Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

Biologia Cellulare e Molecolare

Ciclo XXX

Settore Concorsuale di afferenza: 05/E2 BIOLOGIA MOLECOLARE

Settore Scientifico disciplinare: BIO/11 BIOLOGIA MOLECOLARE

G-quadruplex binders cause DNA damage by inducing R-loops in human cancer cells

Presentata da:

ALESSIO DE MAGIS

Supervisore

Prof. Giovanni Capranico

Coordinatore Dottorato

Prof. Giovanni Capranico

Esame finale anno 2018

CONTENTS

ABSTRACT

G-Quadruplexes (G4s) and R-loops are non-B DNA structures that can regulate transcription and replication. G4s are formed from four guanine residues that are held together in the same plane by Hoogsteen hydrogen bonds and further stabilized by the presence of monovalent cations. R-loops are triple-strand structures that contain an RNA-DNA hybrid and displaced single-stranded DNA. One of the most important features that influence these DNA structures is the GC content. Indeed, R-loop structures can be favoured by a high guanine density in the non-template DNA strand and this is specifically due to the higher thermodynamic stability of RNA-DNA hybrid. R-loops and G4s are generally regarded as highly deleterious, indeed the structures can block both transcription and DNA replication, creating replicative stress and potentially causing DNA damage. Here, we used immunofluorescence analysis in order to identify the increase of G4s and R-loops in cancer cells treated with specific G4 binders at long and very short time. At long time, the increase of these two non-B DNA structures triggers genomic DNA damage as established by the formation of γH2AX foci and other markers of cellular DNA damage response. Interestingly, stable and transient overexpression of RNaseH, an enzyme that specifically removes R-loop structures, induce a rescue of G4 binder-induced DNA damage and genome instability. Our study provides the first direct evidence of a mechanistic link between G4s and R-loops in human cancer cells.

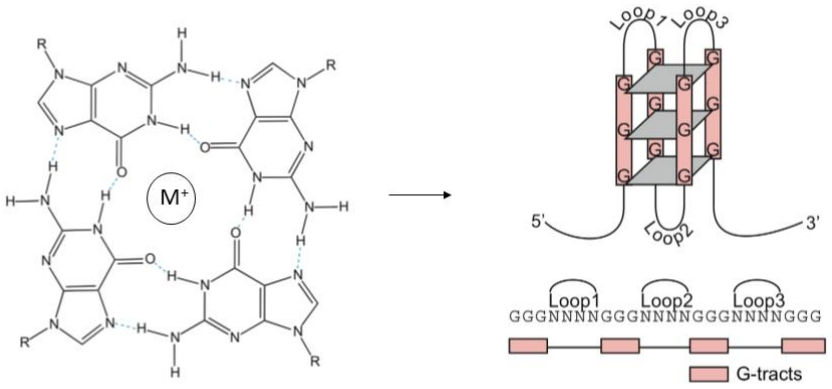
1. INTRODUCTION

1.1 Non-B DNA Structure

The discovery of the canonical B-DNA structure, formed by a right-handed double-helical structure (Watson & Crick, 1953), is a milestone in the history of molecular biology. In the 70s, the results obtained with circular dichroism (CD), single-crystal X-ray and other biophysical techniques documented that DNA double helix is highly polymorphic. Several experimental data established that DNA can adopt alternative conformations such as Z-DNA (Gessner, Frederick, Quigley, Rich, & Wang, 1989), cruciform (Paleček, 1991), G-Quadruplex (Gellert, Lipsett, & Davies, 1962) and R-loop structures (Dson, 1975) (Thomas, White, & Davis, 1976) showing that, rather than being a stable and constant structure, DNA is highly dynamic. Several laboratories have then focused on the question of which function(s) all these non-canonical secondary DNA structures have in living cells. My PhD research project has been focused on two different non-B DNA structures: G-Quadruplex and R-loop, and their interactions in relation to DNA damage and genome instability.

1.1.1 G-Quadruplex Structure

G-Quadruplex or G4 are four guanines that are held together in the same plane by Hoogsteen hydrogen bonds and further stabilized by the presence of monovalent cations. The first demonstration of this structural motif was reported in 1962, when Gellert and colleagues used X-ray diffraction to show that guanines can adopt a tetrameric structure (Figure 1) on three-dimensional space (Gellert et al., 1962).



Edited by (Capra, Paeschke, Singh, & Zakian, 2010)

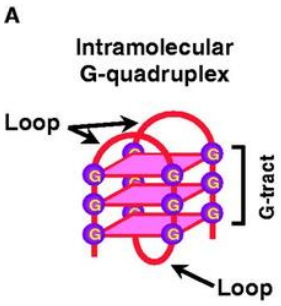
Image under CC BY permission use

Figure 1. At left, four guanine residues connected in the same plane by hydrogen bond and stabilized by a monovalent cations. At right G-Quadruplex structure in three dimensional space formed by three stack of guanine connected with 1-7 nucleotide loops.

The existence in solution of the planar tetramers could then result in the formation of linear aggregates formed by stacking the tetramers on top of each other, since the large planar surfaces would result in strong van der Waals attractions (Gellert et al., 1962).This strong attraction between each stack caused a high thermodynamic stability of G4 structures.

G4s can adopt different conformations on 3D space, which depends on the orientation of the DNA strands. CD analyses allow to discriminate between parallel and anti-parallel strand orientation that depends on different arrangements of anti/syn glycosidic angles (Burge, Parkinson, Hazel, Todd, & Neidle, 2006).

G4s can be grouped into two different categories, i.e. intramolecular and intermolecular structures, according to the number of DNA (or RNA) strands involved in the assembly of the structures. A single nucleic acid strand connecting four G-tracts can fold into an intramolecular G-quadruplex containing a stack of guanine quartets (G-quartet) linked by three loops (Figure 2).

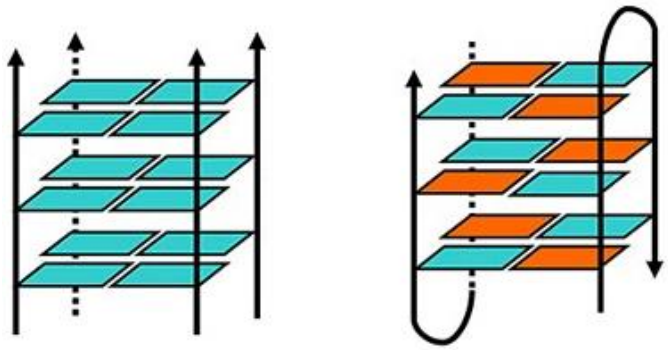


Edited by (Xiao et al., 2013)

Image under CC BY permission use

Figure 2. Intramolecular G-Quadruplex containing a stack of guanine quartets (G-quartet) linked by three loops

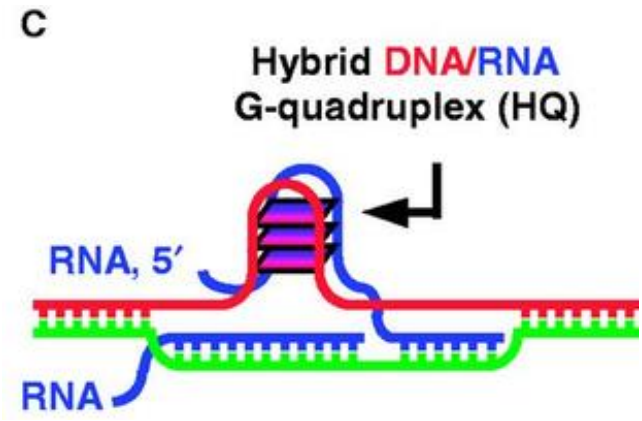
On the other hand, four G-tracts on multiple nucleic acid strands can form intermolecular G-quadruplex (Figure 3) (Xiao et al., 2013).



Permission requested

Figure 3. Intermolecular G-Quadruplex containing stack of guanine quartets (G-quartet) derived for different nucleic acid strands

It has also been suggested that G4 can be constituted by both RNA and DNA strands. Such intermolecular G-quadruplexes can arise when transcription of double-stranded DNA (dsDNA) readily produces DNA:RNA hybrid G-quadruplexes (HQ) by G-tracts from both the non-template DNA strand and the nascent RNA transcript (Figure 4) (Xiao et al., 2013).



Edited by (Xiao et al., 2013)

Image under CC BY permission use

Figure 4. DNA:RNA hybrid G-Quadruplex containing a stack of guanine quartets (G-quartet) in which one strand derived from a RNA nascent and the other strand from a DNA template or not template strand

Topology and stability of all these different G4 conformations depend on many factor, including the nature of the binding cations and the size of base loops between guanine triplets. Their presence in in vitro systems is clear, however it is important to establish the presence of G-Quadruplex in living cells and their physiological functions.

1.1.1.1 G4 motifs in the genome

In the last two decades, much effort has been put into the computational predictions of G-quadruplex structures based on nucleotide sequences. In 2005, Todd and colleagues using simple algorithms developed a tool called Quadparser and identified over 300,000 motif sequence in the human genome that are prone to form G4 structure (Todd, Johnston, & Neidle, 2005). Quadparser can identify canonical G4 structures as it searches for genomic sequences corresponding to a consensus made of four runs of guanine separated by few nucleotides:

$$G_{3-5} - N_{1-7} - G_{3-5} - N_{1-7} - G_{3-5} - N_{1-7} - G_{3-5}$$

Where G_{3-5} are a guanine residues between 3- and 5-base long, and N_{1-7} is a stretch of generic nucleotides (may or may not themselves contain guanines) that creates a loop between the stack of guanines involved in G tetrads.

This is not the only program that allows the prediction of the presence of G4 motifs. Another recent algorithm, called G4 Hunter, predicted that the number of putative G4 sequences is substantially higher than G4 sequences predicted by Quadparser in the human genome. However, G4 Hunter called more G4 motifs with a greater false positive rate than Quadparser (Bedrat, Lacroix, & Mergny, 2016).

These computational studies confirmed that G4 motifs are enriched in telomeres where it is well established that a unimolecular G4 is present and increases the stability of the single-strand DNA (ssDNA).

Furthermore, Sahakyan and co-workers describe a machine learning-based computational tool to predict DNA G4 formation using very large datasets of experimental G4-forming sequences. Their model was able to avoid the prediction of many previously-accepted sequences that do not however form G4 structures, while correctly assessing G4 folding of more than 700,000 sites in the human genome. Furthermore, their approach reveals the relative importance of sequence features from both G4 motifs as well as adjacent regions. The algorithm can be applied to any DNA or genome sequence to characterize the intramolecular sequence-guided G4 formation propensities (Sahakyan et al., 2017).

However, computational findings must be supported and confirmed by experimental data. For this reason, in the last decade, several G4-selective probes have been developed in order to recognize G4s in human cells through different techniques. In 2013, Biffi and colleagues describe an engineered, structure-specific antibody that binds with high selectivity and low nanomolar affinity to DNA G-quadruplex structures. Using ELISA, they showed that BG4 has a high affinity for intramolecular and intermolecular DNA G-quadruplexes with no detectable binding to an RNA hairpin, single-stranded DNA or double-stranded DNA. Other analyses indicate that BG4 is a G-quadruplex structure-specific antibody that does not have a preference for any particular G-quadruplex conformation. In the end, BG4 was used to visualize DNA G-quadruplex structures in human cells showing a punctate nuclear staining (Biffi, Tannahill, McCafferty, & Balasubramanian, 2013). This antibody is still the most widely-used tool to recognize these structures in immunofluorescence and other assay.

A first genome-wide determination of G4 structures was published in 2015 (Chambers et al., 2015). They used a high-resolution sequencing–based method adapting a polymerase-stop assay to detect G4 structures in human genomic DNA extracted from primary human B lymphocytes. The method

rationale is based on polymerase stalling when it encounters an obstacle, which is constituted in this case by a G4 structure in the template strand. They compared the sites of polymerase stops before and after G4 stabilization by adding K+ or the G4-stabilizing ligand pyridostatin (PDS) to the buffer reactions. By high-throughput sequencing, they identified 716.310 putative G4 structures, 451646 of which were not predicted by computational methods.

Next, Balasubramanian's lab has published the determination of G4 structures in the genome of human cells by performing a ChiP-Seq procedures with the use of a specific antibody (BG4) against guanine tetrads (Hänsel-Hertsch et al., 2016). The G4 ChIP findings are somewhat different form previously published method as the determination of G4 is made directly in chromatin of living cells. They found ~10000 G4 structures in human chromatin, which were enriched in regulatory, nucleosome-depleted regions. Furthermore, they find that G4 structures are more frequent in the promoters and 5′ UTRs of highly transcribed genes, and the enrichment is more significant in genes related to cancer, such as MYC. Moreover, some oncogenic promoter regions show a high level of G4 motifs in comparison to the genome overall. This observation suggested the possibility to interfere with oncogene expression to develop novel molecules with the potential to kill tumor cells selectively (Todd et al., 2005) (Bedrat et al., 2016).

1.1.1.2 G4 Binders

In literature there are more than 700/800 G4 binders selected for their ability to bind G4 with different topology or to tie in different positions on the genome. In the 1990, it was already known that G4 structures were able to stabilize single-stranded DNA at the telomeric level. By mid-1990s, there was an increase of interest in the unusual topology at the ends of telomeric DNAs from a wide-range of eukaryotic species. At the same time, many laboratories have been focused on the discovery of small molecules that could specifically stabilize G4 structures at telomeres. Since 1997, Stephen Neidle's laboratory has synthesized a series of regioisomers of disubstituted amidoanthraquinones that have been well characterized for their G4 binding properties (Neidle, 2015). Then, the availability of NMR structure of human telomeric G4 led to a rational design of effective and selective binders of telomeric G4. The Neidle's group synthesized for the first time a ligand with more than two substituents as a side chain. (Neidle, 2015). Among those, a most selective compound was a 3,6,9 - trisubstituted acridine derivate, called BRACO-19. The compound presents a pyrrolidine end group on each side of the acridine moiety and an uncharged aninilino substituent at the central 9-position (Figure 5) (Read et al., 2001).

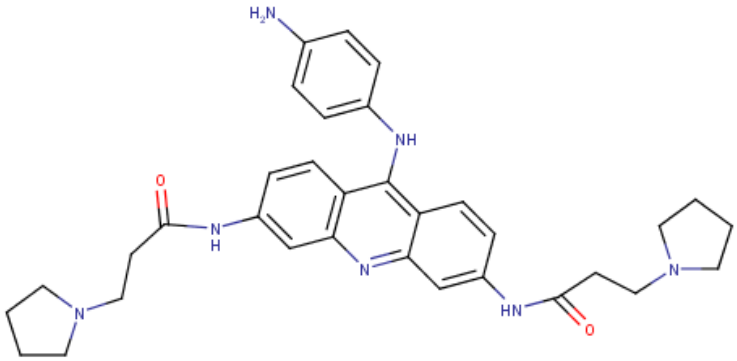


Figure 5. Braco-19 structure, a 3,6,9 - trisubstituted acridine derivate containing two pyrrolidine end groups and an uncharged aninilino substituent at the 9-position.

A different general pharmacophore was then developed by the Balasubramanian group in the late 10s of this century. The intensive search for G4 ligands led to the production of a small molecule with particular features: the participation of hydrogen bonding and the creation of an electron-rich scaffold that creates a "pocket" recognizing G4 structures. The leading compound in this series is Pyridostatin (PDS, Figure 6) (Rodriguez et al., 2008). PDS is one of the most widely-used G4 stabilizers in the literature.

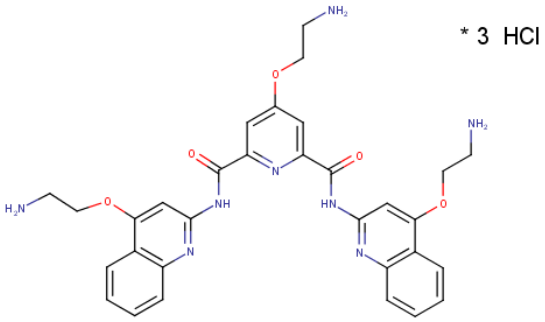


Figure 6. PDS structure, a 4-(2-Aminoethoxy)-N2,N6-bis(4-(2-aminoethoxy)quinolin-2-yl)pyridine-2,6-dicarboxamide; 4-(2-aminoethoxy)-N2,N6-bis[4-(2-aminoethoxy)quinolin-2-yl]pyridine-2,6-dicarboxamide.

In 2010, a collaboration between Rambaldi's (University of Bologna) and Neidle's laboratories produces a paper reporting on a screening of a small group of diimidazo[1,2-a:1,2-c]pyrimidine derivatives for G4 binding activity. FRET analyses showed that the bis-guanylhydrazone derivative of diimidazo[1,2-a:1,2-c]pyrimidine (Figure 7) was a potent stabilizer of several G4 DNAs, whereas

marginal interaction with duplex DNA was found. In this same work, the authors found that the compound 3 (Figure 7B), called FA, did not bind G4 structure selectively.

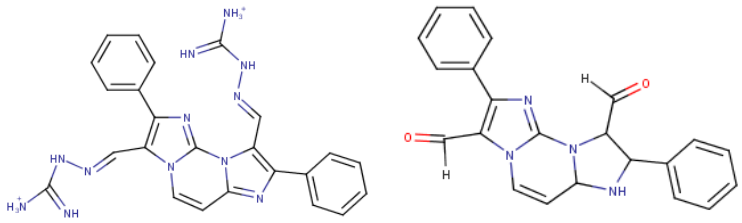


Figure 7. Two different derivative of diimidazo[1,2-a:1,2-c]pyrimidine. On the left, the chemical structure of FG with a bis-guanylhydrazone as a side chain. On the right, the chemical structure of FA with no side chain.

Starting from the lead compound FG, a series of analogues with structural modifications of the substituents have been designed, synthesized, and evaluated as effective and selective G4 ligands (Amato et al., 2016). In this paper, our laboratory collaborated with Rambaldi's and Randazzo's (University of Naples) teams, and biological experiments demonstrated that hydrazone derivatives can effectively trap and stabilized G4 structures in living human cancer U2OS and HCT116 cells (Amato et al., 2016). Remarkably, we found that a derivative (**3**, in the article) with only one side chain (Figure 8) is more effective than FG (**1**, in the article) in inhibiting human U2OS and HCT116 cancer cell growth as well as in stabilizing G4s in living cells (Amato et al., 2016).

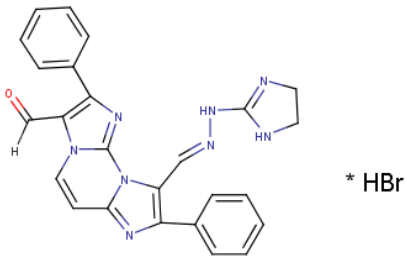


Figure 8. Chemical structure of Compound 3, an other derivative of diimidazo[1,2-a:1,2-c]pyrimidine.

In recent years, a plethora of novel structures, non-canonical G-dependent structures, have been demonstrated, and new molecules have also been developed that bind them. In collaboration with Randazzo's group, we have recently published a paper reporting on a compound, B3F10 (Figure 9), able to bind to G4 as well as to G-triplets (G3). CD analysis showed a significant interaction with G3 and G4 structure, but no significant interaction with duplex DNA. The G4 interaction was support by immunofluorescence experiment in living cells. The dual G3/G4 stabilizing properties may represent a new route for the design of G-rich DNA-targeting compounds, thus stimulating further studies aimed at the development of novel anticancer drugs (Amato et al., 2017).

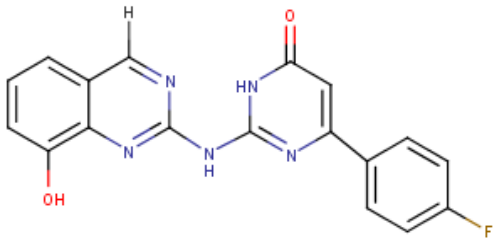


Figure 9. B3F10 structure. The chemical structure presents a dihydropyrimidin-4-one derivative (6-(4-fluorophenyl)-2-[(8-methoxy-4-methylquinazolin-2-yl)amino]-3,4-dihydropyrimidin-4-one

All these compounds with different structures lead to the stabilization of G-Quadruplex. This stabilization seems to play an important role in key processes for the cell, such as replication and transcription.

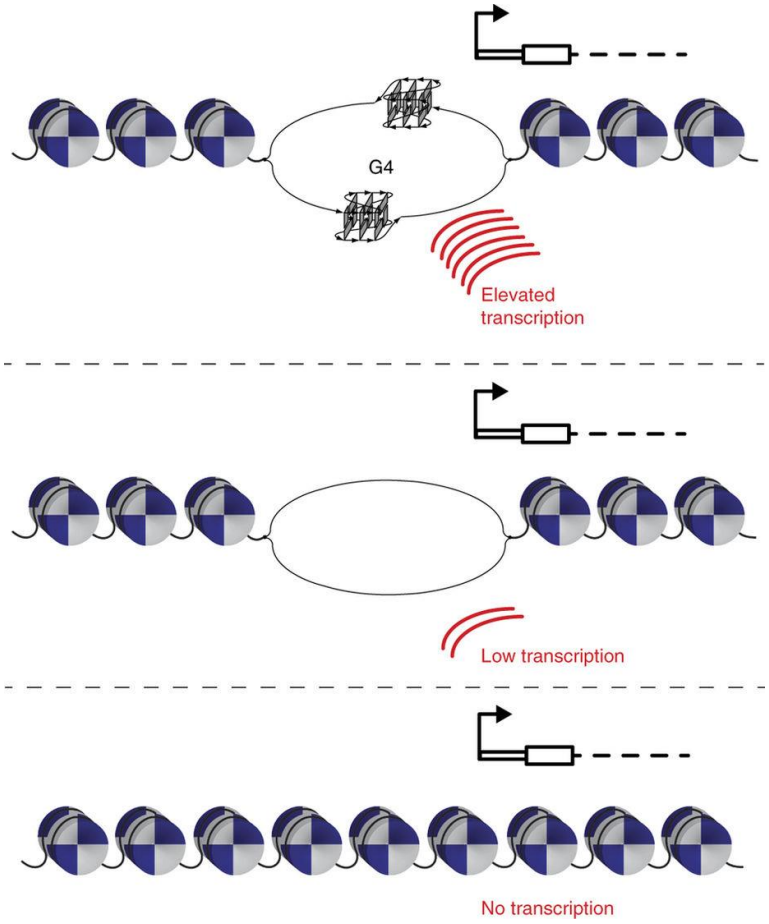
1.1.1.3 G4 and DNA replication

Over the last few decades, G4s have raised intensive interest and research from scientists in different fields because of their presence at chromosome telomeres, oncogenic promoters and 5'-untranslated regions (5'-UTR), which may influence DNA replication, transcription and translation regulation. During DNA replication the double helix are separated in two complementary DNA strand, used as template for the synthesis of the leading and lagging strands, which can have differences in replication timing, in particular if a G4 is formed in the template strand of the lagging strand (Bochman, Paeschke, & Zakian, 2012).

As replicative DNA polymerases are not able to synthesize through G4s, these may cause a replication block (Woodford, Howell, & Usdin, 1994). In order to restore DNA synthesis, G4 structures must be unfolded by a helicase or a different DNA polymerase. G4 unwinding activity can be carried out by several, non-specialized DNA helicases including FANCJ, PIF1, WRN, and BLM, or by DNA polymerases including Polh, Polk, and REV1 (Sauer & Paeschke, 2017). Interestingly, the Doherty's laboratory investigated the capacity of the human PrimPol, a recently described primase-polymerase, to play a role in the bypass of leading strand G4 structures at different ion conditions. Their results showed that PrimPol is required for replicative tolerance of G4 structures in cells (Schiavone et al., 2016).

1.1.1.4 G4 and transcription

In addition to replication interference, G4s may also affect transcription and transcription regulation. In 2007, Huppert and colleagues showed that sequences 1-kb upstream to gene TSSs (Transcription Start Site) are strongly enriched in G4 motifs in comparison to the whole genome. Furthermore, G4-enriched regions are strictly associated with nuclease hypersensitivity (Huppert & Balasubramanian, 2007). Interestingly, in 2014 Gray and colleagues, by ChIP-Seq analyses, showed that 40% of genomic binding sites of transcription-associated helicases, XPB and XPD, overlap with G4 motifs in human cells (Gray, Vallur, Eddy, & Maizels, 2014). Last year, the Balasubramanian's laboratory showed by G4 ChIP-Seq analyses that G4 structures are enriched in promoters and 5′ UTRs of genes with high transcription rate. Furthermore, they showed that *de novo* and enhanced G4 formation are strongly associated with nucleosome-depleted and nuclease-sensitive regions (Figure 10), chromatin markers of active transcription (Hänsel-Hertsch et al., 2016).



