chemistry.

Redon, C. *et al.* (2002) 'Histone H2A variants H2AX and H2AZ', *Current Opinion in Genetics and Development*. doi: 10.1016/S0959-437X(02)00282-4.

Regairaz, M. *et al.* (2011) 'Mus81-mediated DNA cleavage resolves replication forks stalled by topoisomerase I-DNA complexes', *Journal of Cell Biology*. doi: 10.1083/jcb.201104003.

Richard, P. and Manley, J. L. (2017) 'R Loops and Links to Human Disease', *Journal of Molecular Biology*. doi: 10.1016/j.jmb.2016.08.031.

Rideout, M. C., Raymond, A. C. and Burgin, A. B. (2004) 'Design and synthesis of fluorescent substrates for human tyrosyl-DNA phosphodiesterase I', *Nucleic Acids Research*. doi: 10.1093/nar/gkh796.

Roberts, R. W. and Crothers, D. M. (1992) 'Stability and properties of double and triple helices: Dramatic effects of RNA or DNA backbone composition', *Science*. doi: 10.1126/science.1279808.

Rossi, F. *et al.* (1996) 'Specific phosphorylation of SR proteins by mammalian DNA topoisomerase I', *Nature*. doi: 10.1038/381080a0.

Roy, D. *et al.* (2010) 'Competition between the RNA Transcript and the Nontemplate DNA Strand during R-Loop Formation In Vitro: a Nick Can Serve as a Strong R-Loop Initiation Site', *Molecular and Cellular Biology*. doi: 10.1128/mcb.00897-09.

Roy, D. and Lieber, M. R. (2009) 'G Clustering Is Important for the Initiation of Transcription-Induced R-Loops In Vitro, whereas High G Density without Clustering Is Sufficient Thereafter', *Molecular and Cellular Biology*. doi: 10.1128/mcb.00139-09.

Sakai, A. *et al.* (2012) 'PARP and CSB modulate the processing of transcription-mediated DNA strand breaks', *Genes and Genetic Systems*. doi: 10.1266/ggs.87.265.

Salas-Armenteros, I. *et al.* (2017) ' Human THO –Sin3A interaction reveals new mechanisms to prevent R-loops that cause genome instability ', *The EMBO Journal*. doi: 10.15252/embj.201797208.

Sanjana, N. E., Shalem, O. and Zhang, F. (2014) 'Improved vectors and genome-wide libraries for CRISPR screening', *Nature Methods*. doi: 10.1038/nmeth.3047.

Santos-Pereira, J. M. and Aguilera, A. (2015) 'R loops: New modulators of genome dynamics and function', *Nature Reviews Genetics*. doi: 10.1038/nrg3961.

Sanz, L. A. *et al.* (2016) 'Prevalent, Dynamic, and Conserved R-Loop Structures Associate with Specific Epigenomic Signatures in Mammals', *Molecular Cell*. doi: 10.1016/j.molcel.2016.05.032.

Sas, K. *et al.* (2007) 'Mitochondria, metabolic disturbances, oxidative stress and the kynurenine system, with focus on neurodegenerative disorders', *Journal of the Neurological Sciences*. doi: 10.1016/j.jns.2007.01.033.

Scully, R. *et al.* (2019) 'DNA double-strand break repair-pathway choice in somatic mammalian cells', *Nature Reviews Molecular Cell Biology*. doi: 10.1038/s41580-019-0152-0.

Sekiguchi, J. A. and Shuman, S. (1997) 'Site-specific ribonuclease activity of eukaryotic DNA topoisomerase I', *Molecular Cell*. doi: 10.1016/S1097-2765(00)80010-6.

Seol, Y. *et al.* (2012) 'A kinetic clutch governs religation by type IB topoisomerases and determines camptothecin sensitivity', *Proceedings of the National Academy of Sciences of the*

United States of America. doi: 10.1073/pnas.1206480109.

Shen, J. *et al.* (2010) 'Mutations in PNKP cause microcephaly, seizures and defects in DNA repair', *Nature Genetics*. doi: 10.1038/ng.526.

Shiloh, Y. (2006) 'The ATM-mediated DNA-damage response: taking shape', *Trends in Biochemical Sciences*. doi: 10.1016/j.tibs.2006.05.004.

Skourti-Stathaki, K., Kamieniarz-Gdula, K. and Proudfoot, N. J. (2014) 'R-loops induce repressive chromatin marks over mammalian gene terminators', *Nature*. doi: 10.1038/nature13787.

Skourti-Stathaki, K. and Proudfoot, N. J. (2014) 'A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression', *Genes and Development*. doi: 10.1101/gad.242990.114.