(Hänsel-Hertsch et al., 2016)

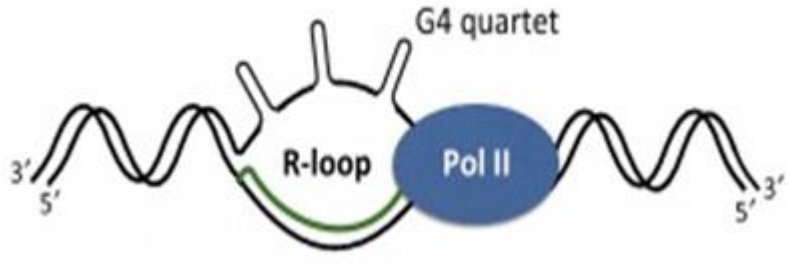
Permission requested

Figure 10. G4 DNA formation is highly correlate on chromatin structure and is frequently found in regulatory, nucleosome-depleted regions in proximity to the transcription start sites of genes.

1.1.1.5 G4 and DNA damage

In 2012, Rodriguez and colleagues shows that PDS treatment induces an accumulation of cells in G2/M phase of the cell cycle with activation of DNA damage response (DDR). In different cancer cell lines, they found phosphorylation of histone H2AX on Ser-139 (γH2AX), activation of the transcriptional repressor KAP1,the checkpoint effector kinase Chk1 and replication protein A (RPA). Furthermore, long time treatments with PDS induced PARP-1 protein cleavage and consequently apoptosis in some cells. Interestingly, they showed that PDS induces DNA damage in G1, S and G2/M cells through transcription-dependent or replication-dependent mechanisms (Rodriguez et al., 2012). Cells pre-treated with 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB), a transcriptional inhibitor, show γH2AX foci in S phase only. Consistently, when cells are pre-treated with aphidilcolin, an inhibitor of DNA replication, in association with DRB, the number and intensity of γH2AX foci are strongly reduced in all cell cycle phases. Overall, these results indicate that PDS triggers double-strand breaks in living cancer cells (DSB) (Rodriguez et al., 2012). An important clinically-relevant finding was then published by Zimmer and colleagues in 2016 (Zimmer et al., 2016). Homologous Recombination (HR)-deficient human cancer cells have increased telomere fragility and accumulate more DSB than HR-proficient cells after G4 binder treatment. In particular, they showed that BRCA2-lacking cells are hypersensitive to PDS treatment. (Zimmer et al., 2016).

Interestingly, a high guanine density in the non-template DNA strand can also favor the formation of R-loop and the DNA/RNA hybrid duplex on the template strand (Figure 11). Therefore, it has been speculated that G4 can favor the formation of R loops at highly transcribed genes (Duquette, Handa, Vincent, Taylor, & Maizels, 2004).

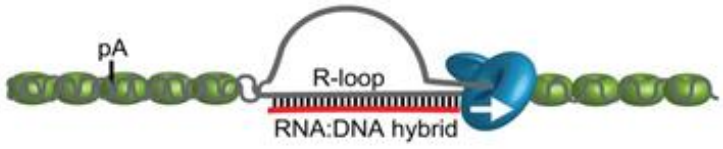


edited by (Sollier & Cimprich, 2015) Permission requested

Figure 11. Physiological roles of R-loops in cells. Speculative scheme of interaction between G-quadruplex and R-loop in which G-Quadruplex are proposed to form on the non-template strand.

1.1.2 R loop structure

R-loops are a triple-stranded structures that contain a DNA/RNA hybrid and a displaced DNA strand (Reaban, Lebowitz, & Griffin, 1994)(Ginno, Lott, Christensen, Korf, & Chédin, 2012). These structures form co-transcriptionally when an RNA molecule emerging from the transcription machinery hybridizes with the DNA template (Figure 12).



edited by (Skourti-Stathaki, Proudfoot, & Gromak, 2011)

Permission requested

Figure 12. Representative scheme of R-loops in cells. In detail, RNA molecule emerging from the transcription machinery hybridizes with the DNA template.

An R-loop structure has been first characterized in vitro over 35 years ago. In *in vitro* studies, a first characterization of R-loops structures (Dson, 1975) (Thomas et al., 1976) was reported by authors investigating transcription activity of RNA polymerases with a defined DNA template. Even in these first papers, it was established that R loops form co-transcriptionally and negative supercoiling tension of the DNA template greatly enhances R loop formation. After 20 years, Drolet and co-workers reports for the first time convincing evidence that R loop structures form in living bacterial cells and that its formation is strongly affected by DNA topoisomerase activity (Drolet et al., 1995).

R loops arise naturally in organisms from bacteria to humans, and they have a multitude of roles in the cell. As reported previously, another important feature that influences R loop formation is the GC content; in particular, R-loop formation can be favoured by a high guanine density in the non-template DNA strand and this is specifically due to the higher thermodynamic stability of RNA-DNA hybrid (Ginno et al., 2012). Indeed, R loop formation in G-clusters have been linked to immunoglobulin class-switch recombination (Yu, Chedin, Hsieh, Wilson, & Lieber, 2003). The presence of G4 structures on the ssDNA induces a higher thermodynamic stability of RNA-DNA hybrid, which is thermodynamically favored as compared with DNA/DNA duplex (Roberts & Crothers, 1992).

1.1.2.1 R loops in the genome

R loops are detectable with selective probes such as the S9.6 antibody, a murine IgG that specifically recognizes DNA/RNA hybrids (Figure 13) (Boguslawski et al., 1986).

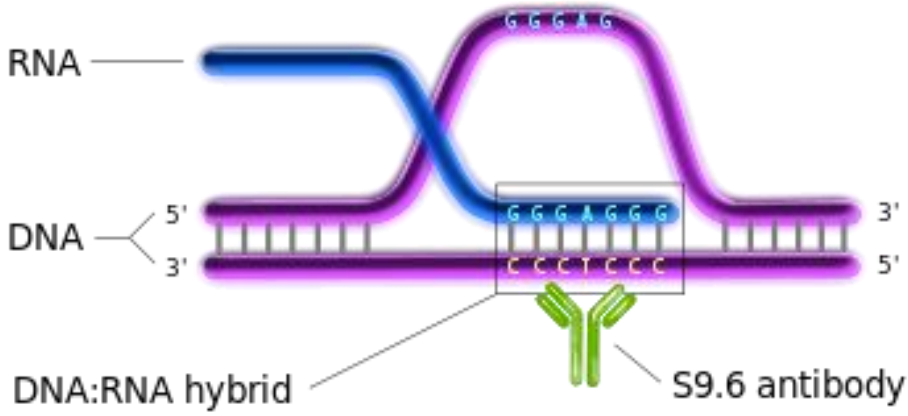


Figure 13. Representative scheme of the interaction between R-loops and S9.6 antibody. In details, the antibody recognizes a hybrid duplex formed by DNA and RNA strands.

The efficiency of S9.6 in detecting R loops in living cells has been established by several results by Chedin's lab. In 2012, Ginno and colleagues (Ginno et al., 2012) developed and compared two different methodologies to measure R loop levels genome-wide:

- DNA:RNA ImmunoPrecipitation (DRIP) technique, which uses the S9.6 antibody to recognize and precipitate R loops from genomic DNA gently isolated from cells;
- DNA:RNA In Vitro Enrichment (DRIVE) technique, which isolates R loops using an inactive human RNaseH1, a mutant enzyme that still binds but cannot degrade hybrid duplexes.

Both methods produce similar genome-wide R-loop profiles, however DRIP identified 20,862 R loop peaks, whereas DRIVE produced 1,224 peaks only. Thus, the DRIP technique appears to be more efficient than DRIVE in R loop detection in cellular genomes. A significant number of peaks detected with the two techniques overlapped with gene promoter regions, from -500 to +1500 bp relative to Transcription Starting Site (TSS). These promoters are strongly correlated with CpG islands and with a pronounced GC skew downstream of the TSS. These findings established that the 5'-ends of active genes are a major site of R loop formation in nuclear chromatin of living cells, and that GC skew is an accurate predictor of R loop formation at promoters of active genes (Ginno et al., 2012) (Ginno, Lim, Lott, Korf, & Chédin, 2013).

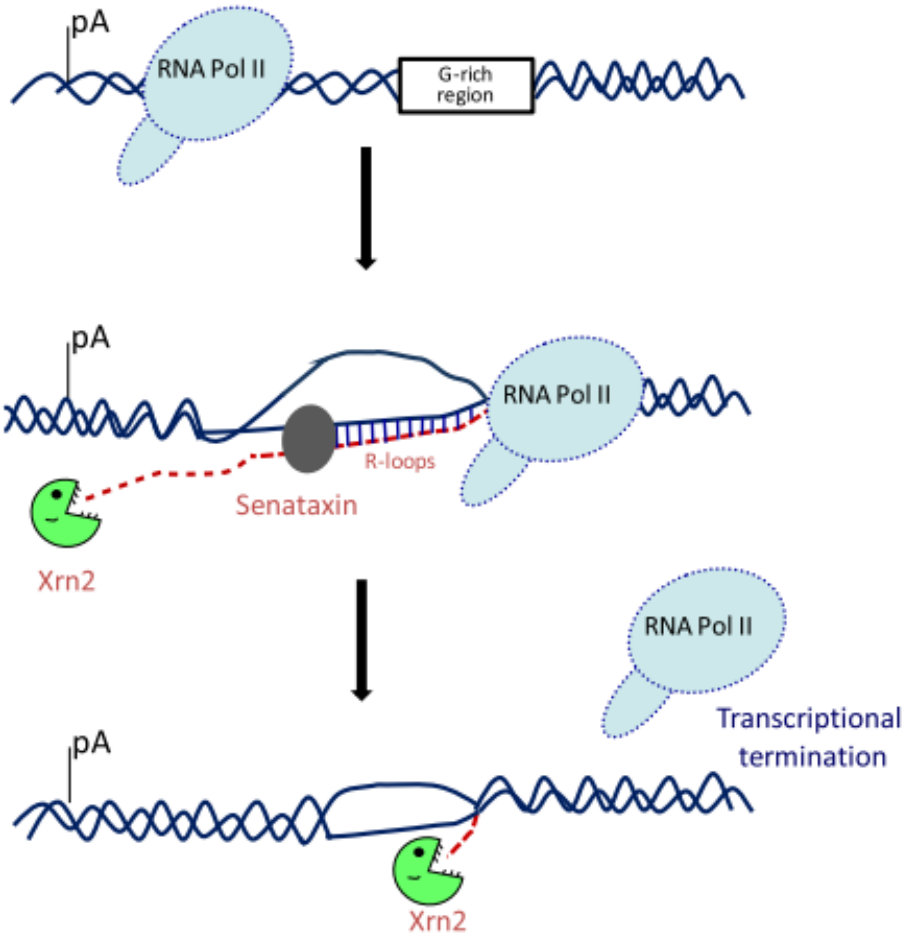
Published works in the last two decades have also shown a cooperative role of R loop structures in DNA replication in different organism. In 1996, Xu and colleagues showed that the RNA-DNA hybrids formed during transcription across the mitochondrial DNA heavy-strand origin, in Saccharomyces cerevisiae cells, provide RNA primers for initiation of mitochondrial DNA replication (Xu1 & Clayton2, 1996). Another example is constituted by the bacteriophage T4 replication mechanisms. Here, the replicative apparatus uses the RNA strand of an RNA-DNA hybrid presents at replication origin to start the synthesis of DNA strands (Mueser, Hinerman, Devos, Boyer, & Williams, 2010). Even though a role of R loops in replication mechanisms is not established in higher eukaryotic cells, however an R loop function in human DNA replication is a matter of intense investigation in several laboratories.

1.1.2.2 R loops and transcription regulation

R loop structures can be involved in the regulation of transcription and gene expression. In 2013, Sun and colleagues showed that R loop formation in the promoter region induced the silencing of the long non-coding COOLAIR RNA gene in *Arabidopsis*. COOLAIR RNA is implicated in the regulation of the flowering locus in response to cold temperature. They conclude that differential stabilization of R loops could be a general mechanism influencing gene expression in many organisms (Sun, Csorba, Skourti-Stathaki, Proudfoot, & Dean, 2013).

As discussed above GC skew is an accurate predictor of R loop formation at promoter region (Ginno et al., 2013). In addition, Chedin's group showed evidence that R loop regions are highly un-methylated and has suggested that R loop can participate in the mechanism of maintenance of un-methylated CpG islands at promoter regions (Ginno et al., 2012).

In Proudfoot's laboratory, they showed a role of R loop formation in transcriptional pausing of RNA polymerases at G-rich stretches closed to poly(A) signals for transcriptional termination (Figure 14) (Skourti-Stathaki et al., 2011). In addition, the same laboratory has shown that R loop formation can induce chromatin remodeling at promoters, gene body and terminal gene regions, and for the latter, chromatin remodeling was specifically mediated by the efficient transcription termination due to R loops (Skourti-Stathaki, Kamieniarz-Gdula, & Proudfoot, 2014).



(Skourti-Stathaki et al., 2011)

Image under CC BY permission use

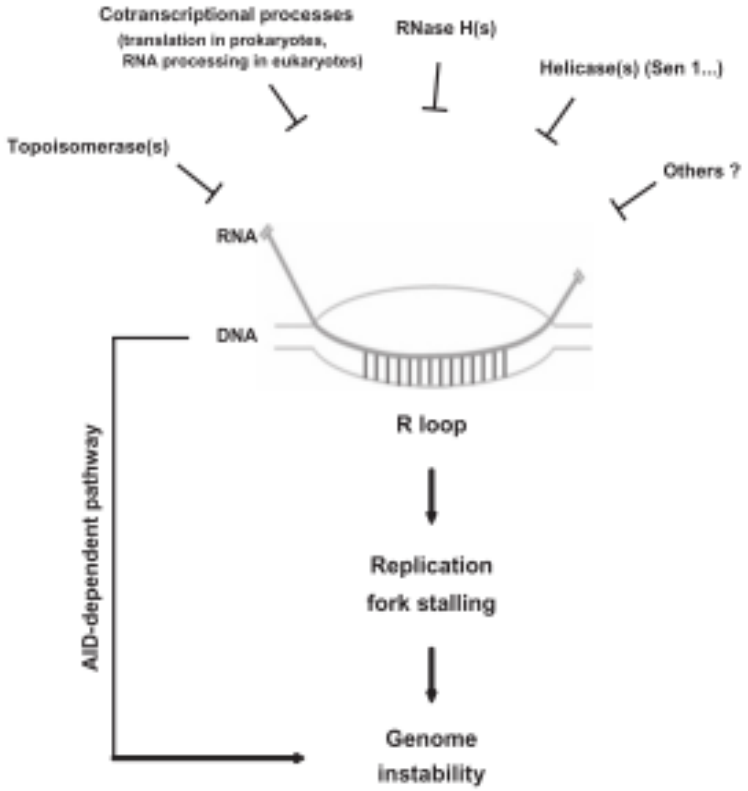
Figure 14. Role of Senataxin and R-Loops in Transcriptional Termination. R-loops formed over the G-rich pause region of human b-actin gene are necessary for Pol II to pause downstream of the poly(A) site. Senataxin is needed to resolve R-loop structures and so allow 50–30 exonuclease Xrn2 to degrade the nascent RNA from the site of poly(A) cleavage and catch up with paused Pol II, causing itstranscriptional termination.

1.1.2.3 R loops and DNA damage

Unscheduled R-loops are generally regarded as highly deleterious as they can readily lead to mutations and genome instability. For instance, the displaced single-stranded DNA is more prone than double stranded DNA to different mutagenesis events due to either exogenous chemical agents or endogenous enzymes, such as AID (Chaudhuri & Alt, 2004) (Petersen-Mahrt, Harris, & Neuberger, 2002) (Gomez-Gonzalez & Aguilera, 2007) (Muers, 2011).

Interestingly, Beletskii and colleagues showed that R loops increase base mutations of the ssDNA much more than of the DNA strand annealed to RNA (Beletskii & Bhagwat, 1996). Experimental evidence for R loops as a source of uncontrolled genome instability was initially provided with the THO mutants in yeast cells (Huertas & Aguilera, 2003).

Yeast THO and THSC/TREX-2 R loop-forming mutants show a transcription-associated hyper-recombination phenotype and elevated chromosome and plasmid losses (Aguilera & Gómez-González, 2008). Interestingly, R loop structures may constitute a special impediment to the progression of the replication fork and this causes a marked increase of genome instability (Figure 15). In 2011, Gan and colleagues showed in two different genomic systems, *E. coli* and human cancer cells (HeLa cells), that the genome instability induced by R loops depends on the DNA replication status. Indeed, at active DNA replication regions R loop structures induced a strong impediment to the progression of replication forks, thus causing an increase of DNA damage and genome instability (Gan et al., 2011).

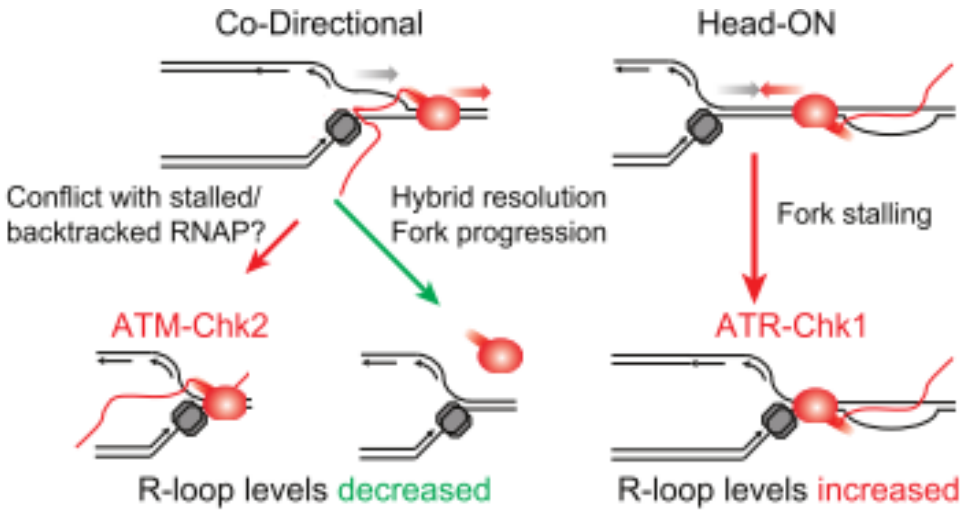


(Gan et al., 2011)

Image under CC BY permission use

Figure 15. Multiple pathways involved in the transcription induced R-loop resolution and that determine genomic instability. One of the most efficient pathway involved in the R-loop resolution provides the RNaseH1 action.

A very recent work has further supported that R loops are responsible for replication fork stalling when it collides with the transcription elongation apparatus. Interestingly, Hamperl and colleagues also showed that there is a reciprocal modulation between R loops and DNA replication, and this is orientation dependent relative to transcription. Indeed, the replisome machinery reduces R loops in the co-directional orientation, whereas it promotes R loop stabilization in head-on orientation (Figure 16) (Hamperl, Bocek, Saldivar, Swigut, & Cimprich, 2017).



(Hamperl et al., 2017)

Image under CC BY permission use

Figure 16. Model for how head-on and co-directional transcription-replication conflicts regulate R-loop homeostasis and induce distinct DNA damage responses in human cells.

R loops have been associated to several markers of double-stranded DNA cleavage, such as phosphorylated H2AX histone (Skourti-Stathaki & Proudfoot, 2014) and DNA damage checkpoint activation, by an extensive literature (Skourti-Stathaki & Proudfoot, 2014) (Sollier & Cimprich, 2015). Therefore, R loops can lead to DNA cleavage that will eventually be repaired by error-prone mechanisms leading to chromosomal rearrangements and genome instability. Although the mechanism responsible for the generation of DSBs from R loops is still unknown, it has been suggested that R loops can be processed into DNA strand breaks by nucleotide excision repair (NER) endonucleases XPF and/or XPG (Sollier et al., 2014). It must be pointed out that other structure-specific endonucleases, such as Mus81/Eme1, can also be involved, and the mechanisms of DNA cleavage generation remains to be fully established.

1.1.2.4 R loops resolving

The organisms have developed many mechanisms for “hybrid dissolution” and reduce R loop in the genome. RNase H1 and RNase H2 are endonuclease enzymes conserved from bacteria to human that specifically cut the RNA in a DNA:RNA hybrid. RNaseH1 is monomeric, whereas RNaseH2 is a trimeric enzyme. Both enzymes present two different domains:

- a hybrid-binding domain (HBD) that binds the RNA:DNA substrate
- an RNase H domain that is involved in the catalysis of RNA hydrolytic cleavage

In 2011, Wahba and colleagues showed that inactivation of both RNasesHs in yeast induced an increase of R loops (Wahba, Amon, Koshland, & Vuica-ross, 2011) demonstrating their involvement in maintaining the R loop balance in chromatin. In some cases, overexpression of RNaseH1 is sufficient to reduce R loop levels in cultured cells and is often used to investigate R loop role in molecular mechanisms. For instance, R loop increases induced by defects in RNA biogenesis factors in bacteria, yeast and mammalian cells can be suppressed by overexpression of RNaseH1 (Domínguez-Sánchez, Barroso, Gómez-González, Luna, & Aguilera, 2011).

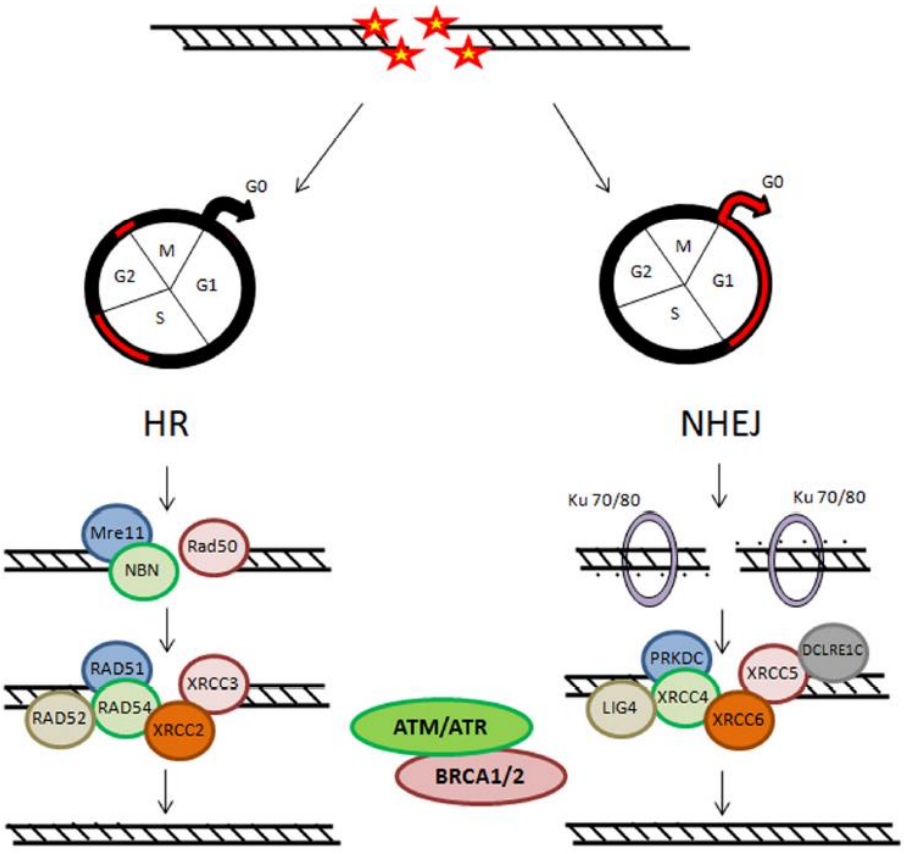
In addition to RNaseH1/2 enzymes other factors affect R loop levels in cells, including helicases and DNA topoisomerases. Negative torsional stress favors R loop formation, and this is conserved from prokaryotes to mammalian systems. In E. coli, distruption of Type IA topoisomerase leads to growth defects that are rescued by overexpression of RNase H, indicating that the increase in negative supercoils induces R loops accumulation (Phoenix, Raymond, Massé, & Drolet, 1997) (Massé, Phoenix, & Drolet, 1997) (Baaklini, Hraiky, Rallu, Tse-Dinh, & Drolet, 2004). In yeast, loss of type IB Topoisomerase (Top1) is associated with R loop accumulation in the very highly transcribed ribosomal RNA gene clusters (El Hage, French, Beyer, & Tollervey, 2010) (French et al., 2011). The function of Top1 in preventing R loop accumulation is conserved in mammalian systems. Top1-deficient mouse lymphocytes accumulate stalled replication forks and DNA breaks at actively transcribed regions (Tuduri et al., 2009). Both replication-fork stalling and DNA breaks are reduced by overexpression of RNase H1, implicating transcription.

Both replication-fork stalling and DNA breaks are reduced by overexpression of RNase H1. Prokaryotic and eukaryotic cells presents multiple RNA-DNA helicases that are able to unwind the hybrid duplexes thus preventing genome instability. Such helicases are Rho, DHX9 and Senataxin (Chakraborty & Grosse, 2011) (Mischo et al., 2011).

1.2 DNA Damage Response (DDR)

The study of model organisms has substantially contributed to our understanding of DNA repair mechanisms, particularly DNA double strand break (DSB) repair. The DSB presence generates a complex cascade of events aimed to blocking the cell cycle and recruiting repair factors (Jeggo, n.d.). One of the first protein involved in this process is the product of the ATM gene, a kinase that through p53 stops the cell G1 phase and causes phosphorylation of the H2AX in the chromatin domain near the DSB. Then, the repair can take place in two different ways, the Homologous Recombination (HR) pathway or the Non-Homologous End Joining (NHEJ).

These two mechanisms compete with each other, but it has been demonstrated that in S and G2 phases, HR is active, whereas it is inhibited in G1 phase. The NHEJ pathway is instead operative in all cell cycle phases (Figure 17).



(Cirincione et al., 2017)

Image under CC BY permission use

Figure 17. Schematic representation of the two mechanistically distinct pathways involved in the mitotic repair of DSBs: homologous recombination (HR) and non-homologous end joining (NHEJ). At right is shows the HR mechanism, the error free mechanism most active in S/G2 phase. In this pathway is involved the MRE11, NBN and RAD50 protein and the RAD51, RAD52, RAD54 and XRCC2, XRCC3 that reconstitute the double strand helix. At left is shown the NHEJ mechanism that seems to be mostly active in G1 phase of the cell cycle. The proteins involved in the first instance is Ku70 and Ku80, after that there is the recruitment of PRKDC, LIG4, XRCC4, XRCC5, XRCC6 and DCLRE1C that reform the double strand helix.

1.2.1 Homologous Recombination (HR)

In the HR, DNA ends produced by DSB are processed by the RAD50-MRE11-NBS1 complex. This complex through the MRE11 nuclease activity leads to the production of protruding 3' ends (Symington & Gautier, 2011). Other specific esonucleases are then recruited to resect DNA end more extensively producing a long 3'-end protruding strand. RPA proteins polymerize along the single-strand DNA forming specific nucleoprotein structures. This nucleoprotein complex identify the homologous regions creating a supercomplex with the DNA. At this level, there is the recruitment of RAD51 protein that catalyzes the exchange between the two DNA strand (Pellegrini et al., 2002). The exchange produces the matching between the homologous sequence in the other strand and the DNA break. The 3'-ends of the cut strand annealed to the complementary strand of the homologous chromosome is then used to start DNA synthesis and fill the gap. The neo-synthesized stretch may be carried out by the homologous propeller and return to fit with the filament that contained the other rupture generated by the DSB (Figure 18) (Kakarougkas & Jeggo, 2014) (Zimmermann & De Lange, 2014). Thus, HR is a repair pathway that does not introduce errors or deletion/insertion in the broken DNA sequence as it uses the genetic information contained in the sister chromatid.

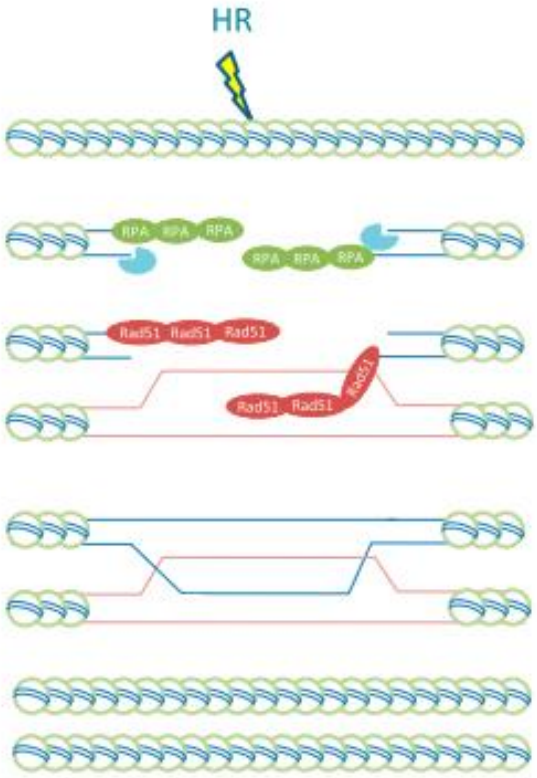
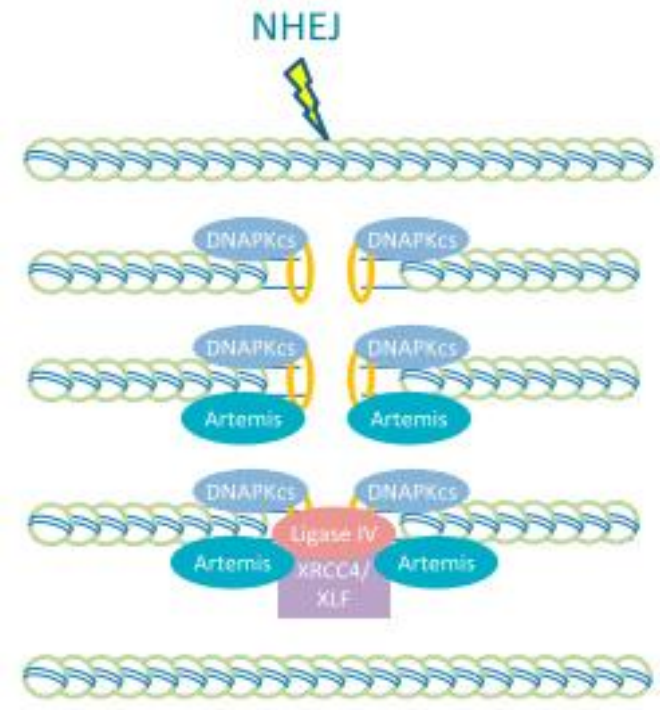


Figure 18. Schematic representation of the homologous recombination (HR) pathways involved to repair mitotic DSBs. In the first instance RPA recognized the double strand break and recruit RAD51. RAD51 acts by allowing the filament reconstitution and by DNA ligases reforming the double helix of the DNA.

1.2.2 Non-Homologous End Joining (NHEJ)

NHEJ is a compact process that most likely requires small changes to the chromatin at the double strand break (DSB) proximity. DSBs are recognized by the Ku80-Ku70 heterodimer, that then recruits DNA-PKcs generating DNA-PK complex. The DNA-PK complex activates the activity of DNA-PK. This kinase regulates end-processing and NHEJ through auto-phosphorylation and also facilitates recruitment of a ligation complex (Cottarel et al., 2013). The ligation complex includes the DNA ligase IV (LigIV), the X-ray cross complementing group 4 (XRCC4) and XRCC4-like factor. Additional proteins also contribute to end processing, including polynucleotide kinase 39 phosphatase (Bernstein et al., 2009). The structure-specific nuclease, Artemis, is also required for re-joining a subset of DNA ends (Figure 19) (Riballo et al., 2004).



(Kakarougkas & Jeggo, 2014)

Image under CC BY permission use

Figure 20. Schematic representation of the non-homologous end joining (NHEJ). pathways involved to repair mitotic DSBs. In the first instance DNApK recognized the double strand break and recruit Artemis. The complex DNApK-Artemis induce recruitment of XRCC4, XLF and Ligase IV that have a synergistic activity in reforming the DNA double helix.

However, the connection between these two structures is highly acclaimed and already confirmed *in silico* but has not yet been confirmed in human cells. The aim of my project was the definition of the relationship between G4s and R loops in living human cancer cells and the effects on DNA damage and genome stability.

1.3 AIMS

The aim of the present PhD Thesis is the definition of the molecular mechanism of G4 binder induction of DNA damage and genome instability. In particular, the experimental work has been focused on the involvement of R loops in mediating the DNA cleavage production by G4 binders in human cancer cells.

For my work, we selected commercial molecules known for their G4 binding activity (PDS and Braco-19) and compounds synthesized by Professor M. Rambaldi's group at the University of Bologna. The latter compounds were selected for their activity (FG, Compound 3) or inactivity (FA) as specific G-quadruplex binders relative to ssDNA or dsDNA. The most suitable cell lines were selected for molecular and cellular investigations based on published works (Balasubramian). Preliminary analyzes carried out in the lab confirmed the efficiency of immunofluorescence microscopy images with the used antibodies. In detail, the BG4 antibody (specific for G4 structures) showed a clear dotted pattern in agreement with published papers. Similarly, clear results were obtained with the antibody against DNA:RNA hybrids (S9.6). Preliminary analyses with MTT assay allowed the determination of optimal concentrations of compounds to be analyzed. Finally, the immunoprecipitation analysis (DRIP) with Ab S9.6 allowed to investigate R loops in the mechanism of action of G4 binders with a different methodological approach.

2. MATERIALS AND METHODS

2.1 BG4 plasmid

BG4 plasmid containing the sequence specific for G4 structure was given as a gift from Professor Balasubramanian, University of Cambridge (Biffi, Di Antonio, Tannahill, & Balasubramanian, 2013). BG4 was then transfected in E.coli bacterial expression system to express antibody specific for the detection of G4 (Figure 21).

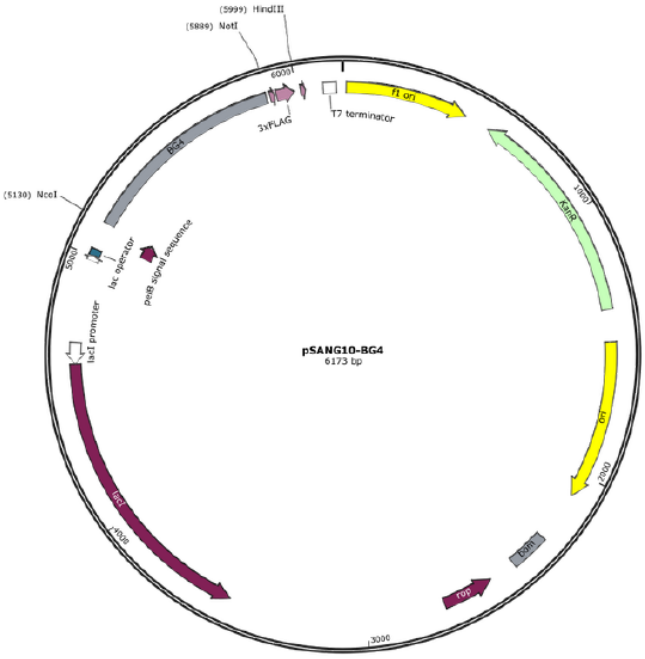


Figure 21. pSANG10-BG4. Plasmid containing BG4 sequence.

2.1.2 Transformation of BG4 plasmid

The strain is a specially modified BL21 that will express genes from the T7 promoter. Competent cells kept in - 80 °C was placed on ice for 5 mins or until just thawed. A DNA concentration of 1-50 ng (in a volume not greater than 5µl) was added to the competent cells by moving the pipette tip through the cells while dispensing and quickly flicked the tube several times without vortexing. The tubes were immediately returned to ice for 5–30 mins. Cells were heat-shocked for 15 seconds in a water bath at exactly 42 °C without shaking. The tubes were immediately placed on ice for 2 mins after which 450 µl of room temperature SOC medium was added to each transformation reaction, and incubated for 60 mins at 37 ˚C with shaking (approximately 225 rpm). The tubes were laid on their sides and taped to the platform for best transformation efficiency. 100µl of undiluted cells and 1:10 and 1:100 cell dilutions were plated on antibiotic plates and incubated overnight at 37°C.

2.1.2 Preparation of BG4 antibody

Before start:

- TES buffer : 50 mM Tris-HCL pH 8.0

  1 mM EDTA pH 8.0

  20 % sucrose

- Auto-Induction Medium, supplement basic ZY medium as follow :
  - 2 mM MgSO4
  - 0.2X Metals Mix
  - 1X 5052
  - 1X M
  - 50 µg/ml Kanamycin

- 1000X metals mix (stock solution) – add in order written
  - ddH_2O to volume
  - 50 mM FeCl3 * 6H_2O (dissolved in 0.1M HCL)
  - 20 mM $CaCl_2$
  - 10 mM $MnCl_2$ * 4H_2O
  - 10 mM $ZnSO_4$ * 7H_2O
  - 2 mM $CoCl_2$ * 6H_2O
  - 2 mM $CuCl_2$ * 2H_2O
  - 2 mM NA_2MoO_4 * 2H_2O
  - 2 mM NA_2SeO_3 * 5H_2O
  - 2 mM H_3BO_3

- 50X 5052
    - 25% (w/v) Glycerol
    - 2.5 % (w/v) Glucose
    - 10% (w/v) α-Lactose

- 50X M
    - 1.25 mM KH_2PO_4
    - 2.5 mM NH_4Cl
    - 0.25 mM Na_2SO_4

The induction of BG4 protein expression in E.coli was performed using the auto-induction method as described by Studier (ref 2005), which requires a high bacterial density. For initiation culture, transformed-BL21 cells were inoculated in 2 ml of 2x TY media + 2 % glucose + 50 mg/ml kanamycin. Cells were grown overnight at 200 rpm at 30 °C. The initiation culture was inoculated in 100 ml auto induction media and let to grow at 37 °C at 250 rpm for 6 hours. The cell culture was pelleted for 30 min at 4 °C at 4000 g. The pellet was resuspended in 8ml TES and left on ice for 10 mins. A further 12 ml TES diluted 1:5 was added into mix and left for 15 mins on ice. Cells debris was discarded by spinning down for 10 mins at 8000 g at 4 °C. At this point, the solution contained a protein mixture of endogenous protein and BG4 antibody. BG4 was purified by using silica-based resin (Protino® Ni-IDA) pre-charged with Ni2+ ions. Since BG4 is tagged with polyhistidine, the protein would be bound by the immobilized Ni2+ on the resins. One gram of resin was weighed and packed into a column. The resin was then equilibrated in PBS, pH 8.0. The protein solution was added to the pre-equilibrated column and was allowed to drain by gravity. To ensure higher BG4 binding, the flowthrough was collected and was re-added to the column. The column was washed twice with cold PBS/100 mM NaCl/10 mM imidazole, pH 8.0 and drained by gravity. Elution was done in a new collecting tube by adding PBS/250 mM imidazole, pH 8.0. The eluted antibody solution was dialysed overnight in PBS, pH 8.0 in cold room. The BG4 antibody is stored at 4 °C for few weeks.

2.2 Cell lines

U2OS, and HeLa were grown in DMEM, WI38 in MEM supplemented with 1mM Sodium Piruvate, 1500 mg/L of Sodium Bicarbonate and non essential aminoacids. Both growth mediums were supplemented with 10% heat-inactivated foetal bovine serum and 2 mM glutamine. The cells were maintained by trypsinization and passed 2-3 times a week and incubated at 37 °C, 5 % CO2.

2.3 Drug preparation

FG and Braco-19 (B-19) (Sigma) was initially dissolved in 100% DMSO to make up 10 mM stock solutions while FA in 2 mM stock solutions (FA is less soluble than FG and B-19). Pyridostatin (Sigma) was dissolved in 100% ddH_2O at 5 mM stock solution. The stock solutions were aliquoted in small vials of 30 μL to avoid repeating freeze-thaw. Stock solutions were diluted at the work concentration in growth medium.

2.4 Immunofluorescence

Cell seeding was performed in a different manner for the different culture plate. For U2OS cells:

- 150000/200000 cells in a 35 mm dish
- 50000 cells in a 4-well Nunc™ Lab-Tek™ II Chamber Slide™ System (Nalge Nunc International)
- 2000 cells in a 96 well plate for operetta

For HeLa cells: 150000/200000 cells in a 35 mm dish.

For WI38 cells: 2000 cells in a 96 plate for operetta.

24 hours after seeding cells were treated for a different time/concentration of PDS, FG, FA, B-19.

2.4.1 BG4 Immunofluorescence

Before start:

- Fix solution : Methanol/Acetic Acid 3:1 solution
- Permeabilization solution : Triton 0,1% in PBS 1X
- Blocking buffer : 2% Milk in PBS 1X
- Wash solution : Tween 0.1% in PBS 1X

24h post seeding + Treatment, cells were pre-fixed with DMEM + Fix solution 50/50 5 min at room temperature (RT). After a brief wash with Fix solution, cells were fixed with Fix solution at RT for 10 min. Cells were permeabilized with a Permeabilization solution 3 min at RT. 1h after incubation with Blocking buffer at RT, cells were incubate 2h at RT with 2µg per slide of BG4 antibody in Blocking buffer. Cells were then incubated 1h at RT with Anti-Flag antibody (1:800 dilution) in Blocking buffer. Next , cells were incubated 1h at RT with Anti-Rabbit antibody (1:1000 dilution) in Blocking buffer. For nuclear staining, cells were incubated with 2µg/µl of DAPI. The cover glasses were mounted with Mowiol-488.

2.4.2 S9.6 Immunofluorescence

Before start:

- Blocking buffer: 3% BSA, 0.1% Tween-20 in SSC 4X
- Wash solution: SSC 4X

24h post seeding + Treatment, cells were fixed with freezer-cold Methanol at RT for 10 min. Cells were permeabilized with Acetone for 3 min at RT. After 30 mins of Blocking buffer at RT, cells were incubated 2h at RT with 2µg per slide of S9.6 antibody in Blocking buffer. Cells were then incubated 1h at RT with Anti-Mouse secondary antibody (1:1000 diluition) in Blocking buffer. For nuclear staining, cells were incubated with 2µg/uL of DAPI. The cover glasses were mounted with Mowiol-488.

2.4.3 γH2AX Immunofluorescence

Before start:

- Fix solution: 3.7% PFA solution 3.7% in ddH_2O
- Permeabilization solution: Triton 0.5% in PBS 1X
- Blocking buffer: 8% BSA in PBS 1X
- Antibodies buffer: 1% BSA in PBS 1X
- Wash solution: PBS 1X

24h post seeding + Treatment, cells were fixed with PFA solution at RT for 15 min. Cells were permeabilized with the Permeabilization solution for 15 min at RT. After 1h of Blocking buffer at

RT, cells were incubated 2h at RT with 1:500 of γH2AX antibody in Antibodies buffer. Cells were then incubated 1h at RT with Anti-Mouse antibody (1:500 dilution) in Antibodies buffer. For nuclear staining cells were incubated with 2µg/uL of DAPI. The cover glasses were mounted with Mowiol-488.

2.4.4 Rad51 Immunofluorescence

Before start:

- Permeabilization solution: Triton 0.5% in PBS 1X
- Blocking buffer : 10% BSA, 0.3% Triton in PBS 1X
- Antibodies buffer : 1% BSA, 0.3% Triton in PBS 1X
- Wash solution : PBS 1X

24h post seeding + Treatment, cells were fixed with freezer-cold methanol on ice for 15 min. Cells were permeabilized with a Permeabilization solution for 5 min at RT. After 1h of Blocking buffer at RT, cells were incubated ON at 4°C with 1:500 of RAD51 antibody in Antibodies buffer. Cells were then Incubate 1h at RT with Anti-Mouse antibody (1:500 dilution) in Antibodies buffer. For nuclear staining, cells were incubated with 2µg/uL of DAPI. The cover glasses were mounted with Mowiol-488.

2.4.5 53BP1, p53BP1, pATM Immunofluorescence

Before start:

- CSK buffer : 10 mM PIPES pH 6.8, 100 mM NaCl, 300 mM Sucrose, 3 mM MgCl2, 0.5% Triton X-100 and protease inhibitors
- Fix solution : PFA solution at 2% in ddH_2O
- Antibodies buffer : 5% BSA in PBS 1X
- Wash solution : PBS 1X

24h post seeding + Treatment, CSK buffer was added dropwise and cells were incubated 3 min at RT. Cells were then fixed with Permeabilization solution for 15 min at RT. Cells were incubated 1h at RT with 1:500 of p53BP1/53BP1/pATM in Antibodies buffer. Cells were then incubated 1h at RT

with Anti-Mouse/Anti-Rabbit antibodies (1 : 500) in Antibodies buffer. For nuclear staining cells were incubate with 2µg/uL of DAPI. The cover glasses were mounted with Mowiol-488.

2.4.6 BG4/S9.6 Immunofluorescence

Before start:

- Permeabilization solution : Triton 0.5% in PBS 1X
- Blocking buffer : 8% BSA in PBS 1X
- Antibodies buffer : 1% BSA in PBS 1X
- Wash solution : PBS 1X

24h post seeding + Treatment, cells were fixed with freezer-cold methanol at RT for 10 mins.
Cells were then permeabilized with Permeabilization solution for 15 min at RT. After 1h of Blocking buffer at RT, cells were Incubate 1h at RT with 2µg per slide of BG4 and 2µg per slide of S9.6 in Antibodies buffer. Cells were then incubated 1h at RT with Anti-Flag (1 : 800 dilution) in Blocking buffer. Finally, cells were incubated 1h at RT with Anti-Mouse and Anti-Rabbit secondary antibody (1 : 500 dilution) in Antibodies buffer. For nuclear staining, cells were incubated with 2µg/uL of DAPI. The cover glasses were mounted with Mowiol-488.

2.4.7 S9.6 - γH2AX Immunofluorescence

Before start:

- Pre-extraction solution: NP-40 0.5% in PBS 1X
- Fix solution : 3.7% PFA in PBS 1X
- Blocking buffer : 2% BSA in PBS 1X
- Wash solution : PBS 1X

24h post seeding + Treatment, cells were treated with Pre-extraction solution on ice for 3 mins. Cells were then fixed with Fix solution for 15 min at RT. After 1h of Blocking buffer at RT, cells were incubated 1h at RT with 2µg per slide of S9.6 and 1 : 500 of γH2AX antibody in Blocking buffer. Cells were then incubated 1h at RT with Anti-Mouse and Anti-Rabbit secondary antbodies (1 : 500 dilution) in Antibodies buffer. To stain nuclear DNA, cells were incubated with 2µg/uL of DAPI. The cover glasses were mounted with Mowiol-488.

2.5 DNA : RNA Immunoprecipitation (DRIP)

The DNA:RNA immunoprecipitation technique allows to immunoprecipitate R-loop structures at genome wide level by using the S9.6 as specific antibody for the detection and the binding of these hybrids in living cells. (Ginno et al., 2013).

2.5.1 DNA Extraction and Restriction Digestion for DRIP

Before start:

- Lysis Buffer: 30 mM Tris-HCL pH 8.0, 1 mM EDTA, 1% SDS
- 1X PBS
- Tris-EDTA (TE) pH 8.0: 10 mM Tris-HCl, 1 mM EDTA
- Proteinase K

Cells seeding was performed in a 10 cm plate at a concentration of 100000 cells/ml. 24h post seeding + Treatment, cells were resuspended in Lysis Buffer, scraped and collected in a 15ml tube. 100 µg of Proteinase K were added to the tubes and the samples were incubated overnight at 37°C.

The day after, DNA was precipitated adding 2.5 volume of 100% Ethanol, 1/10 volume of 3M NaOAc and 1 ul of Glycogen (the latter is optional). After gently inversion, the DNA was completely precipitated. The DNA was washed 5 times with 70% EtOH and then allowed to air dry. Pellet was resuspended in TE and incubated in ice for 1h or moreGenomic DNA was then digested ON at 37°C using a restriction enzyme cocktail (80U HindIII, 40U XbaI, 40U BsrgI, 40U SSpI) in Tango Buffer 1X. The day after, cells were incubated again with 80U EcoR1 in 2X Tango Buffer.

DNA quantification was performed with a Thermo Scientific™ NanoDrop™ spectrophotometers.