Skourti-Stathaki, K., Proudfoot, N. J. and Gromak, N. (2011) 'Human Senataxin Resolves RNA/DNA Hybrids Formed at Transcriptional Pause Sites to Promote Xrn2-Dependent Termination', *Molecular Cell*. doi: 10.1016/j.molcel.2011.04.026.

Sobek, S. *et al.* (2013) 'Negative regulation of mitochondrial transcription by mitochondrial topoisomerase i', *Nucleic Acids Research*. doi: 10.1093/nar/gkt768.

Sollier, J. *et al.* (2014) 'Transcription-Coupled Nucleotide Excision Repair Factors Promote R-Loop-Induced Genome Instability', *Molecular Cell*. doi: 10.1016/j.molcel.2014.10.020.

Sollier, J. and Cimprich, K. A. (2015) 'Breaking bad: R-loops and genome integrity', *Trends in Cell Biology*. doi: 10.1016/j.tcb.2015.05.003.

Sordet, O. *et al.* (2008) 'Hyperphosphorylation of RNA Polymerase II in Response to Topoisomerase I Cleavage Complexes and Its Association with Transcription- and BRCA1-dependent Degradation of Topoisomerase I', *Journal of Molecular Biology*. doi: 10.1016/j.jmb.2008.06.028.

Sordet, O. *et al.* (2009) 'Ataxia telangiectasia mutated activation by transcription- and topoisomerase I-induced DNA double-strand breaks', *EMBO Reports*. doi: 10.1038/embor.2009.97.

Sordet, O. *et al.* (2010) 'DNA double-strand breaks and ATM activation by transcription-blocking DNA lesions', *Cell Cycle*. doi: 10.4161/cc.9.2.10506.

Sridhara, S. C. *et al.* (2017) 'Transcription Dynamics Prevent RNA-Mediated Genomic Instability through SRPK2-Dependent DDX23 Phosphorylation', *Cell Reports*. doi: 10.1016/j.celrep.2016.12.050.

Staker, B. L. *et al.* (2002) 'The mechanism of topoisomerase I poisoning by a camptothecin analog', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.242259599.

Stavnezer, J., Guikema, J. E. J. and Schrader, C. E. (2008) 'Mechanism and Regulation of Class Switch Recombination', *Annual Review of Immunology*. doi: 10.1146/annurev.immunol.26.021607.090248.

Stewart, L. *et al.* (1998) 'A model for the mechanism of human topoisomerase I', *Science*. doi: 10.1126/science.279.5356.1534.

Stork, C. T. *et al.* (2016) 'Co-transcriptional R-loops are the main cause of estrogen-induced DNA damage', *eLife*. doi: 10.7554/eLife.17548.

Straub, T. *et al.* (1998) 'The RNA-splicing factor PSF/p54(nrb) controls DNA-topoisomerase I activity by a direct interaction', *Journal of Biological Chemistry*. doi: 10.1074/jbc.273.41.26261.

Strumberg, D. *et al.* (2000) 'Conversion of Topoisomerase I Cleavage Complexes on the Leading Strand of Ribosomal DNA into 5'-Phosphorylated DNA Double-Strand Breaks by Replication Runoff', *Molecular and Cellular Biology*. doi: 10.1128/mcb.20.11.3977-3987.2000.

Stuckey, J. A. and Dixon, J. E. (1999) 'Crystal structure of a phospholipase D family member', *Nature Structural Biology*. doi: 10.1038/6716.

Sun, Q. *et al.* (2013) 'R-loop stabilization represses antisense transcription at the arabidopsis FLC locus', *Science*. doi: 10.1126/science.1234848.

Symington, L. S. and Gautier, J. (2011) 'Double-Strand Break End Resection and Repair Pathway Choice', *Annual Review of Genetics*. doi: 10.1146/annurev-genet-110410-132435.

Takashima, H. *et al.* (2002) 'Mutation of TDP1, encoding a topoisomerase I-dependent DNA damage repair enzyme, in spinocerebellar ataxia with axonal neuropathy', *Nature Genetics*. doi: 10.1038/ng987.

Thomas, M., White, R. L. and Davis, R. W. (1976) 'Hybridization of RNA to double stranded DNA: Formation of R loops', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.73.7.2294.

Tuduri, S. *et al.* (2009) 'Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription', *Nature Cell Biology*. doi: 10.1038/ncb1984.

Uemura, T. and Yanagida, M. (1984) 'Isolation of type I and II DNA topoisomerase mutants from fission yeast: single and double mutants show different phenotypes in cell growth and chromatin organization.', *The EMBO Journal*. doi: 10.1002/j.1460-2075.1984.tb02040.x.

Vance, J. R. and Wilson, T. E. (2002) 'Yeast Tdp1 and Rad1-Rad10 function as redundant pathways for repairing Top1 replicative damage', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.202242599.