2.5.2 DRIP immunoprecipitation

Before start:

- Tris-EDTA (TE): 10 mM Tris-HCl, 1 mM EDTA
- Binding Buffer: 100 mM $NaPO_4$ pH 7.0, 1.4 mM NaCl, 0.5% Triton
- Elution Buffer: 50 mM Tris-HCl, 10 mM EDTA, 0.5% SDS

4 µg of collected gDNA and 10% of the equivalent input were resuspended in TE buffer. In order to have a control for a quantitative PCR (qPCR) the input was removed and stored at 4°C. 10 µg of S9.6 Antibody and 1X Binding Buffer were added to the samples and incubated ON at 4°C under rotation. The day after, the complex DNA/S9.6 was incubated with a mixtures of Protein G Sepharose® 4 Fast Flow (GE Healthcare ref #17-0618-01) and Protein A-Sepharose® CL-4B (GE Healthcare ref #17-0780-01) beads, that were previously washed and equilibrated with Binding Buffer 1X. After 2h in rotation at 4°C, the DNA/S9.6/beads complex was washed and then resuspended in Elution Buffer (50 mM Tris-HCl, 10 mM EDTA, 0.5% SDS). Samples were incubated with 100 µg of Proteinase K at 55°C for 45 mins or more. The supernatant was collected in a new tube and DNA was extracted using phenol-chloroform extraction protocol followed by DNA precipitation ON at -20°C. The last day the DNA was precipitated, washed with 70% Ethanol and then resuspend in TE buffer.

2.6 Western Blotting

Before performing Western Blotting analysis, cells were lysed and protein extracted following the protocol described below.

2.6.1 Protein extraction

<u>WCE Protocol</u>

Boiling Buffer

- 1% SDS
- 10 mM Tris pH 7.4
- 1X Protease/Phosphatase inhibitors

Whole-cell lysates were prepared by lysing the cells in boiling buffer in the presence of protease and phosphatase inhibitors. The viscosity of the samples was reduced by brief sonication cycles.

<u>Cellular extracts for ATM and P-ATM</u>

Lysis Buffer

- 50 mM Tris-HCl pH 8.0
- 300 mM NaCl
- 0.4% NP-40

- 10 mM MgCl2
- 5 mM DTT
- 20 mM β-Gly
- 1X Protease/Phosphatase inhibitors

Dilution Buffer

- 50 mM Tris-HCl pH 8.0
- 0.4% NP-40
- 5 mM DTT

Cells were lysed with cold Lysis Buffer for 15 min. After centrifugation the supernatant was transferred in another tube containing dilution buffer.

Hystones Extraction

Lysis Buffer

- 10 mM Tris-HCl pH 7.5
- 0.5% NP-40
- 1 mM MgCl2
- 1X Protease/Phosphatase inhibitors

Extraction Solution

- 200 mM H_2SO_4 MgCl2
- 20% Glycerol
- 5% MercaptoEthanol

Cells were lysed with cold Lysis Buffer for 15 min. After centrifugation pellet was resuspended in Lysis Buffer + 400 mM NaCl and then pelleted. Pellet was resuspended in Extraction Solution and incubated for 2 hours at 4°C. After centrifugation the supernatant was resuspended in 20% Trichloroacetic acid and centrifugated. Pellet was resuspended in 100% Ethanol and after in 10% Glycerol solution.

2.6.2 SDS-PAGE

Polyacrylamide gel electrophoresis (PAGE), describes a technique widely used in biochemistry, molecular biology and biotechnology to separate biological macromolecules, usually proteins or nucleic acids, according to their electrophoretic mobility. Mobility is a function of the length, conformation and charge of the molecule. As with all forms of gel electrophoresis, molecules may be run in their native state, preserving the molecules' higher-order structure. SDS-PAGE (Sodium dodecyl sulfate polyacrylamide gel electrophoresis), is a method of separating molecules based only on the difference of their molecular weight.

Before start:

- Tris-HCL 1.5 M pH 8.8
- Tris-HCL 0.5 M pH 6.8
- Ammonium persulfate (APS) 10 %
- SDS 10 %
- Laemnli Buffer 4X : Tris-HCL 0.2 M , SDS 8% , Bromophenolblu 0.4% , Glycerol 40 %
- Running Buffer : Trizma-Base 0.3% , Glycine 1.5% , SDS 0.1%

The first step is the gel preparation, the gel is constituted by a resolving and stacking gel.

Resolving gel:

- 405 mM Tris-HCL pH 8.8
- 0.1 % SDS
- 0.1 % APS
- 0.1 % TEMED
- Achrilammide - different concentration for a different protein molecular weight

Stacking gel:

- 40% Acrilammide
- 175 mM Tris-HCl pH 6.8
- 0.2 % SDS
- 0.1 % APS
- 0.1% TEMED

The second step consists of sample preparation: x µg of each sample was mixed with 1XLaemnli Buffer and 2.5% Mercaptoethanol and incubated 10 mins at 95°C. After run, gel was used for wet transfer to the nitrocellulose membrane

2.6.3 Immunoblotting

Before start:

- Transfer Buffer : 25 mM Tris-HCl pH 7.5 , 192 mM Glycine , 20 % Methanol
- TBS 1X : 50 mM Tris-HCl pH 7.5 , 150 mM NaCl
- Blocking buffer: BSA / Milk in TBS 1X - Tween 0.1%

After wet transfer, the nitrocellulose membrane was stained with Red Ponceau to analyse protein concentration among the different samples. Next, membrane was used for incubation with primary and secondary antibodies.

2.7 MTT assay

The MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-2H-tetrazolium bromide) colorimetric assay is commonly used to determine mitochondrial reductive function and hence it is useful as indicator of cell death or inhibition of growth. MTT was performed on cells treated with increasing doses of drug and incubated for 72 hrs after drug exposure for recovery. This last step was performed to provide a more accurate indication of the growth inhibition caused by drug cytotoxicity rather than other situations, such as the cell quiescence, metabolic stopping, or induction of apoptosis. All assays were performed in duplicate. 50 µl of MTT solution (5 mg/ml in PBS, Sigma) was added to each well and incubated for 2 hrs. Medium was subsequently removed from wells and resulting formazan crystals solubilised in 300 µl of DMSO. Culture plates were gently rocked for 1 hour to solubilise the formazan crystals before measuring the optical density using a microplate reader at 540 nm.

2.7.1 Interpretation of MTT assay

Cell survival for each treatment condition was directly correlated with absorbance values. Each absorbance value from treated cells was then normalized against a negative control (cells with no drug treatment) and normalized values depicting the survival rate of cells was then used to compute the IC50 values. IC50 is the molar concentration of inhibitor that produces 50 % of the maximal possible inhibitory effect of that agonist. The action of the agonist may be stimulatory or inhibitory. IC50 of all cell lines was calculated by using GraphPad Prism 5. Data from proliferation assay were analysed using nonlinear regression of log dose-response. The program finds the bestfit values of the parameters from the model to the obtained data from MTT assay.

2.8 Click-iT® EdU Assay

The Click-iT® EdU Assay from Invitrogen is a novel alternative to the BrdU assay, the most common technique used for cell cycle identification. EdU (5-ethynyl-2´-deoxyuridine) provided in the kit is a nucleoside analog of thymidine and is incorporated into DNA during active DNA synthesis. Detection is based on a click reaction, 2-5 a copper-catalyzed covalent reaction between an azide and an alkyne. In this application, the EdU contains the alkyne and the Alexa Fluor® dye contains the azide. (Figure 22)

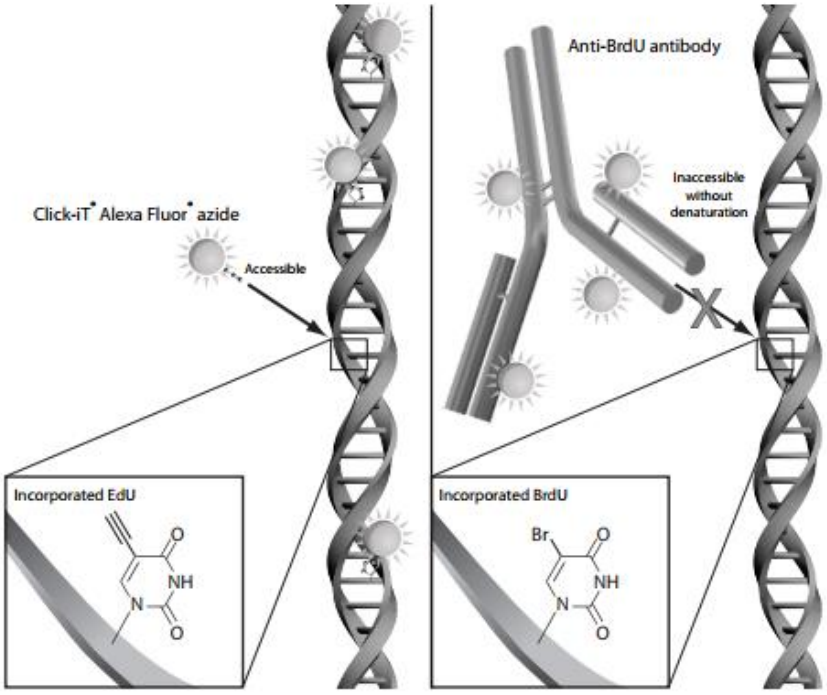


Figure 22. Click-iT® EdU Imaging Kit

This experiments was followed using the Click-iT® EdU Imaging Kit manufacturer protocol

2.9 Citofluorimetry

Before start:

PI solution: 1 g/L Sodium Citrate

10 mg/ml RNase

1. mL/L NP40

50 mg/L Propidium Iodide (PI)

24h post seeding + Treatment cells were trypsinized, centrifugated and resuspended in PI solution. After 30 mins, or more, at 37°C, all RNAs were digested and sample analyzed at cytofluorimeter.

2.10 Antibodies's list

1. γH2AX mouse : Millipore cat #05-636
2. γH2AX Rabbit : Cell signalling cat #20E3
3. 53BP1 : Novus cat #NB-100-305
4. p53BP1 : Cell signalling cat #S1778
5. ATM : SantaCruz cat #SC23921
6. pATM : Cell signalling cat #S1981
7. pATR : Cell signalling cat #2853S
8. Chk2 : Cell signalling cat #2662S
9. pChk2 : Cell signalling cat #2661S
10. mCherry : Genetex cat #GTX128509
11. AntiNucleolin rabbit: Abcam cat #22758

3. RESULTS

This Chapter is divided in two main sections: 1) the first part is related to a study aimed at showing that selected compounds of a series of hydrazone derivatives and other agents are able to stabilize G4 structure in living cultured human cancer cells; 2) the second part is related to establish that G4 binders induce R loop in living cells and that G4 binders trigger genome instability with a mechanism mediated by R loop stabilization. The results of 4.1) have been included in two publications (Amato et al., 2016) (Amato et al., 2017), while the findings of 4.2) are currently unpublished (manuscript in preparation).

3.1 Section 1 : G-Quadruplex stabilization and cytotoxicity by selected hydrazone derivatives in human Osteosarcoma U2OS cells

A series of new compounds have been developed at our Department in Bologna that have been shown to bind G4 structures with higher affinity than duplex DNA (Amato et al., 2016). Thus, we aimed at establishing whether these new G4 ligands are active in living cells and which is their potency in G4 structure stabilization.

First, the cell killing activity of the compounds (FG, FA and Compound 3) was determined by the MTT assay in osteosarcoma U2OS cells (Figure 26) and HCT116 (Figure 27).

Then, we tested whether these compounds are able to stabilize G4 structures in living cells by immunofluorescence microscopy using the specific BG4 antibody. Thus, we evaluated G4 foci formation at 24 hours of treatment with 10 μM of FG and FA (Figure 23a) in human osteosarcoma U2OS cells (Figure 23a). In these experiments, we used PDS (Figure 23a) and Braco-19 (Figure 23a) at 10 and 15 μM, respectively, as positive controls as they are well-known highly selective G4 binders (Rodriguez et al., 2012) (Biffi, Tannahill, et al., 2013).

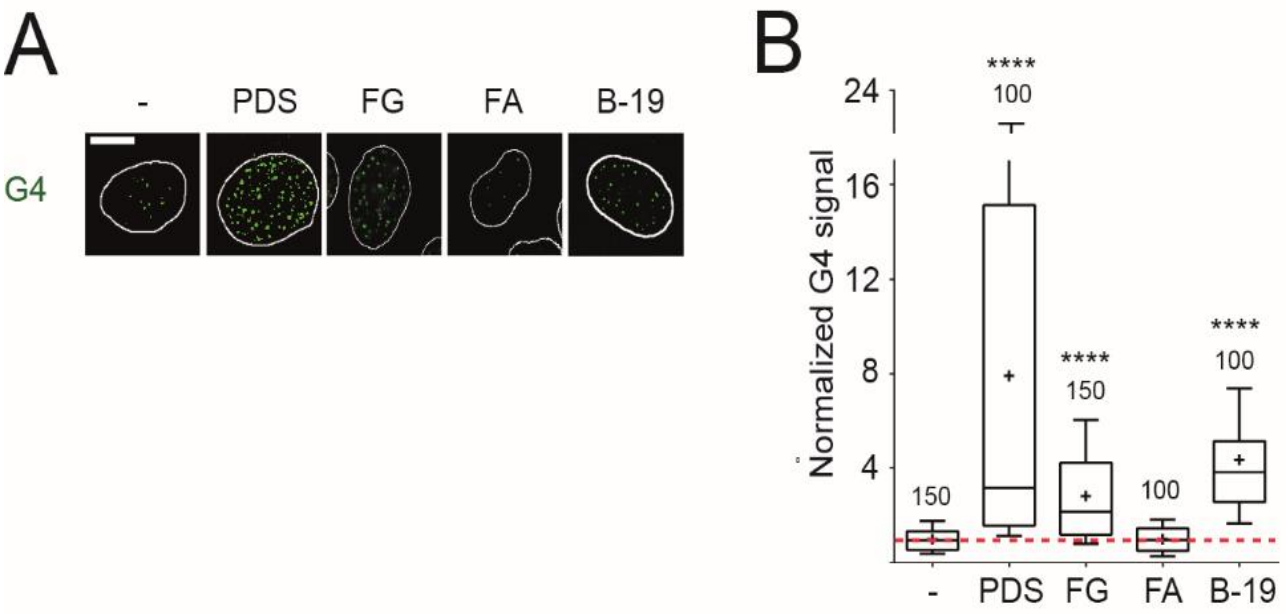


Figure 23. a) Immunofluorescence experiment on Osteosarcoma U2OS cells treated with 10 µM of PDS, FG and FA for 24 hours and labelled with BG4 antibody; b) Quantification of nuclear BG4 signal, The data were plotted after normalization over untreated control cells. The number of nuclei analyzed is reported for each sample on the bars of the graph. The * indicate the significance of the variation compared to the control sample. The number of * is directly correlate with the significant increase calculated with statistical parametric tests "Kolmogorov/Smirnov".

The quantification shows that PDS, B-19 and FG, but not FA, increase the number of G4 spots in U2OS cells (Figure 23b). The data also show that FA is an inactive analogue, and therefore we use it as a negative control in the experiments of section 2.

Next, FG, FA and Compound 3 activities in stabilizing G4 in living cells were compared by immunofluorescence microscopy at two different concentrations (Figure 24)

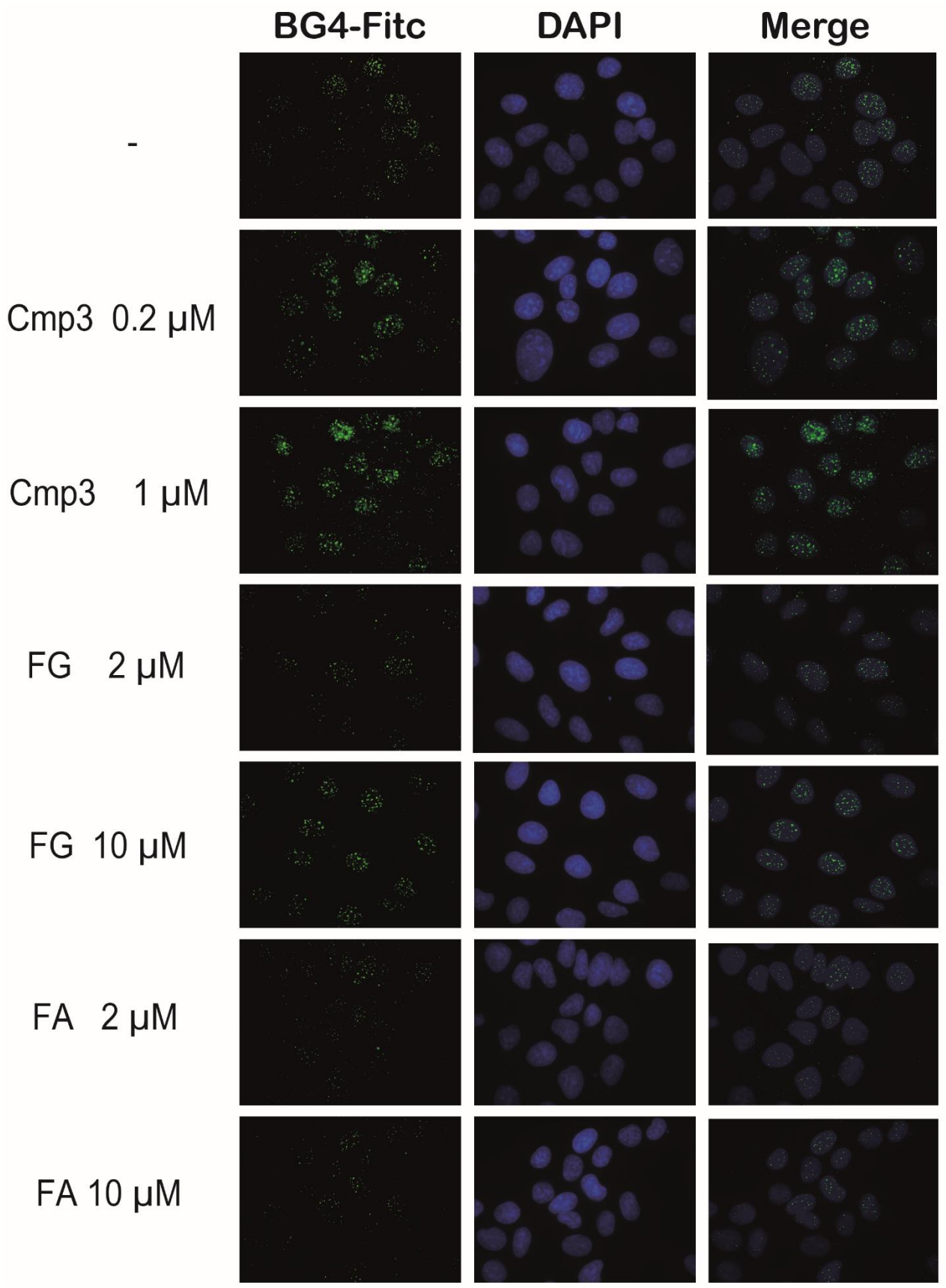


Figure 24. Immunofluorescence experiment on Osteosarcoma U2OS cells treated with 10 µM of FG , FA and Compound 3 at 24 hours at two different concentration and labelled with BG4 antibody.

The compounds effects were very different, indeed the compounds 3 and FG, but not FA, increased the number and intensity of BG4 spots indicating that they stabilize the G4 structures in the nucleus of living cells (Figure 24). We then measured the total fluorescence signal per cell for each sample and the results indicated that compound 3 was the most active G4 stabilizer because it increase the signal already at 0.2 μM and moreover at 1 μM. FG was able to increase BG4 nuclear signal at 2 μM and moreover at 10 μM. In the last case the value is comparable with the levels of compound 3 at 0.2 μM (Figure 25). In contrast, FA did not show any increase in fluorescence intensity at the concentrations tested (Figure 25). Statistical analyzes showed that a significant increase in BG4 signal after 0.2 μM and 1 μM of compound 3 and 10 μM of FG (Figure 25).

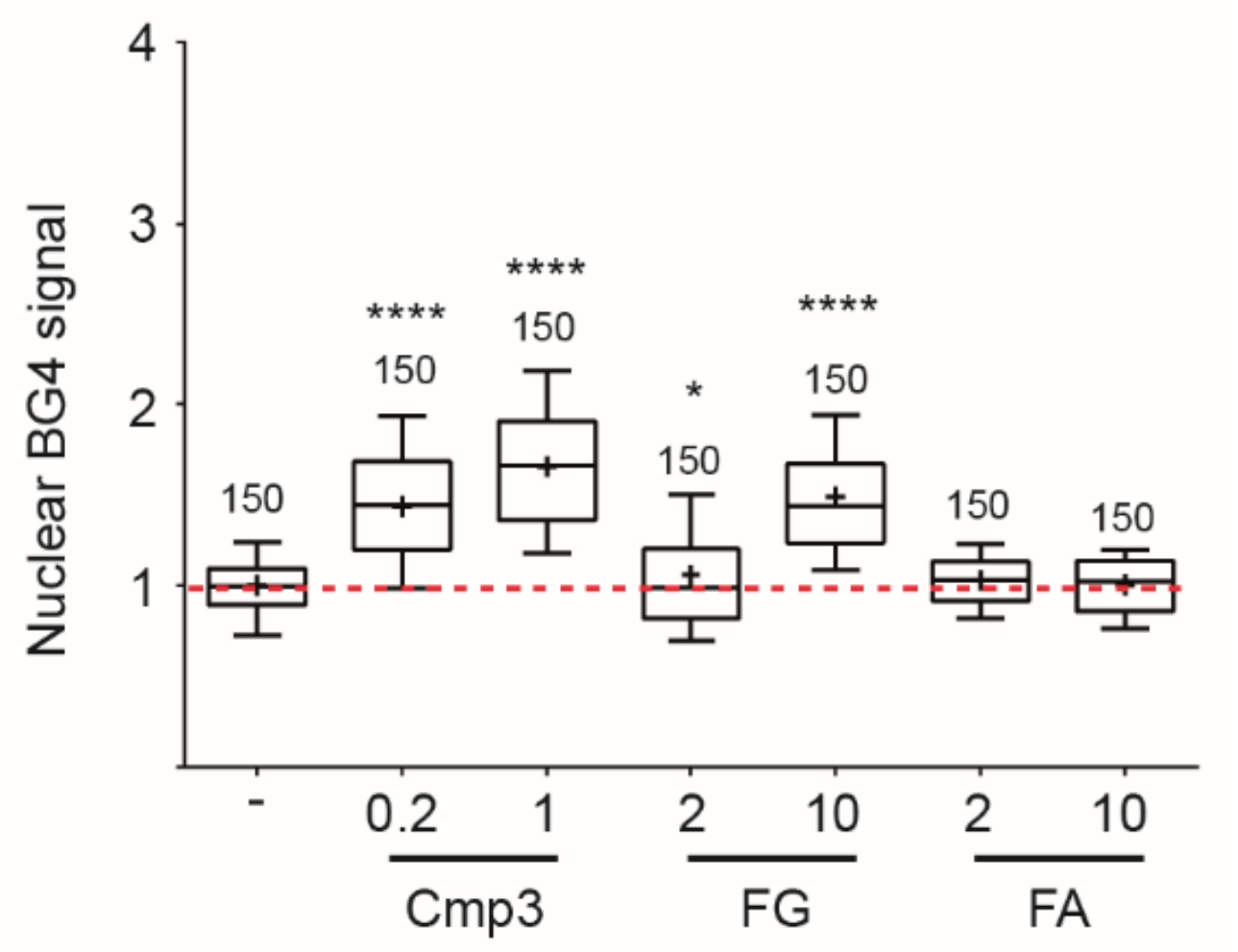


Figure 25. Quantification of nuclear BG4 signal, in cells treated with different concentration of FG, FA and Compound 3. The data were plotted after normalization over untreated control cells. The number of nuclei analyzed is reported for each sample on the bars of the graph. The * indicate the significance of the variation compared to the control sample. The number of * is directly correlate with the significant increase calculated with statistical parametric tests "Kolmogorov/Smirnov".

To be sure that the concentration of FG, FA and Compound 3 tested is not cytotoxic for the cells, it was performed a MTT assays on Osteosarcoma U2OS cells (Figure 26) (Amato et al., 2016). In each experiment, six concentrations (0.1 - 1 - 5 - 10 - 20 - 50 – 100 µM) in triplicate were assessed for each compound (Figure 26).

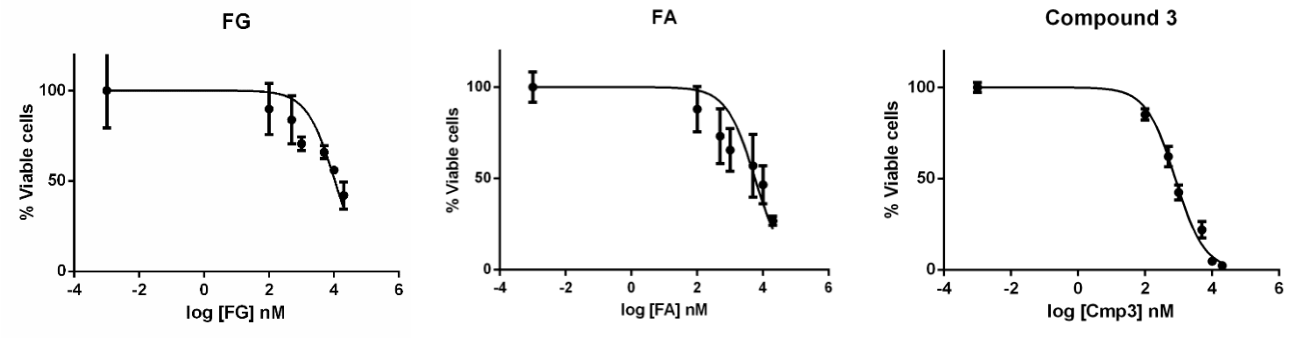


Figure 26. Results of MTT performed by treating Osteosarcoma U2OS cells for 24 hours after seeding with different concentration of FG, FA and Compound 3. Cells were then incubate for 42 hours in drug-free medium and MTT was added. In the graph y axis represents the percentage of viable cells vs Compound concentration in the x axis. Untreated cells represent 100% viable cells.

The results showed high cytotoxicity activity for compound 3 after 24 hours of treatment (IC50=0.773 µM). In addition, compound 3 was 20- and 12-fold more cytotoxic than FG (IC50=15.97 µM) and FA (IC50=9.16 µM), respectively.

The same experiment were performed in another cancer line, HCT 116 (Figure 27)

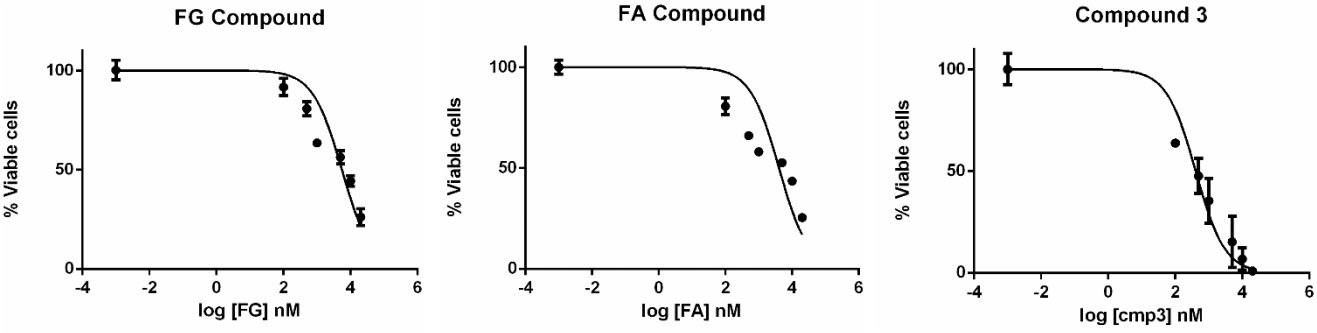


Figure 27. Results of MTT performed by treating HCT116 cells for 24 hours after seeding with different concentration of FG, FA and Compound 3. Cells were then incubate for 42 hours in drug-free medium and MTT was added. In the graph y axis represents the percentage of viable cells vs Compound concentration in the x axis. Untreated cells represent 100% viable cells.

In HCT116 cells, the results showed similar results. For compound 3 after 24 hours of treatment, the IC50 was the lowest (0.407 µM). Indeed, compound 3 was 12- and 10-fold more cytotoxic than FG (IC50=5.122 µM) and FA (IC50=4.062 µM), respectively.

Then, in collaboration with A. Randazzo's laboratory at the University of Naples, we performed biological assays for a new compound, called B3F10, which has been synthesized as a binder of a new non-B DNA structure, the G-Triplex motif. IF experiments showed an increase of G4 signal at different concentrations (Figure 28a) showing that B3F10 can stabilize G4 structure in living cells.

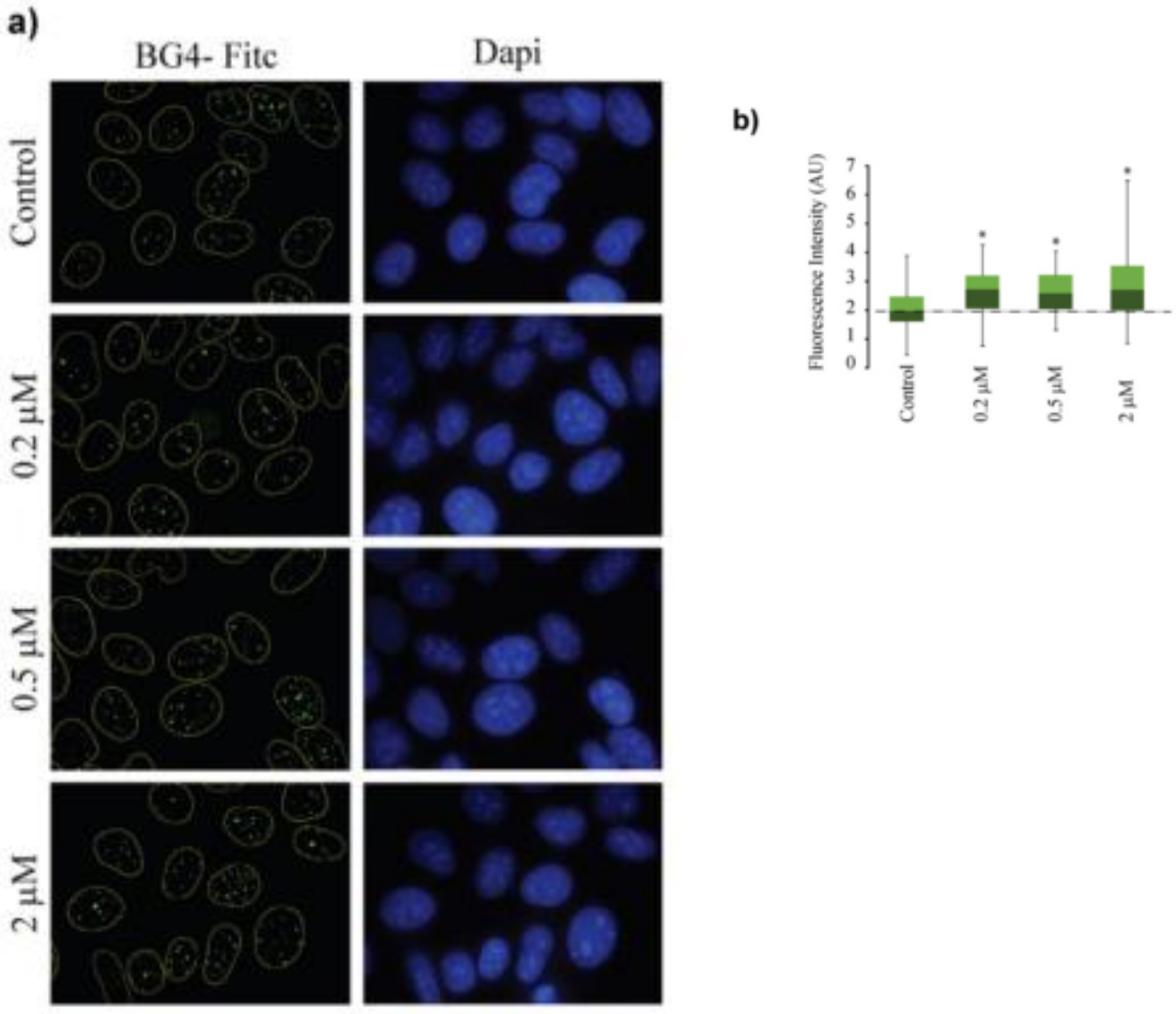


Figure 28. a) Immunofluorescence experiment on Osteosarcoma U2OS treated with different concentration of B3F10 compound at 24 hours and labelled with BG4 antibody b) Quantification of nuclear BG4 signal; 150 nuclei were selected, 75 for each esperiment. The * indicate the significance of the variation compared to the control sample. The number of * is directly correlate with the significant increase calculated with statistical parametric tests "Kolmogorov/Smirnov".

The drug treatment determines a slight but significant increase in the BG4 intensity in the nucleus at all tested concentrations, demonstrating that the compound B3F10 is a stabilizer of G4 structures in human U2OS cells. (Figure 28b).

The whole data confirmed the stabilizer activity of 2 newly synthesized compounds, FG and Compound 3. FA results to be an excellent negative control for future analyzes by it's nature of inactive analogue of FG and Compound 3. Finally, compound B3F10 confirmed double stabilized action on G-Triplex / G-Quadruplex structures; however, the stabilizing activity on G-Triplex could not be confirmed for the absence of specific antibodies that can recognize these structures.

3.2 Section 2: Genome instability caused by G-Quadruplex binders is mediated by R loops in human cancer cells

3.2.1 Immediate stabilization of G4 structures by the studied binders in human cancer cells

In Section 1, the data show that the studied compounds stabilize G4 in chromatin after 24 hours of treatment, a common experimental condition used by several labs (Biffi, Tannahill, et al., 2013). However, a number of pathways are likely activated in cells at that time, which might interfere with the action of G4 binders. In order to better understand the cellular activity of G4 binders, time course experiments have been performed with the tested agents. Surprisingly, the results demonstrate that the three G4 binders (PDS, FG and B-19) but not the inactive FG analog (FA) stabilize G4 structures after few minutes of treatment, followed by a drastic reduction of G4 foci at 30 minutes (Figure 29a). The quantification shows a bell-shaped trend of G4 signal in cells treated with PDS, B-19 and FG but not with FA (Figure 29b). The data clearly show that G4 stabilization is an immediate effect of G4 ligands as detected by immunofluorescence microscopy with BG4 antibody.

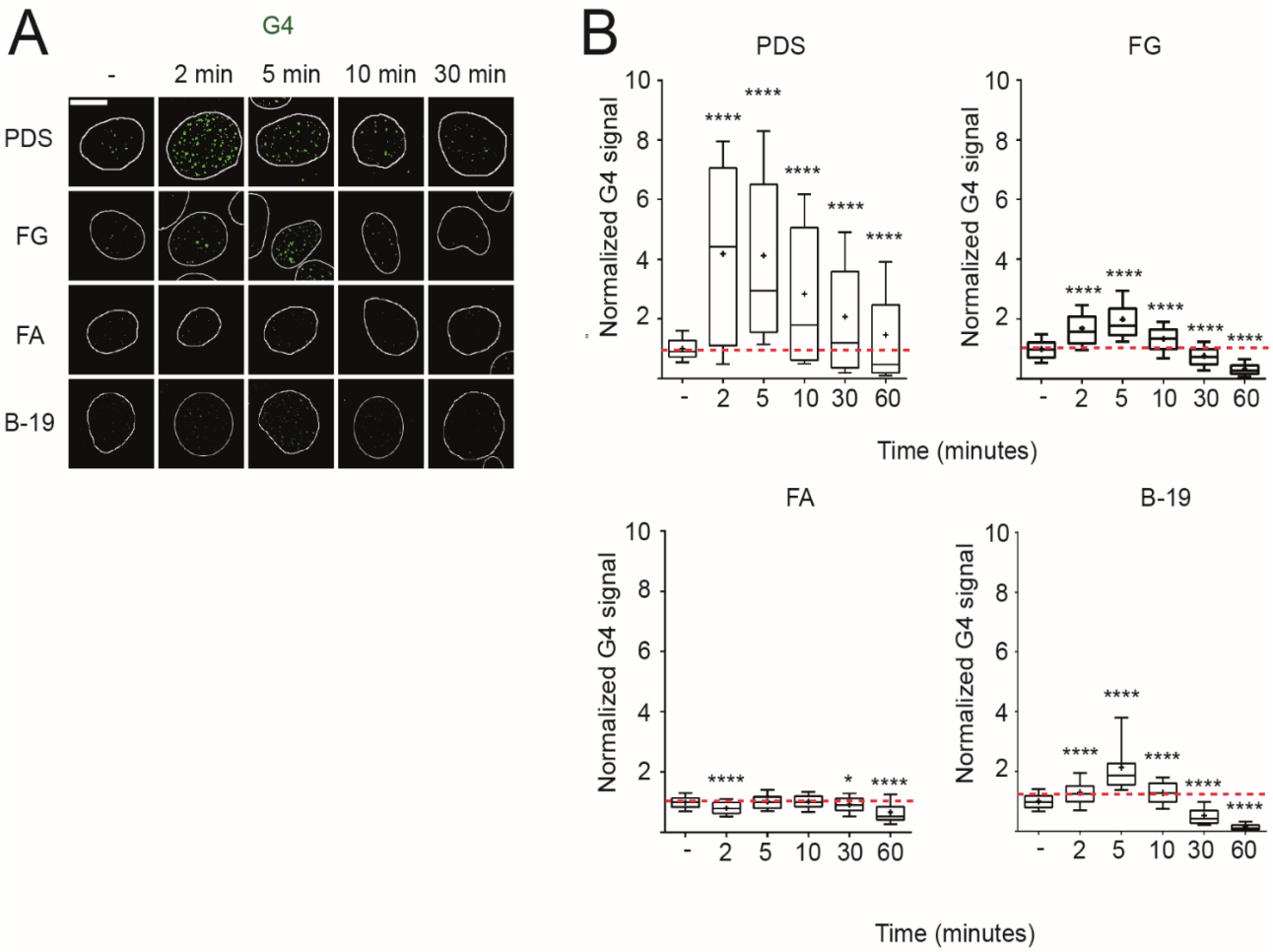


Figure 29. a) Time course experiment on Osteosarcoma U2OS cells treated with 10 µM of PDS, FG, FA and 15 µM of B-19 for and labelled with BG4 antibody; b) Quantification of nuclear BG4 signal; 150 nuclei were selected, 75 for each esperiment, the data was plotted after control normalization. The * indicate the significance of the variation compared to the control sample. The number of * is directly correlate with the significant increase calculated with statistical parametric tests "Kolmogorov/Smirnov".

3.2.2 R loop stabilization by G4 binders in human cancer cells

As discussed on the introduction section, GC skew is an accurate predictor of both G4 and R-loop formation at promoter regions (Ginno et al., 2013) (Hartono, Korf, & Chédin, 2015). Thus, we asked the question of whether G4 ligands can favor R loop stabilization as well. Then, we tested with IF experiments using the S9.6 antibody, the capacity of the G4 binders to stabilize R-loop structures at different time points.

First, we treated the cells for 24 hours with the agents and stained the nucleolus with an anti-Nucleolin antibody to distinguish nucleolar from nucleoplasm R loops. The results of such experiments are shown in Figure 30 and document that G4 ligands, but not an inactive compound, trigger the formation of R loops at higher levels than in control cells. The main part of the G4 ligand-induced R loop signal is in the nucleoplasm of cells (Figure 30a).

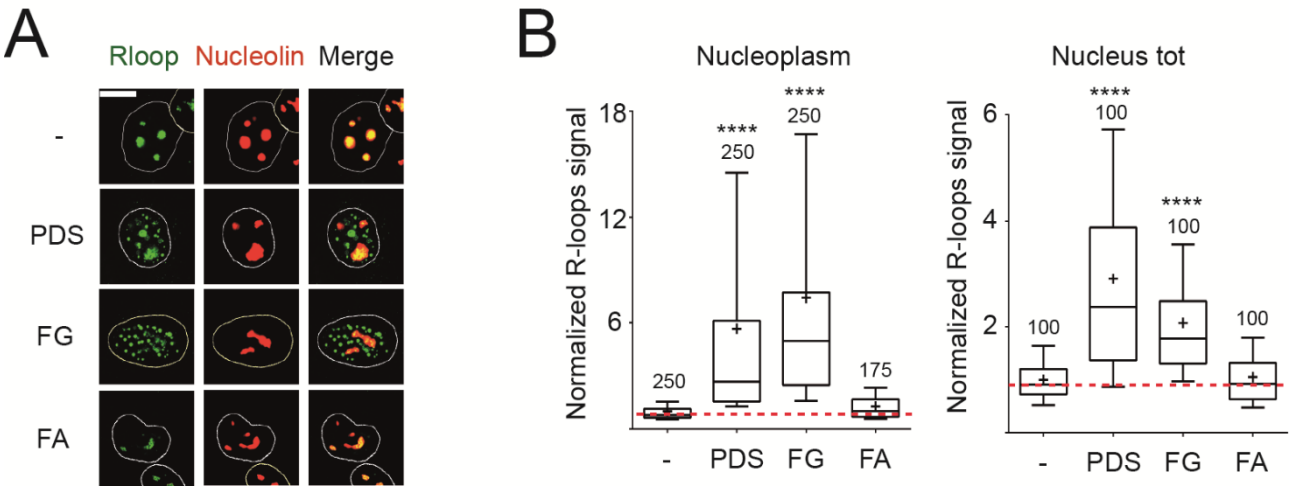


Figure 30. a) Immunofluorescence experiment on Osteosarcoma U2OS cells treated with 10 µM of PDS, FG and FA for 24 hours and labelled with S9.6 antibody; b) Quantification of nuclear (right chart) and nucleoplasmatic (left chart) S9.6 signal; the data was plotted after control normalization. The nucleolus was excluded by overlay the signal obtained by staining with the nucleolar marker (anti-nucleolin). The number of nuclei analyzed is reported for each sample on the bars of the graph. The * indicate the significance of the variation compared to the control sample. The number of * is directly correlate with the significant increase calculated with statistical parametric tests "Kolmogorov/Smirnov".

The signal quantification also shows a significant increase of R loops in the nucleus but an even higher increase when considering the nucleoplasm only (Figure 30b). Moreover, dose response curves for PDS and B-19 show that the R loop increase is dependent on the concentration of the G4 ligand (Figures 31 and 32), supporting the effect of G4 ligands on R loops in living cells.

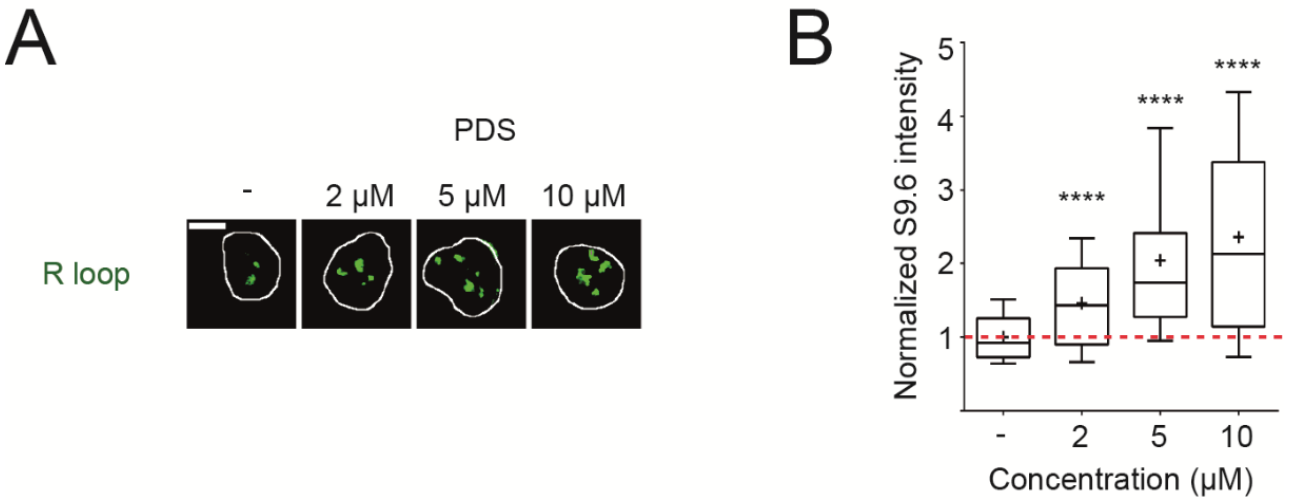


Figure 31. a) Dose response experiment on Osteosarcoma U2OS cells treated with different concentration of PDS at 24 hours and labelled with S9.6; b) Quantification of nuclear S9.6 signal, the data was plotted after control normalization. 150 nuclei were selected, 75 for each experiment. The * indicate the significance of the variation compared to the control sample. The number of * is directly correlate with the significant increase calculated with statistical parametric tests "Kolmogorov/Smirnov".

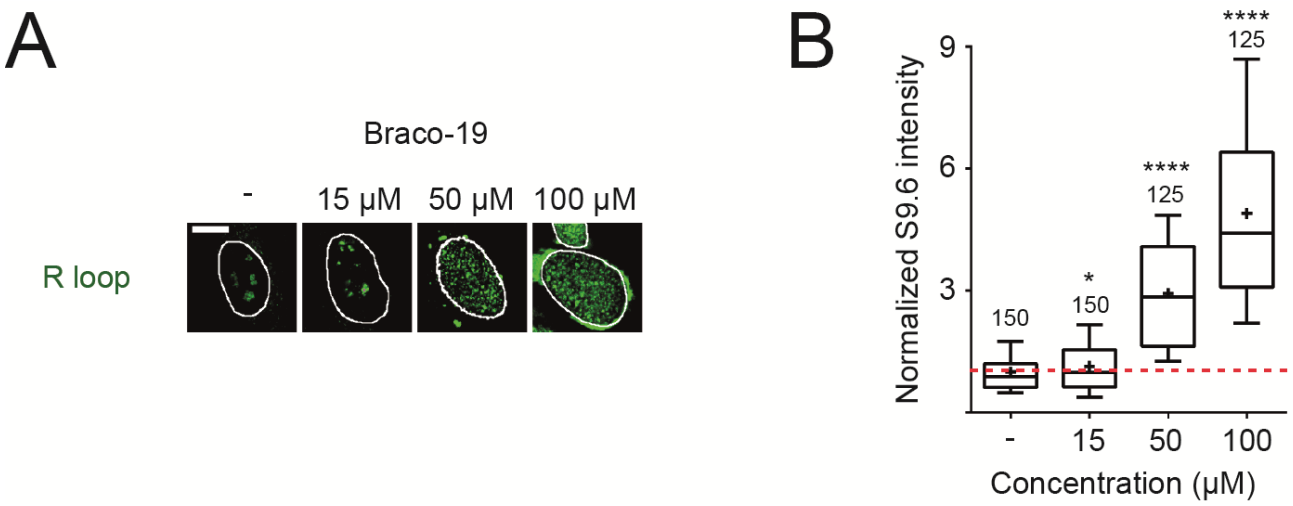


Figure 32. a) Dose response experiment on Osteosarcoma U2OS cells treated with different concentration of B-19 at 24 hours and labelled with S9.6; b) Quantification of nuclear S9.6 signal, the data was plotted after control normalization. The number of nuclei analyzed is reported for each sample on the bars of the graph. The * indicate the significance of the variation compared to the control sample. The number of * is directly correlate with the significant increase calculated with statistical parametric tests "Kolmogorov/Smirnov".

As discussed above, at 24 hours there are a number of pathways that are likely activated in cells, leaving the possibility that the R loop increase is an indirect effect mediated by cellular mechanisms. Thus, we performed a time course experiments with 10 µM of FG and PDS (Figure 33a).