Wahba, L. *et al.* (2016) 'S1-DRIP-seq identifies high expression and polyA tracts as major contributors to R-loop formation', *Genes and Development*. doi: 10.1101/gad.280834.116.

Wall, M. E. *et al.* (1966) 'Plant Antitumor Agents. I. The Isolation and Structure of Camptothecin, a Novel Alkaloidal Leukemia and Tumor Inhibitor from Camptotheca acuminata', *Journal of the American Chemical Society*. doi: 10.1021/ja00968a057.

Walton, C. *et al.* (2010) 'Spinocerebellar ataxia with axonal neuropathy', *Advances in Experimental Medicine and Biology*. doi: 10.1007/978-1-4419-6448-9_7.

Wang, J. C. (1971) 'Interaction between DNA and an Escherichia coli protein ω', *Journal of Molecular Biology*. doi: 10.1016/0022-2836(71)90334-2.

Wang, J. C. (2002) 'Cellular roles of DNA topoisomerases: A molecular perspective', *Nature Reviews Molecular Cell Biology*. doi: 10.1038/nrm831.

Weterings, E. and Chen, D. J. (2007) 'DNA-dependent protein kinase in nonhomologous end joining: A lock with multiple keys?', *Journal of Cell Biology*. doi: 10.1083/jcb.200705106.

Whitehouse, C. J. *et al.* (2001) 'XRCC1 stimulates human polynucleotide kinase activity at damaged DNA termini and accelerates DNA single-strand break repair', *Cell*. doi:

10.1016/S0092-8674(01)00195-7.

Williams, J. S., Lujan, S. A. and Kunkel, T. A. (2016) 'Processing ribonucleotides incorporated during eukaryotic DNA replication', *Nature Reviews Molecular Cell Biology*. doi: 10.1038/nrm.2016.37.

Wu, J. and Liu, L. F. (1997) 'Processing of topoisomerase I cleavable complexes into DNA damage by transcription', *Nucleic Acids Research*. doi: 10.1093/nar/25.21.4181.

Wyman, C. and Kanaar, R. (2006) 'DNA Double-Strand Break Repair: All's Well that Ends Well', *Annual Review of Genetics*. doi: 10.1146/annurev.genet.40.110405.090451.

Xu, W. *et al.* (2017) 'The R-loop is a common chromatin feature of the Arabidopsis genome', *Nature Plants*. doi: 10.1038/s41477-017-0004-x.

Xu, Y. and Her, C. (2015) 'Inhibition of topoisomerase (DNA) I (TOP1): DNA damage repair and anticancer therapy', *Biomolecules*. doi: 10.3390/biom5031652.

Yang, S. W. *et al.* (1996) 'A eukaryotic enzyme that can disjoin dead-end covalent complexes between DNA and type I topoisomerases', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.93.21.11534.

Yasuhara, T. *et al.* (2018) 'Human Rad52 Promotes XPG-Mediated R-loop Processing to Initiate Transcription-Associated Homologous Recombination Repair', *Cell*. doi: 10.1016/j.cell.2018.08.056.

Yu, K. *et al.* (2003) 'R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells', *Nature Immunology*. doi: 10.1038/ni919.

Zeng, Z. *et al.* (2012) 'TDP2 promotes repair of topoisomerase I-mediated DNA damage in the absence of TDP1', *Nucleic Acids Research*. doi: 10.1093/nar/gks622.

Zhang, H. *et al.* (2014) 'Increased negative supercoiling of mtDNA in TOP1mt knockout mice and presence of topoisomerases IIα and IIβ in vertebrate mitochondria', *Nucleic Acids Research*. doi: 10.1093/nar/gku384.

Zhang, Y. W. *et al.* (2011) 'Poly(ADP-ribose) polymerase and XPF-ERCC1 participate in distinct pathways for the repair of topoisomerase I-induced DNA damage in mammalian cells', *Nucleic Acids Research*. doi: 10.1093/nar/gkq1304.

Zhou, T. *et al.* (2005) 'Deficiency in 3′-phosphoglycolate processing in human cells with a hereditary mutation in tyrosyl-DNA phosphodiesterase (TDP1)', *Nucleic Acids Research*. doi: 10.1093/nar/gki170.

Zhou, T. *et al.* (2009) 'Tyrosyl-DNA phosphodiesterase and the repair of 3′-phosphoglycolate-terminated DNA double-strand breaks', *DNA Repair*. doi: 10.1016/j.dnarep.2009.05.003.

Zierhut, C. and Diffley, J. F. X. (2008) 'Break dosage, cell cycle stage and DNA replication influence DNA double strand break response', *EMBO Journal*. doi: 10.1038/emboj.2008.111.