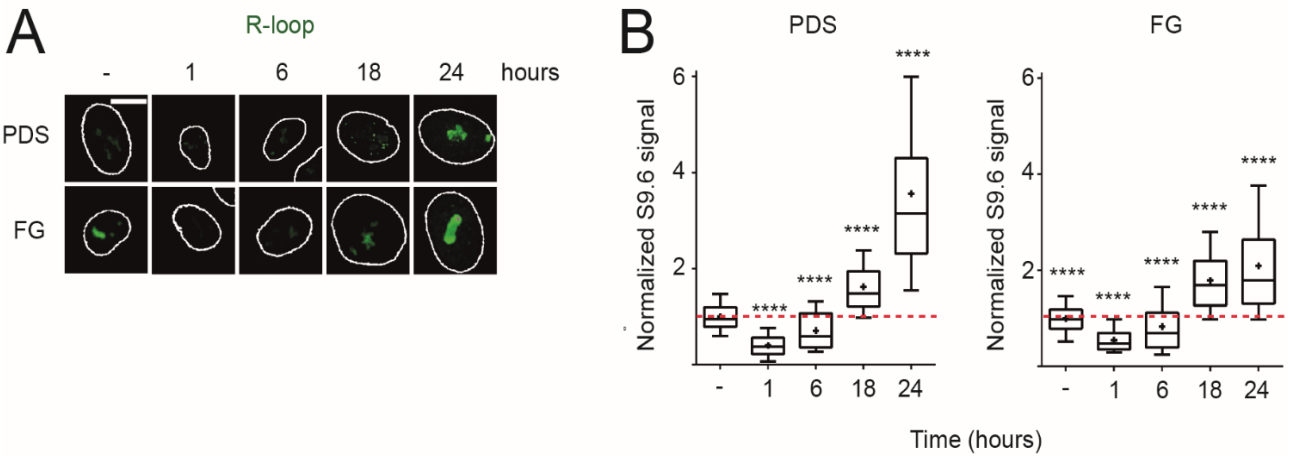


Figure 33. a) Time course experiment on Osteosarcoma U2OS cells treated with 10 µM of PDS and FG and labelled with S9.6 antibody b) Quantification of nuclear S9.6 signal, the data was plotted after control normalization. 150 nuclei were selected, 75 for each esperiment. The * indicate the significance of the variation compared to the control sample. The number of * is directly correlate with the significant increase calculated with statistical parametric tests "Kolmogorov/Smirnov".

The data show for both agents a decrease of R loop signal at 1 hour and 6 hours followed by an increase at 18 and 24 hours (Figure 33b). The same trend was obtained in cells treated with 15 and 50 µM of Braco-19 (Figure 34).

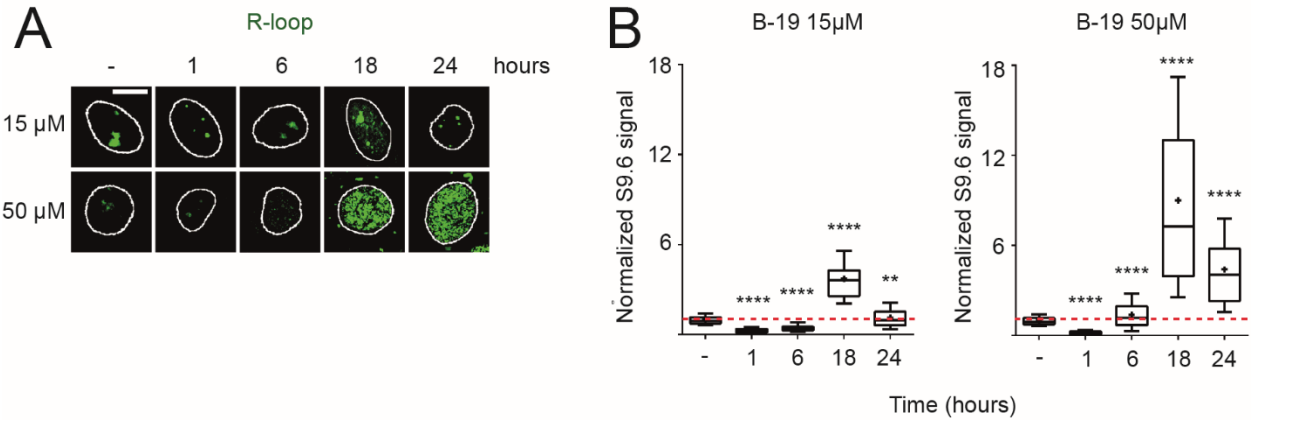


Figure 34. a) Time course experiment on Osteosarcoma U2OS cells treated with 15 and 50 µM of Braco-19 and labelled with S9.6 antibody b) Quantification of nuclear S9.6 signal, the data was plotted after control normalization. 150 nuclei were selected, 75 for each esperiment. The * indicate the significance of the variation compared to the control sample. The number of * is directly correlate with the significant increase calculated with statistical parametric tests "Kolmogorov/Smirnov".

Optical microscopy visualization of cells revealed a marked increase of cell death at 50 μM of B-19. This observation raises the question of whether the R loop signal and pattern changes are determined by cytotoxic effects of the compound or the stabilization of G4s. Thus, we wondered if there was already some change at very short treatments, which were tested for G4 structure formation (see above Figure 35).

Then, time course experiments were performed with PDS, FG, B-19 and the negative control, FA (Figure 35a). The results show that G4 ligands specifically induce an increase of R loop levels as compared to control cells at the same time of the G4 foci formation by the same ligands (Figure 35a). The quantification shows the same bell-shaped trend revealed for the G4 signal in cells treated with PDS, B-19 and FG but not with FA (Figure 35b). The R loop increase is higher when considering only the nucleoplasm (Figure 35b).

Overall, the findings demonstrate that the tested G4 ligands can increase the level of R loops at short (2-10 minutes) and long (18-24 hours) times, and that the R loop kinetics are correlated with a similar kinetic pattern of G4 stabilization by the same ligand in human cancer U2OS cells.

The bell-shaped curve might be due to a prompt cellular response to G4 stabilization, such as activity of some helicases, nucleases or other enzymes, which may attempt to resolve these structures. Such a hypothesis however needs to be confirmed by future experiments involving the silencing of specific enzymes.

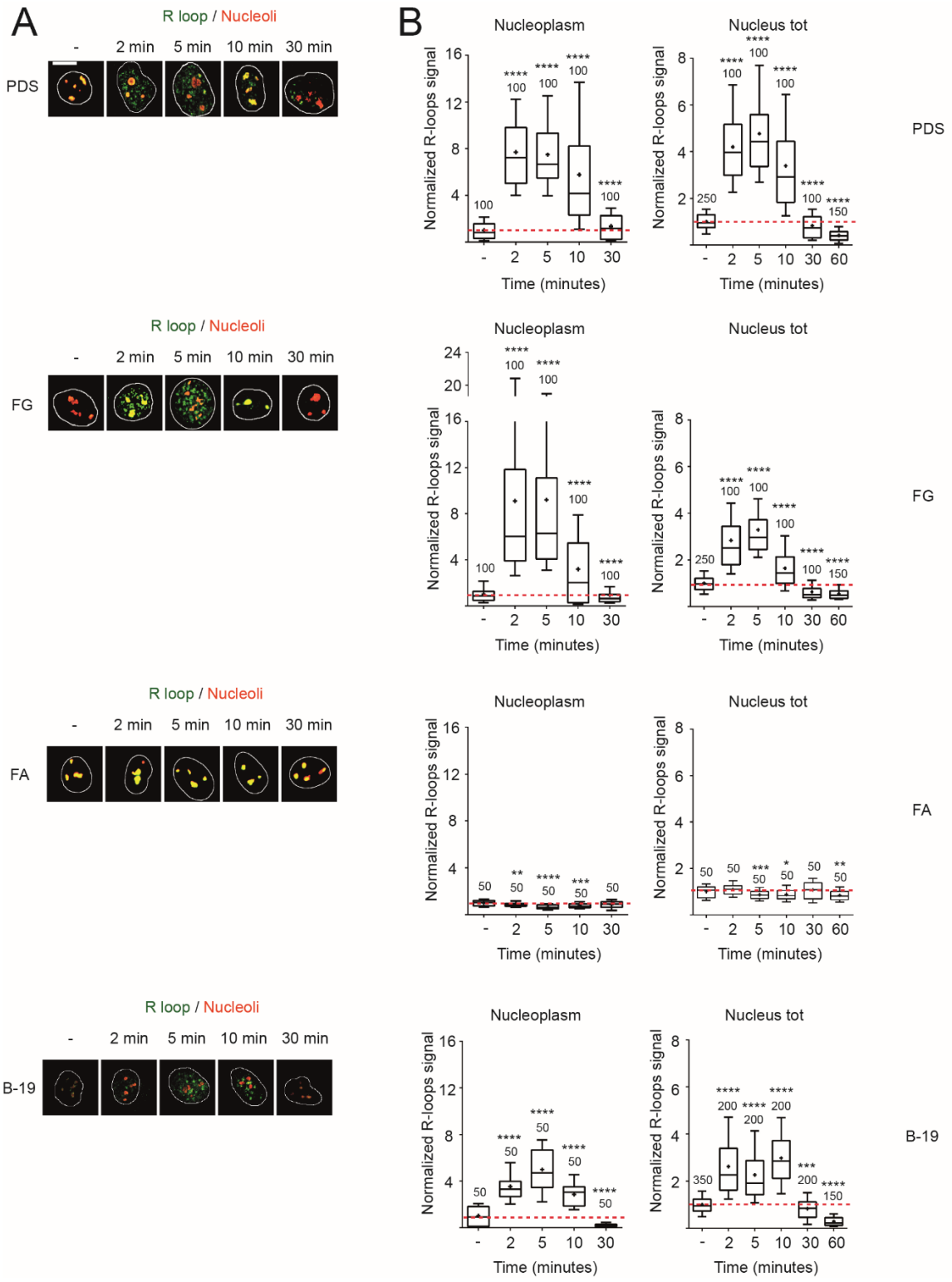


Figure 35. a) Short time course experiment on Osteosarcoma U2OS cells treated with 10 µM of PDS, FG, FA and 15 µ of Braco-19 and labelled with S9.6 antibody; b) Quantification of nuclear (right chart) and nucleoplasmatic (left chart) S9.6 signal; the data was plotted after control normalization. The nucleolus was excluded by overlay the signal obtained by staining with the nucleolar marker (anti-nucleolin). The number of nuclei analyzed is reported for each sample on the bars of the graph. The * indicate the significance of the variation compared to the control sample. The number of * is directly correlate with the significant increase calculated with statistical parametric tests "Kolmogorov/Smirnov".

In order to confirm this finding, it was performed a DNA:RNA immunoprecipitation followed by qPCR determination of fragment enrichment at selected genomic loci following 24 hours treatment with G4 ligands (Figure 36). In these experiments: negative loci is SNRPN and positive loci are RPL13A, EIF5A and MYADM. The primers for SNRPN, EIF5A and RPL13A amplify at the intronic region of the gene. These are regions already validated for the presence or absence of R-loops. The results suggest that PDS and FG can stimulate R loops at selected loci with a less that two-fold increase (Figure 36).

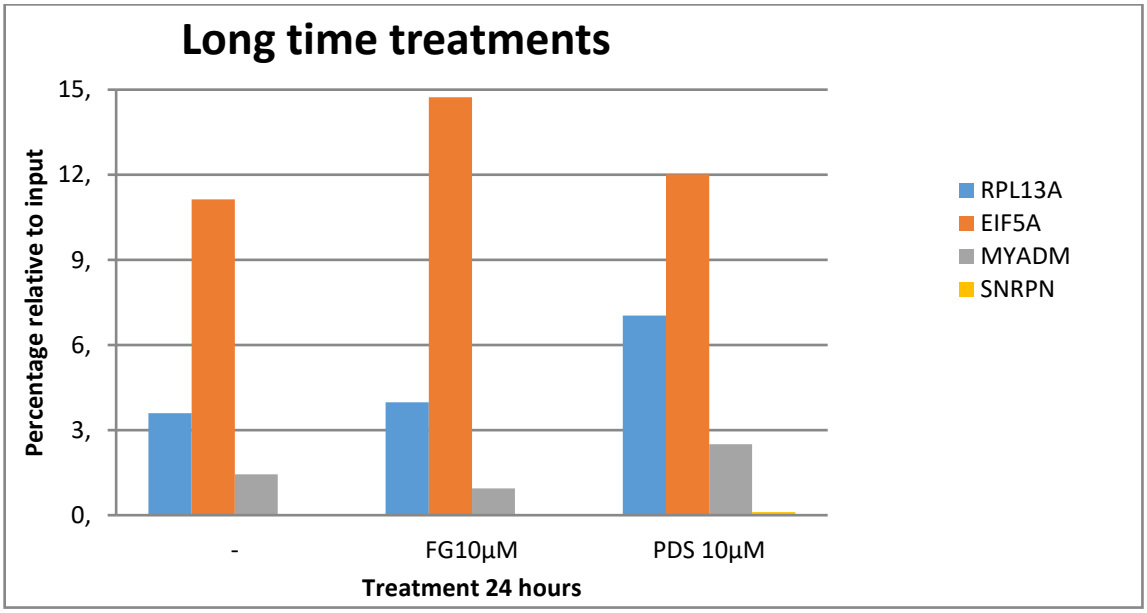


Figure 36. The graph shows the recovery (y axis) compared to the input (non-immunoprecipitated sample) for cells treated with 10 µM of FG and PDS at 24 hours (x axis). Have been selected three positive locus (RPL13A, MYADM and EIF5A) and a negative control (SNRPN).

The same DRIP experiments were performed at short time treatments and R loop levels was determined at a larger set of negative and positive locus (Figure 37). In this experiment, other negative loci, T-Nik and α-Sat and positive loci, EGR1 and LMNB2, already validated for the absence/presence of R-loops, have been added. Furthermore, primers have been designed in promoter regions of oncogenic genes known to have G4 structures, such as VEGF, C-MYC, Bcl-2

The results show that R loops are increased at several loci:

- PDS shows great increase in VEGF (4 fold increase), LMNB2 (2.7 fold increase), C-MYC (2.23 fold increase) and MYADM (2.14 fold increase) loci. Slight increase for RPL13A (1.32 fold increase).
- Braco-19 shows great increase in RPL13A (3.9 fold increase), VEGF (3.85 fold increase), LMNB2 (5.9 fold increase), C-MYC (2.16 fold increase), MYADM (2.88 fold increase) and Bcl-2 (2.14 fold increase) loci.
- FG shows great increase in RPL13A (2.71 fold increase), LMNB2 (4.28 fold increase), C-MYC (2.17 fold increase), MYADM (3 fold increase) and Bcl-2 (2.14 fold increase) loci.
- FA shows only an increase at VEGF locus (4 fold increase).

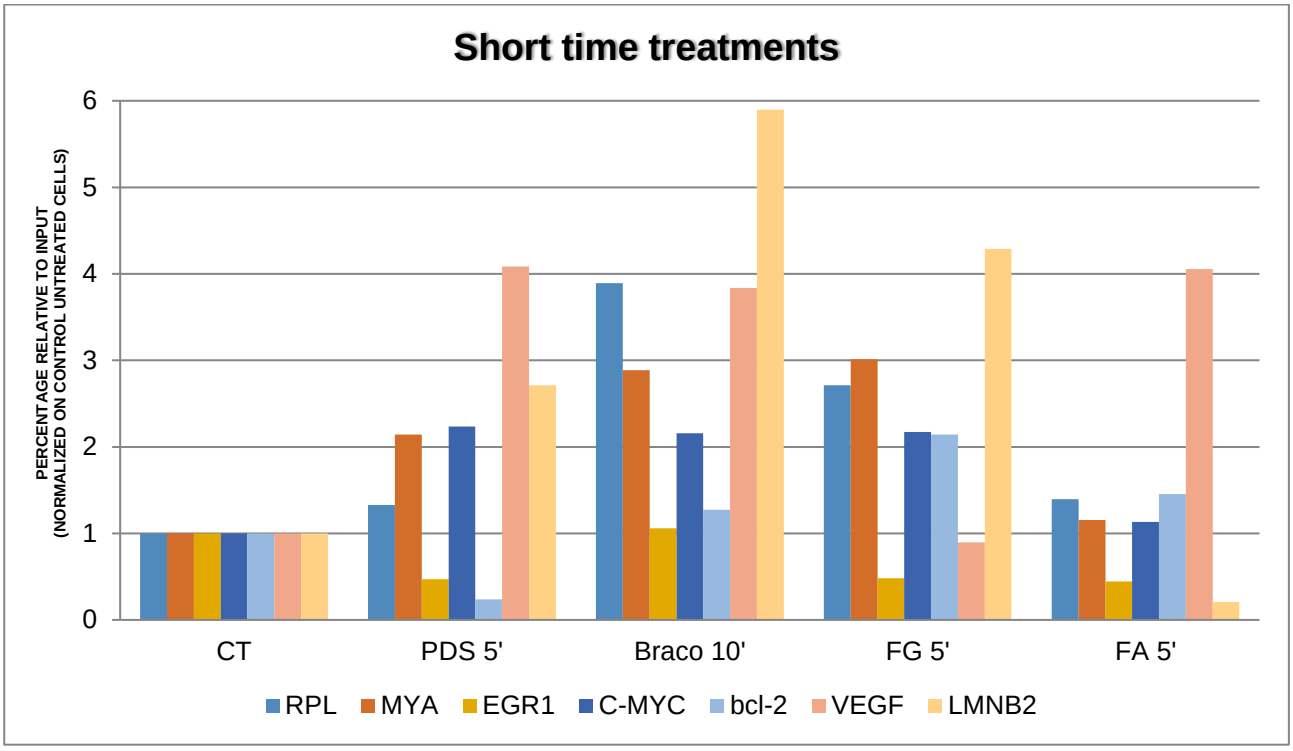


Figure 37. The graph shows the recovery (y axis) compared to the input (non-immunoprecipitated sample) for cells treated with 10 µM of FG, FA and PDS at 5 minutes and 15 µM of Braco-19 at 10 minutes (x axis). CT, control untreated cells.

Genomic regions known to not form R loops (negative loci) showed an enrichment at least 100-fold less (data not shown) than positive loci, shown in Figure 37.

To better characterize the IF signal revealed by the S9.6 antibody, we transiently overexpressed the human RNaseH1 in Osteosarcoma U2OS cells untreated or treated with PDS for 24 hours (Figure 38a). RNaseH1 is an RNA nuclease that specifically degrades RNA annealed to a DNA strand, therefore its overexpression should reduce S9.6-depedent fluorescence signals in living cells. Cells were transfected with a plasmid containing a cytoplasmic GFP or a GFP-RNaseH1 fusion protein. (Figure 38a).

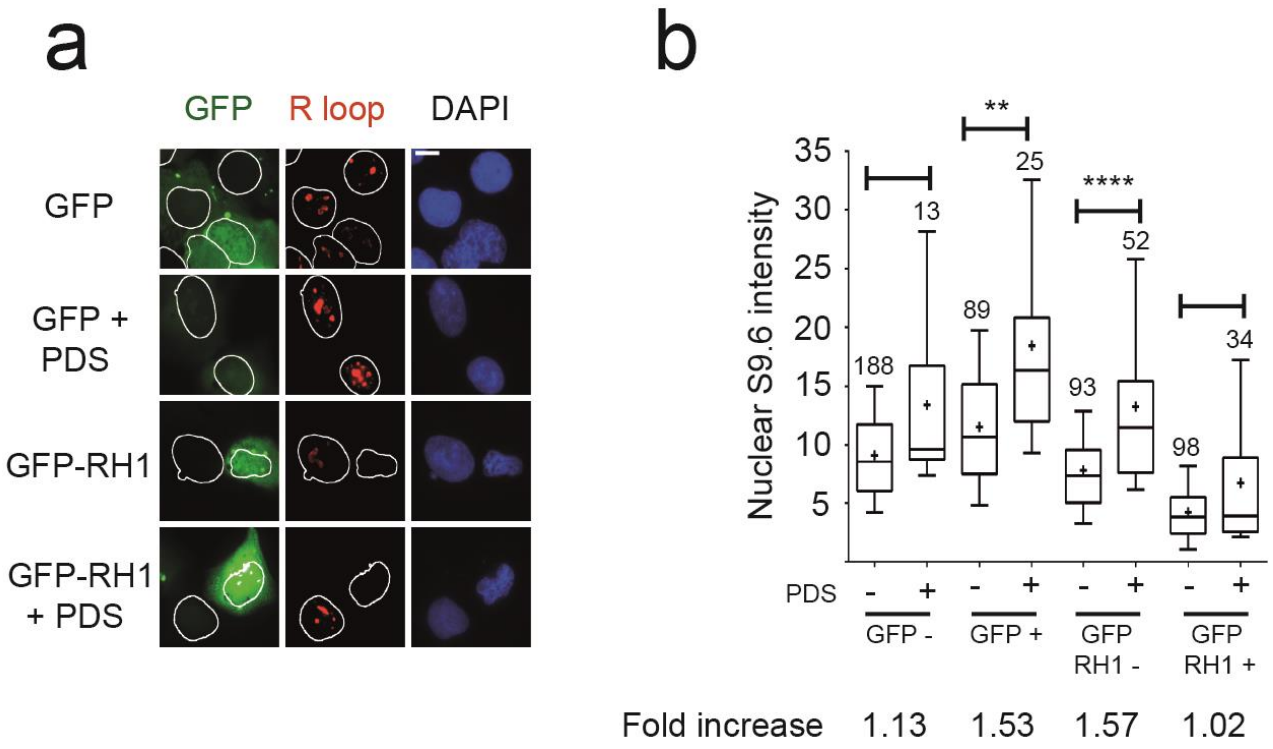


Figure 38. a) Immunofluorescence experiments on Osteosarcoma U2OS cells transfected with transient GFP and GFP-RNaseH1, treated/untreated with 10 µM of PDS at 24 hours and labelled with S9.6 antibody; b) Quantification of S9.6 signal, shows an increase of signal in cells positive e negative for GFP after PDS treatment. RNaseH1 positive cells shows a rescue of r-loops signal compared to RNaseH1 negative cells. The number of nuclei analyzed is reported for each sample on the bars of the graph. The * indicate the significance of the variation compared to the control sample. The number of * is directly correlate with the significant increase calculated with statistical parametric tests "Kolmogorov/Smirnov".

Cells were then quantified by dividing the population into positive and negative cells for transfection. The results show that RNaseH1 almost abolishes R loop signal in cells untreated and treated with PDS 10 µM for 24 hours, and significantly abolishes the PDS-dependent increase of R loop signal (see fold increase in Figure 38b).

Moreover, RNaseH1 treatment of genomic DNA before immunoprecipitation in DRIP experiments fully abolishes detection of R loops at the studied loci (Figure 39).

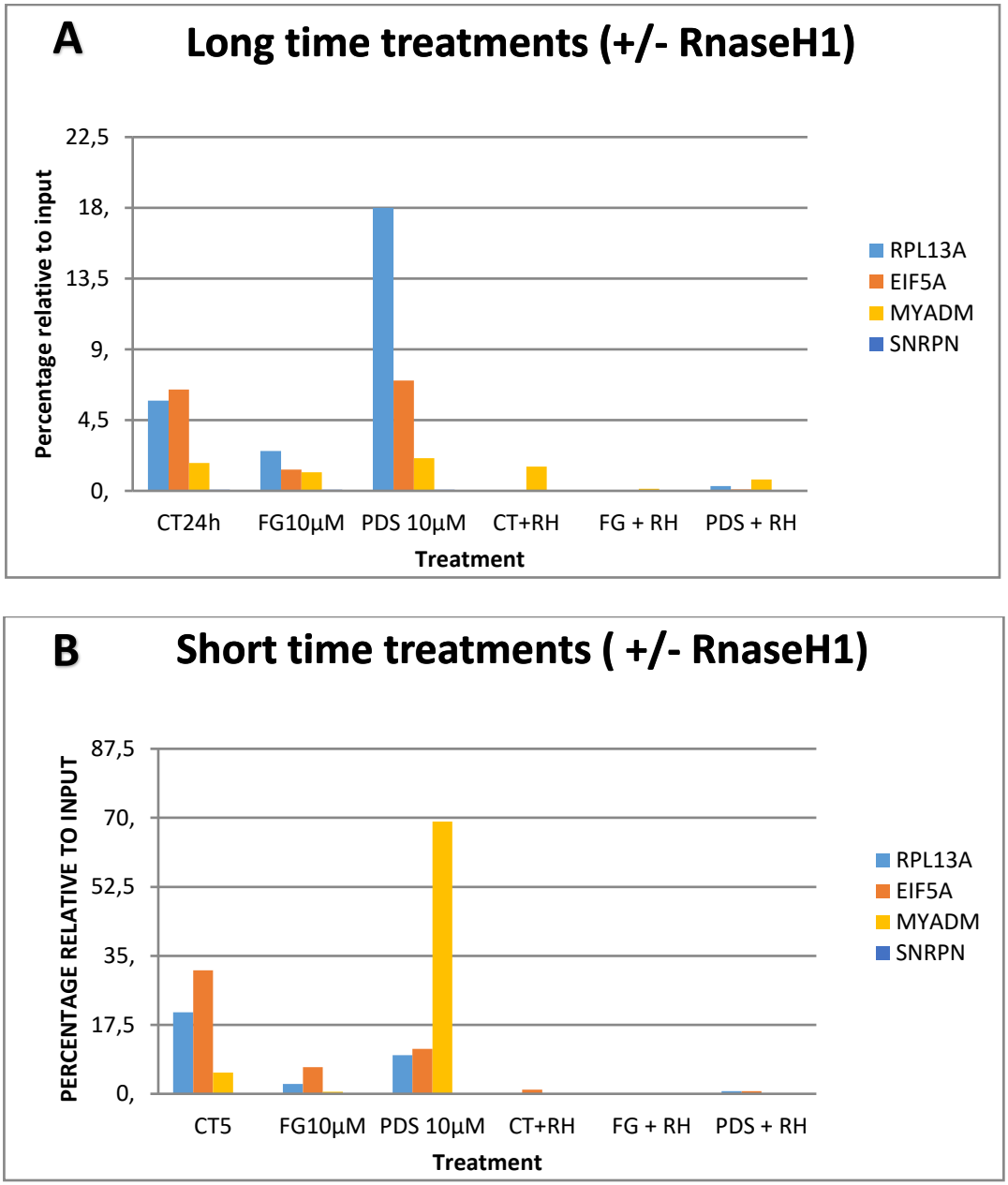


Figure 39. The graphs show the recovery decrease in treated/untreated cells incubated with an exogenous RNaseH1 before immunoprecipitation; a) Osteosarcoma U2OS cells treated with 10 µM of FG and PDS for 24 hours b) Osteosarcoma U2OS cells treated with 10 µM of FG and PDS for 5 minutes.

The data show an increase of R loop signal (as percentage of DNA recovery relative to input) in cells treated with PDS and FG for long (Figure 39a) or short times (Figure 39b). The incubation with an RNaseH1 fully abolishes the R loop signal (Figures 39).

The whole data represent the first demonstration that stabilization of G-Quadruplex structures by G4 binders can determine an increase of R loop levels in human cancer cells

3.2.3 Co-localization of R loop and G4 foci in nuclear chromatin of U2OS cancer cells

The next step was to develop a co-labelling protocol to identify a co-localization between G4 and R-loop structures (Figure 40). The starting point was to analyze the differences in fixation, permeabilization, blocking, incubation and washing between the two protocols. Once the differences were assessed, a screening of several co-labeling protocols was performed in which one or more steps were changed. In the end, we developed a co-labeling protocol that differs from both starting procedures to allow the staining of both R loops and G4s (Figure 40).

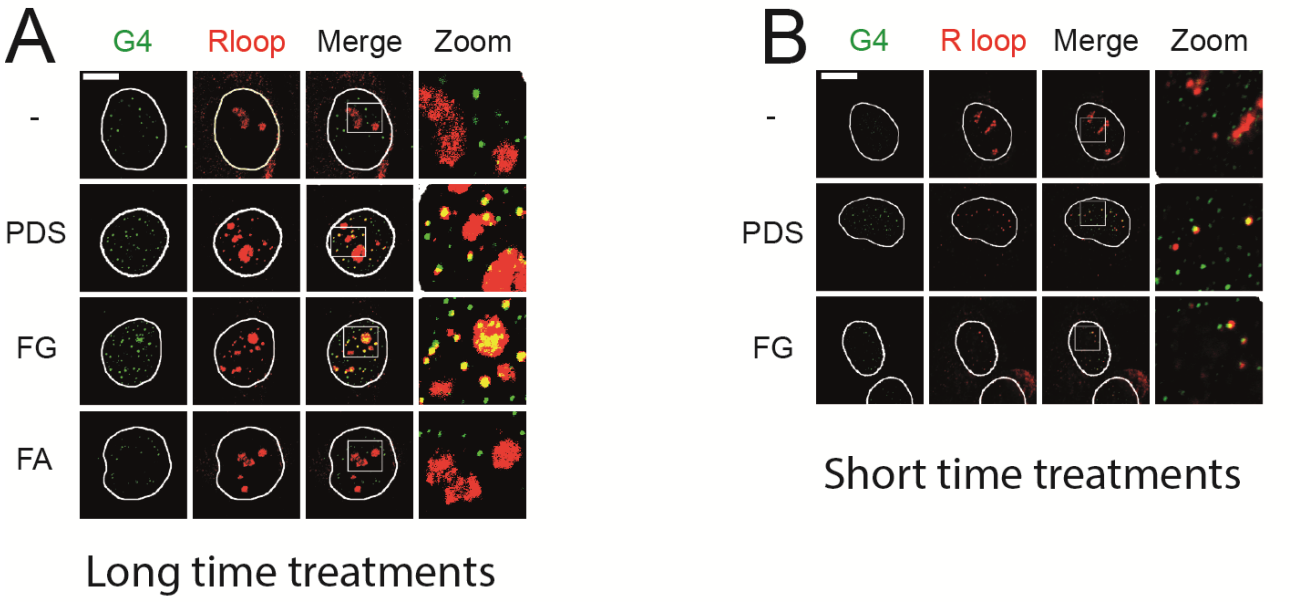


Figure 40. a) Immunofluorescence experiment on Osteosarcoma U2OS cells treated with 10 µM of PDS, FG, FA at 24 hours and co-labelled with BG4 (green) and S9.6 (red) antibodies. b) Immunofluorescence experiment on Osteosarcoma U2OS cells treated with 10 µM of PDS, FG at 5 minutes and co-labelled with BG4 (green) and S9.6 (red) antibodies.

Co-labeling results show an increase of merged spots between the two antibodies after treatment with PDS and FG at short (Figure 40b) and long times (Figure 40a). In the control and the sample treated with FA, the quantity of merged foci is very low either at short and long treatment times. (Figure 40).

3.2.4 Ligand-induced R loop and G4 foci are not detected in normal human cells

It has been reported that cancer cells possess more G4 structures than primary normal cells (Rodriguez et al., 2012). Therefore, we tested G4 and R loop levels by IF after 24 hours treatments with 10 µM of PDS, FG and FA of human normal WI-38 fibroblasts and measured the R loop signal as above. The results show that PDS, FG and FA are completely ineffective in stabilizing G4 structures or R loops in normal cells (Figure 41).

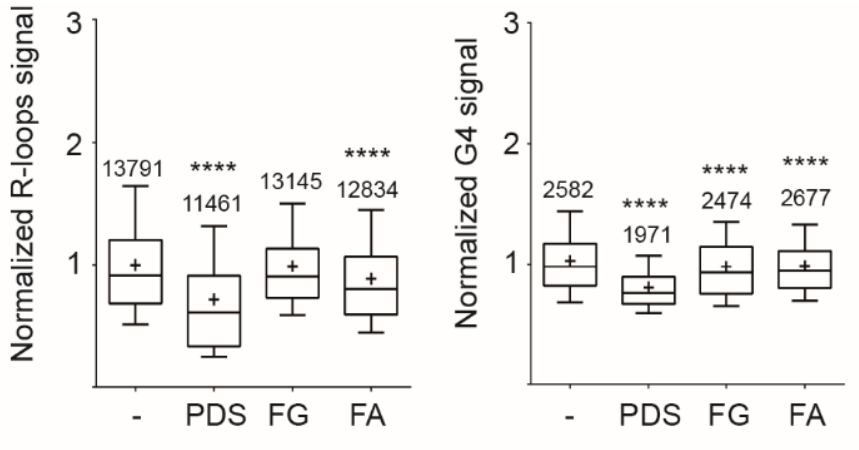


Figure 41. Quantification of S9.6 (left graph) and BG4 (right graph) signal in WI38 fibroblasts treated with 10 µM of PDS, FG or FA for 24 hours. The data were plotted after normalization over untreated control cells. The number of nuclei analyzed is reported for each sample on the bars of the graph. The * indicate the significance of the variation compared to the control sample. The number of * is directly correlate with the significant increase calculated with statistical parametric tests "Kolmogorov/Smirnov".

As we detected a biphasic kinetics within 1 hour treatment in U2OS cells, we repeated the same kinetics in WI-38 cells to further investigate R loop and G4 formation by the studied agents. The results were striking and in full contrast with U2OS cells. Again, PDS, FG, FA and B-19 did not induce G4 foci nor increase R loop levels in WI-38 cells under our experimental conditions (Figure 42). It is interesting to note, however, that PDS, but somewhat also the other G4 ligands, decreases R loops and G4s at the tested times (Figures 41 and 42). These data are in full agreement with published data on G4 formation in cancer vs. normal cells (Rodriguez et al., 2012), in which they showed that cancer cells present more G4s as compared to normal cells and are more susceptible to G4 binder treatment. This could be due to high replication and transcription rates of tumor than normal cells. That may determine that a more dynamical state of the genomic DNA with a consequent greater rate of formation of G4s and R loops. All the data further support the findings that G4 ligands induce at the same time G4 foci and R loops in human cancer cells.

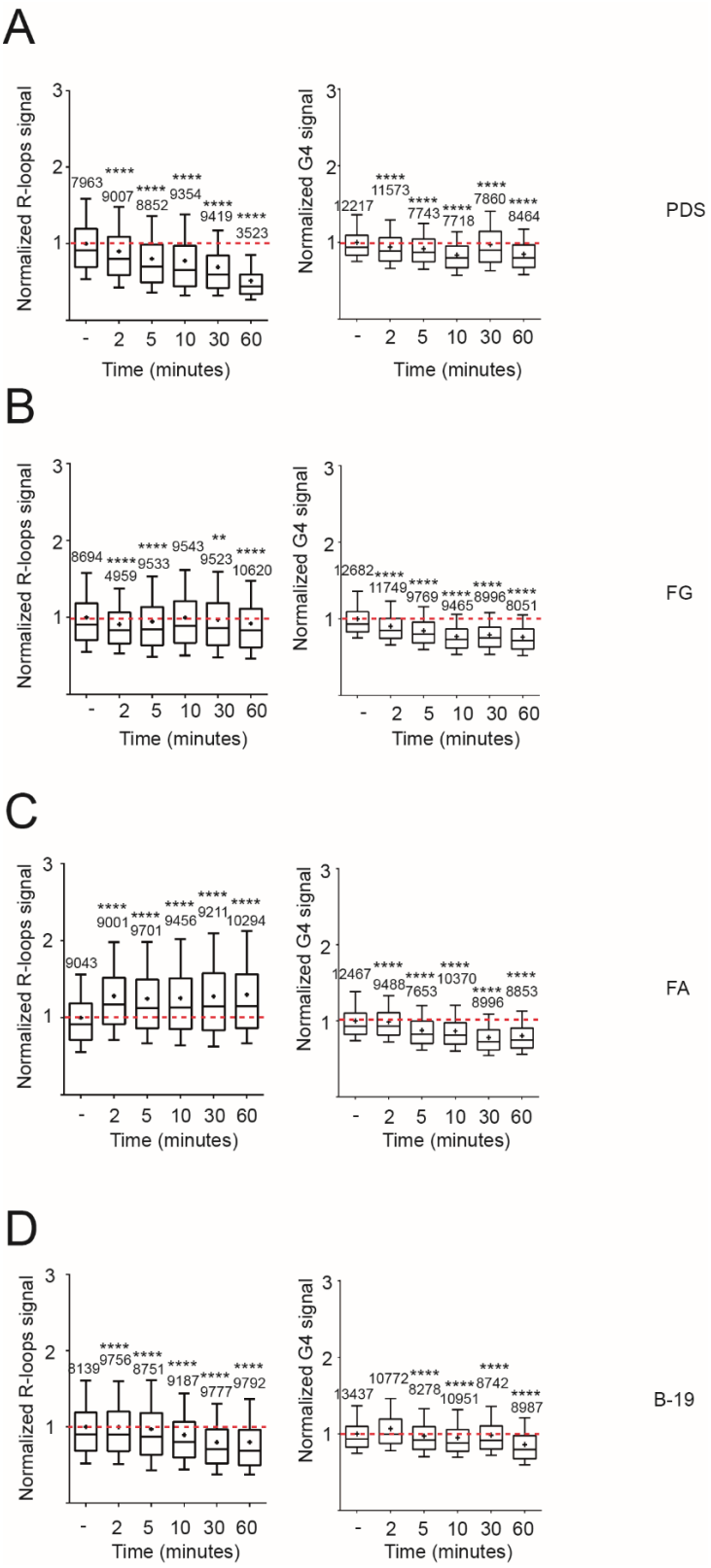


Figure 42. Quantification of S9.6 (left graph) and BG4 (right graph) signal in WI38 fibroblasts treated with 10 µM of PDS, FG, FA and 15 µM of Braco-19 for 2 to 60 minutes. The data were plotted after normalization over untreated control cells. The number of nuclei analyzed is reported for each sample on the bars of the graph. The * indicate the significance of the variation compared to the control sample. The number of * is directly correlate with the significant increase calculated with statistical parametric tests "Kolmogorov/Smirnov".

3.2.5 DNA damage and cell cycle arrest induced by G4 binders in human cancer cells

The above findings demonstrate that specific G4 ligands can stabilize not only G4s but also R loops in nuclear chromatin of human cancer cells. Thus, we next decided to understand the biological consequences of R loop induction. In particular, as G4 ligands are known to induce DNA damage and genome instability, we asked the question of whether this is due to increased levels of R loops.

First, we determined the γH2AX levels after cell treatments with G4 ligands by IF and western blotting. The results (Figure 43) show that γH2AX foci and phosphorylated protein content are increased in cells after treatment for 24 hours and the quantification of γH2AX foci intensity shows an increase in cells treated with 10 µM of PDS and FG, but not with FA. Moreover, the effects of G4 ligands, used at the same concentration, were different as the average number of PDS-induced foci is higher than that of FG.

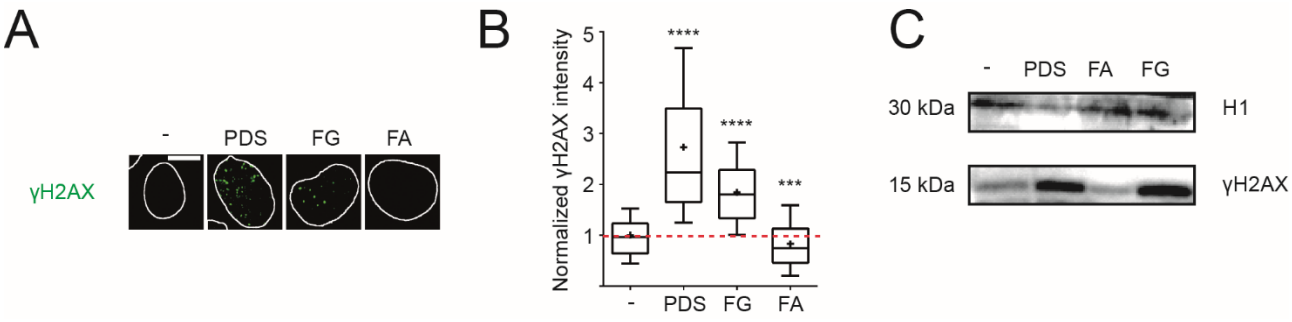


Figure 43. Elevated levels of γH2AX by G4 ligands by IF (A and B) and western blotting (C). U2OS cells were treated with 10 µM of PDS, FG or FA for 24 hours. Cells were then fixed and stained with a specific γH2AX antibody. In panel B, the γH2AX signal is reported after normalization over control untreated cells. 150 nuclei were selected, 75 for each esperiment. The * indicate the significance of the variation compared to the control sample. The number of * is directly correlate with the significant increase calculated with statistical parametric tests "Kolmogorov/Smirnov".

Subsequently, induction of γH2AX foci was determined at shorter times. Time course immunofluorescence experiment was conducted in Osteosarcoma U2OS cells treated with 10 µM of PDS, FG, FA and 15 µM of Braco-19 (Figure 44) .

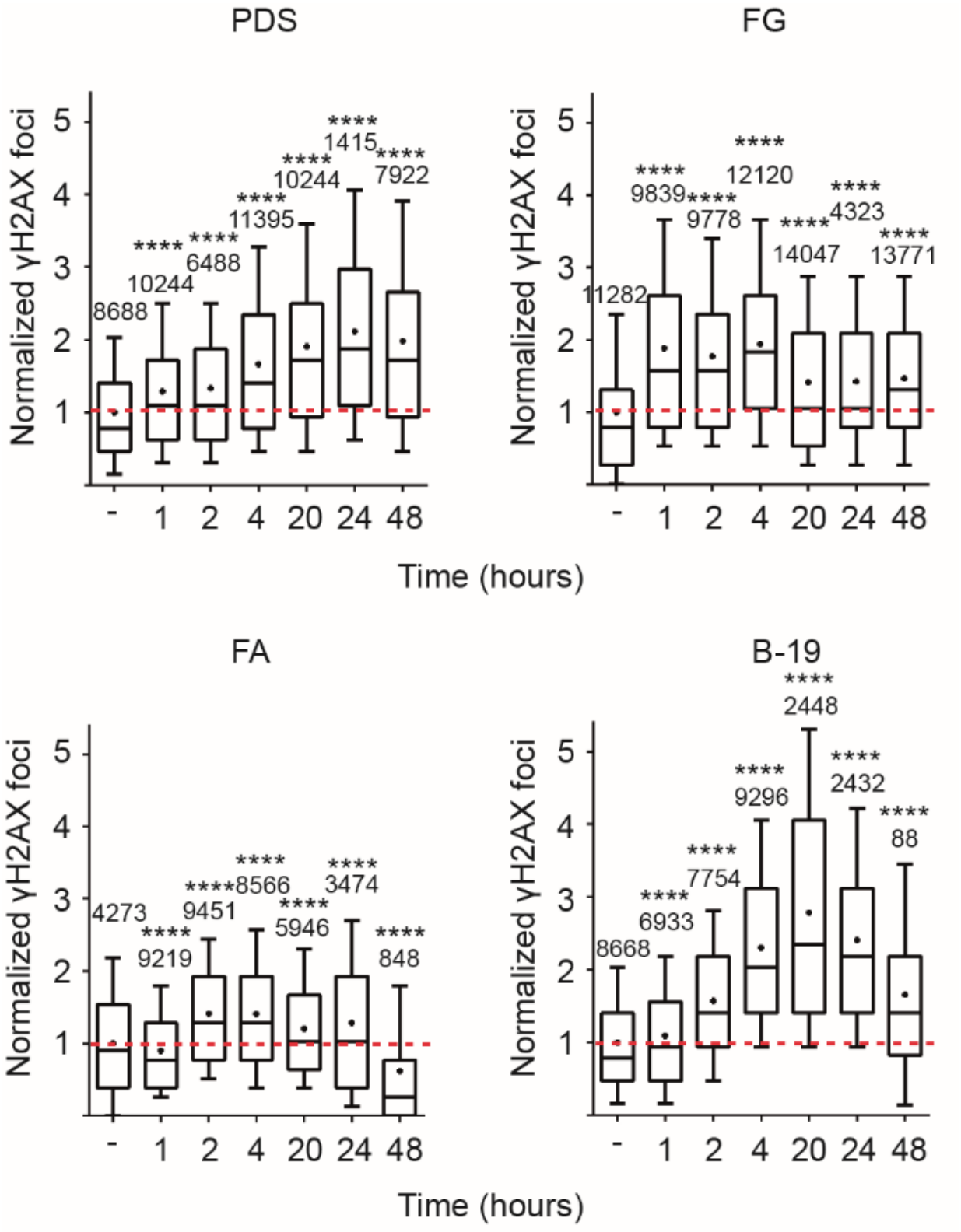


Figure 44. Quantification of γH2AX foci in time course immunofluorescence experiment on Osteosarcoma U2OS cells treated with 10 µM of PDS, FG, FA and 15µM of B-19. The data were plotted after normalization over untreated control cells. The * indicate the significance of the variation compared to the control sample. The number of nuclei analyzed is reported for each sample on the bars of the graph. The number of * is directly correlate with the significant increase calculated with statistical parametric tests "Kolmogorov/Smirnov".

The results show that γH2AX foci are increased at earlier time than 24 hours for all compounds, however with a kinetic somewhat different among the studied G4 ligands. The γH2AX foci linearly increased over time for PDS and Braco-19, while they show a peak between 1 and 4 hours followed by a tendency to decrease at longer times (Figure 44). FA was not able to significantly induce γH2AX foci. The data suggest that γH2AX foci formation follow the immediate increase of R loops in U2OS cells specifically by the studied G4 ligands (Figure 35).

Next, as G4 ligands did not stabilize G4 structures and R-loops in normal WI-38 fibroblast cells (Figure 42), we tested γH2AX foci formation in WI-38 fibroblasts treated for 24 hours with 10 µM of PDS, FG and FA (Figure 45a) and for short times with 10 µM of PDS, FG, FA and 15 µM of Braco-19 (Figure 45b). The results show a complete absence of γH2AX increase in these cells, supporting that R loop structures mediate DNA cleavage induced by G4 ligands.

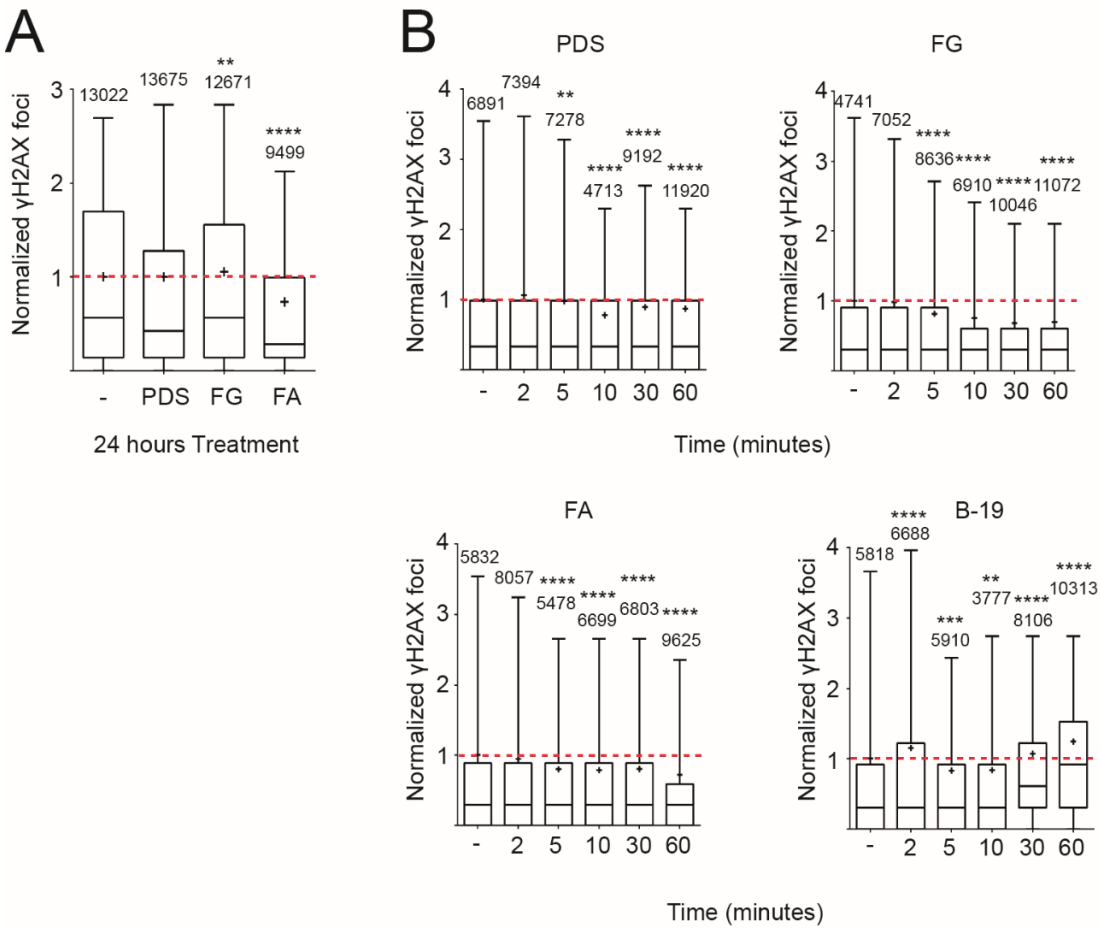


Figure 45. Quantification of γH2AX foci in time course immunofluorescence experiment on WI38 cells treated with 10 µM of PDS, FG, FA and 15µM of B-19. The data was plotted after control normalization. The data were plotted after normalization over untreated control cells. The number of nuclei analyzed is reported for each sample on the bars of the graph. The * indicate the significance of the variation compared to the control sample. The number of * is directly correlate with the significant increase calculated with statistical parametric tests "Kolmogorov/Smirnov".

Next, we examined cell cycle phase distribution by cytofluorimetry in 24-hour treated U2OS cells with 10 μM of PDS, FG, FA and 15 μM of Braco-19 (Figure 46). The results show an increase of the fraction of cells in G2/M phase, a hallmark of DNA damaging agents. However, the G4 ligands did not have the same effects. In particular, G2/M block was significant for PDS and B-19 while FG and FA show a less pronounced block of cells at G2/M.

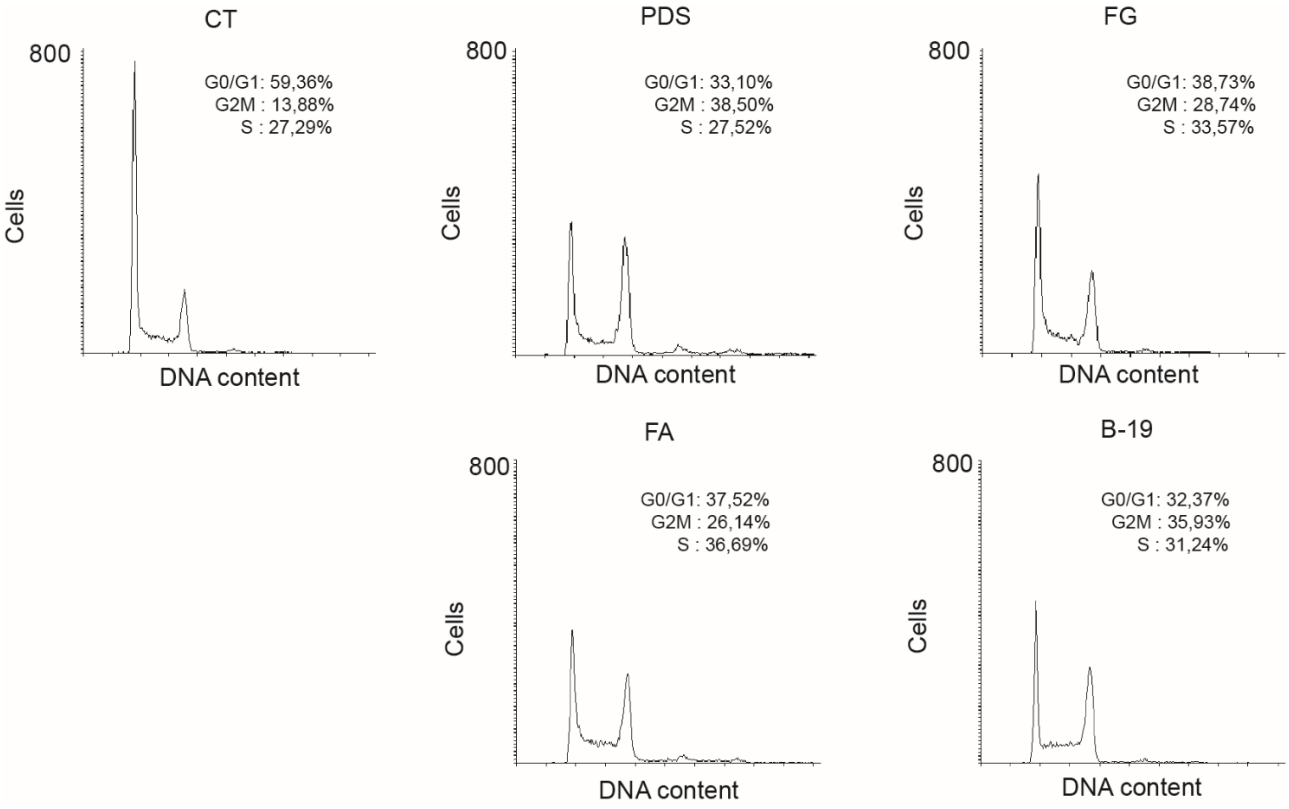


Figure 46. Cytofluorimetric analysis in Osteosarcoma U2OS cells treated for 24 hours with 10 μM of PDS, FG, FA and 15μM of Braco-19 and stained with Propidium Iodide (PI).

These results demonstrate that the studied G4 ligands induce γH2AX foci in human U2OS cancer cells immediately after R loop formation coupled to a cell cycle arrest at G2/M phase. Interestingly, the G2/M block correlated with γH2AX levels at 24 hours as FG induced a minor block at G2/M along with a reduction of γH2AX foci at the same time than PDS and B-19. Moreover, no γH2AX increase was observed in WI-38 normal fibroblasts where the G4 ligands did not induce R loops and G4 foci. Altogether, the findings support that DNA damage induced by G4 ligands can be mediated by an increase of R loops.

3.2.6 DNA damage checkpoint and homologous recombination are activated by G4 binders

Next, we determined whether or not G4 ligands activate DNA damage checkpoints. Therefore, we measured the activation ATM by IF and western blotting. The results (Figure 47) show that pATM is increased by PDS, and somewhat by FG, but not FA. The quantification shows an increase of pATM foci in cells treated with PDS (Figure 47b). In addition, pATM foci fully co-localize with γH2AX foci induced by PDS, suggesting that ATM was activate at specific chromatin regions where it phosphorylated H2AX around the site of DNA damage (Figure 47a). FG effects were much less marked than PDS in these assays.

We also tested another marker of DNA damage, 53BP1, which has also an important role in DSB repair (Zimmermann & De Lange, 2014). We determined by IF the phosphorylated (Figure 48a; 48b) and un-phophorylated (Figure 48c) forms of 53BP1 with specific antibodies. The quantification shows an increase of 53BP1 and p53BP1 foci in cells treated with PDS (Figure 48b, 48c) whereas FG and FA did not affect this marker after 24 hour treatments. Interestingly, we noted a nearly-perfect co-localization between p53BP1 and γH2AX foci induced by PDS (Figure 48a)

The results suggest that the studied G4 ligands have different effects on ATM and 53BP1 after 24 hour treatments. This may be related to different kinetics of γH2AX foci formation (Figure 44) determined in the same cell line. γH2AX induced by FG rapidly declined after 4 hours, whereas γH2AX induced by PDS and B-19 progressively increased up to 24 hours. The data may therefore suggest that activation of checkpoint is reversed more rapidly in the case of FG, in agreement with the lack of detection of pATM and P53BP1 after 24 hour of FG treatments. However, other explanations can be possible, such as differences in DNA repair kinetics, therefore more experiments are needed to clarify this point.

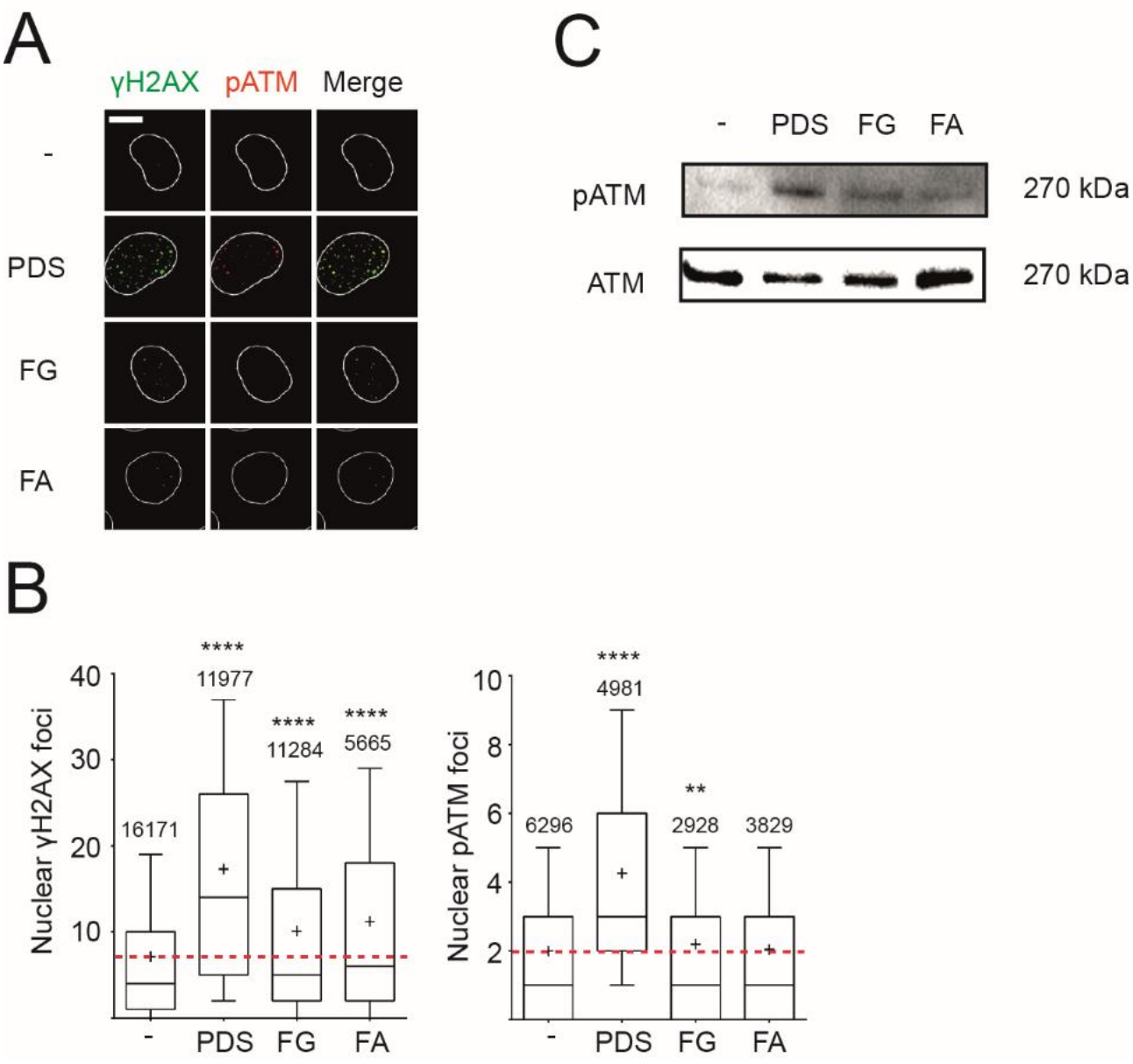


Figure 47. a) Immunofluorescence experiment on Osteosarcoma U2OS cells treated with 10 µM of PDS, FG and FA for 24 hours and co-labelled with pATM (red) and γH2AX (green) antibody; b) Quantification of γH2AX foci (at left) and pATM foci (at right). The number of nuclei analyzed is reported for each sample on the bars of the graph. The * indicate the significance of the variation compared to the control sample. The number of * is directly correlate with the significant increase calculated with statistical parametric tests "Kolmogorov/Smirnov". c) Western Blot experiment on U2OS treated with 10 µM of PDS, FG and FA for 24 hours and labelled with pATM and ATM (control) antibody.

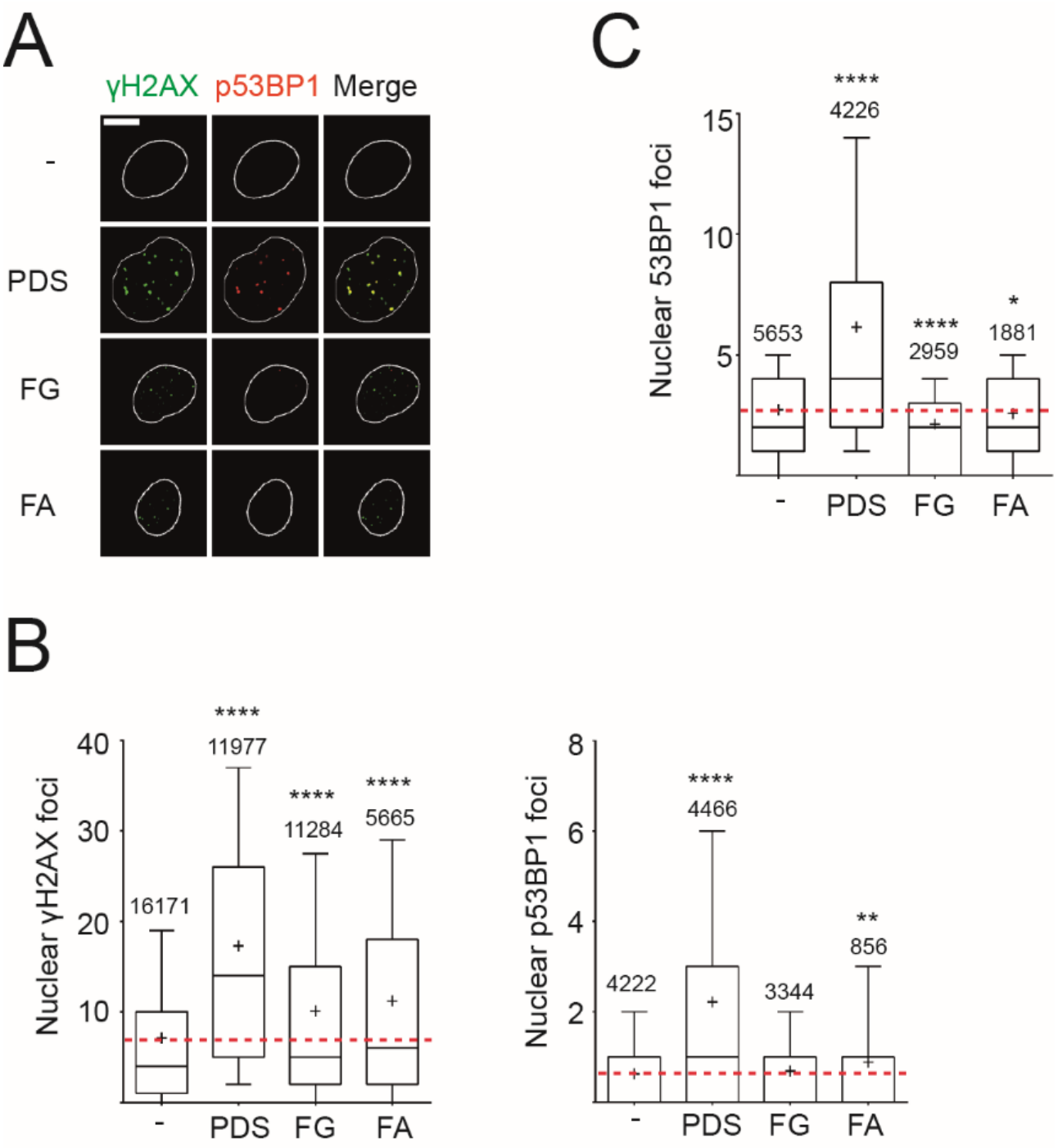


Figure 48. a) Immunofluorescence experiment on Osteosarcoma U2OS cells treated with 10 µM of PDS, FG and FA for 24 hours and co-labelled with p53BP1M (red) and γH2AX (green) antibody; b) Quantification of γH2AX foci (at left) and p53BP1 foci (at right). c) Quantification of 53BP1 foci. The number of nuclei analyzed is reported for each sample on the bars of the graph. The * indicate the significance of the variation compared to the control sample. The number of * is directly correlate with the significant increase calculated with statistical parametric tests "Kolmogorov/Smirnov".

In addition, we also assessed ATM and 53BP1 activation during different cell cycle phases. To this end, we performed a staining with Edu, a marker of S-phase cells. Edu staining associated with DNA content (measured with DAPI staining intensity by Operetta) and cell size can discriminate among G0/G1, S and G2/M phases allowing the grouping of cells in one phase of cell cycle. Three DNA damage markers, γH2AX, p53BP1 and pATM were evaluated in immunofluorescence after discrimination of the cell cycle phase (Figure 49).

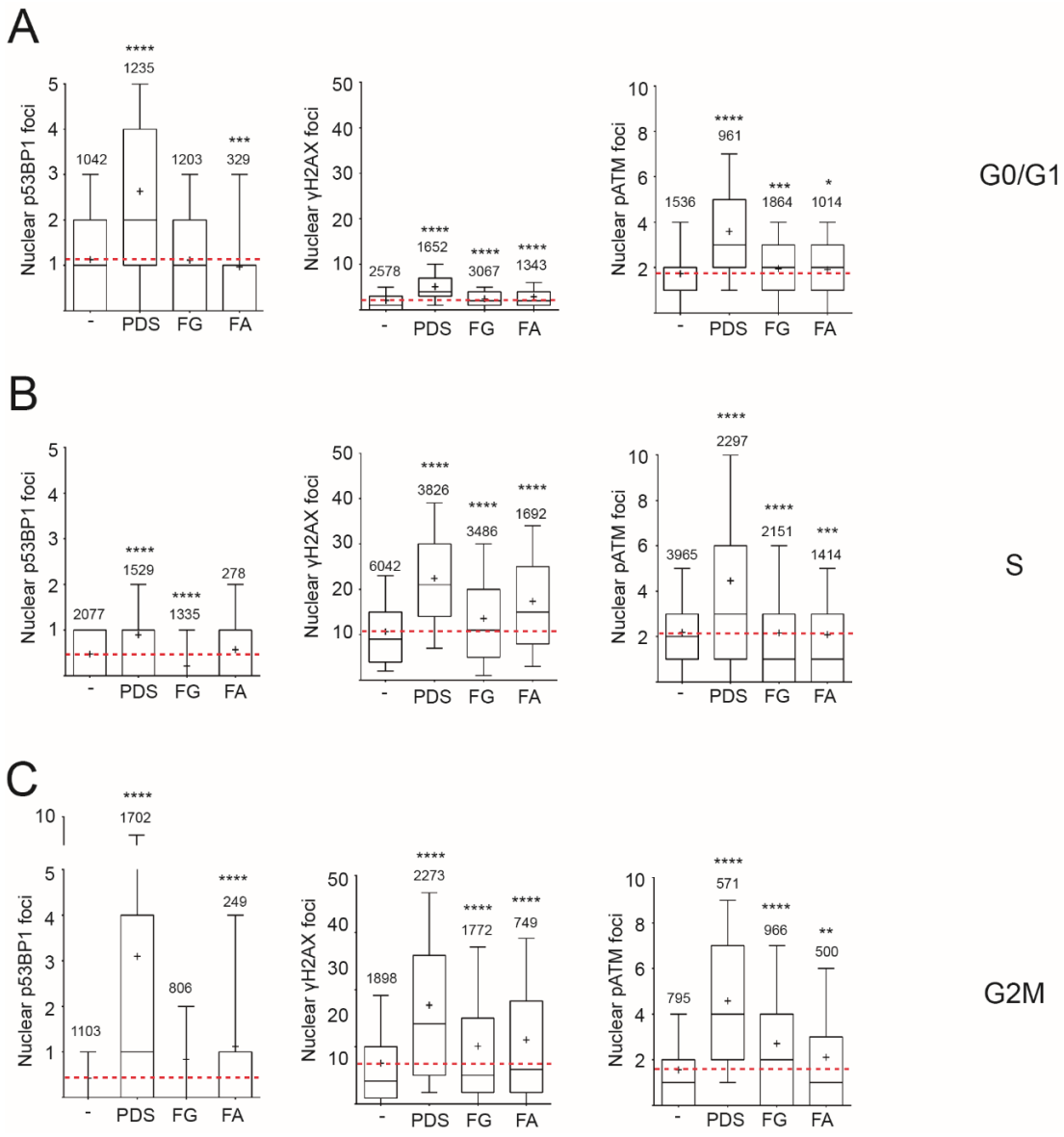


Figure 49. Immunofluorescence experiment on Osteosarcoma U2OS cells treated with 10 µM of PDS, FG, FA and 15 µM of Braco-19 and co-labelled with γH2AX, p53BP1 or pATM and EdU. a) Quantification of nuclear γH2AX, p53BP1 or pATM foci in G0/G1 phase; b) Quantification of nuclear γH2AX, p53BP1 or pATM foci in S phase; c) Quantification of nuclear γH2AX, p53BP1 or pATM foci in G2M phase. The number of nuclei analyzed is reported for each sample on the bars of the graph. The * indicate the significance of the variation compared to the control sample. The number of * is directly correlate with the significant increase calculated with statistical parametric tests "Kolmogorov/Smirnov".

The results show an increase of all the markers in all of the cell cycle phase after PDS treatment. FG could increase pATM and γH2AX in G2M phase only (Figure 46). Furthermore, for FA the results show only a small increase of γH2AX in G2M phase. The results suggest that FG induces an increase of DNA damage markers only in G2M phase, and probably this subpopulation is the one that increases in cytofluorimetric analysis (Figure 46).

Then, we asked whether G4 ligand-induced DNA cleavage was repaired by homologous recombination mechanisms, under our experimental conditions, as reported by other previously (Zimmer et al., 2016). Thus, we determined by IF the formation of RAD51 foci after 24 hours of treatment with 10 µM of PDS, FG or FA in U2OS cells. Rad51 is a critical factor involved in the essential strand invasion step of homologous recombination pathway (Kakarougkas & Jeggo, 2014).

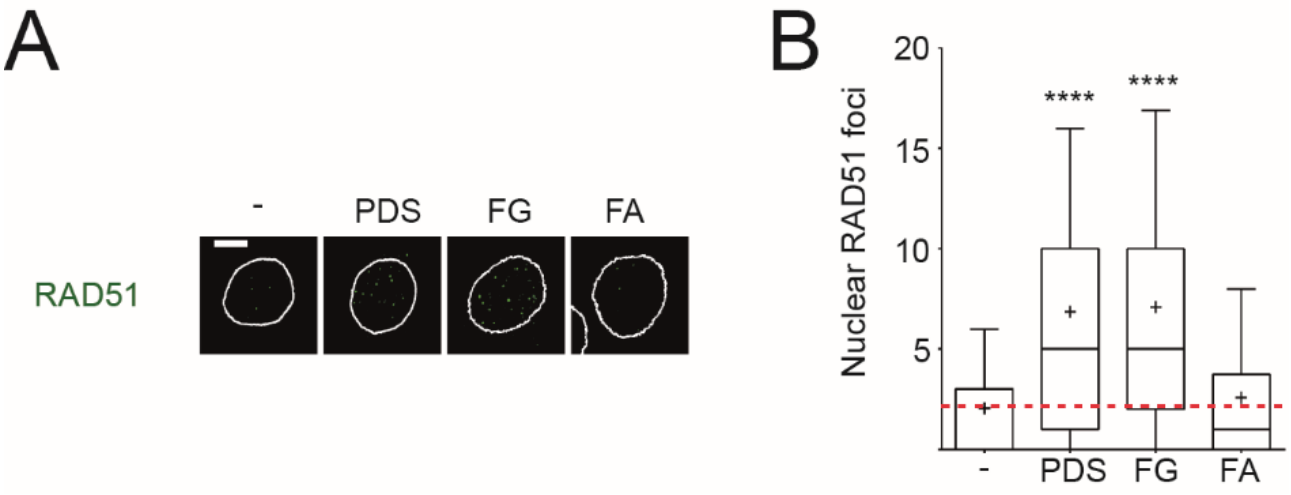


Figure 50. a) Immunofluorescence experiment on U2OS treated with 10 µM of PDS, FG and FA for 24 hours and labelled with RAD 51 antibody; b) Quantification of RAD 51 foci. 150 nuclei were selected, 75 for each esperiment. The * indicate the significance of the variation compared to the control sample. The number of * is directly correlate with the significant increase calculated with statistical parametric tests "Kolmogorov/Smirnov".

IF cell images and quantitative analyses show a clear increase of RAD51 foci in cells treated with PDS and FG but not FA (Figure 50b). Thus, the data demonstrate that PDS and FG, but not FA, trigger homologous recombination repair in U2OS cells suggesting that G4 ligand-induced DSBs are repaired, at least partially, by the homologous recombination repair pathway.

3.2.7 DNA damage induced by G4 binders is mediated by R loop increase

As G4 ligands induce an increase of R loops, the activation of DNA damage cell response and of homologous recombination repair, we wondered whether G4 ligand-induced R loops mediate the generation of DNA damage triggered by G4 ligands.

First, we attempted to establish whether or not R loops co-localize with chromatin regions of DSB and γH2AX by IF assays. We then perform an immunostaining with S9.6 and γH2AX antibodies in U2OS cells treated with 10 µM of PDS, FG and FA for 24 hours (Figure 51a).

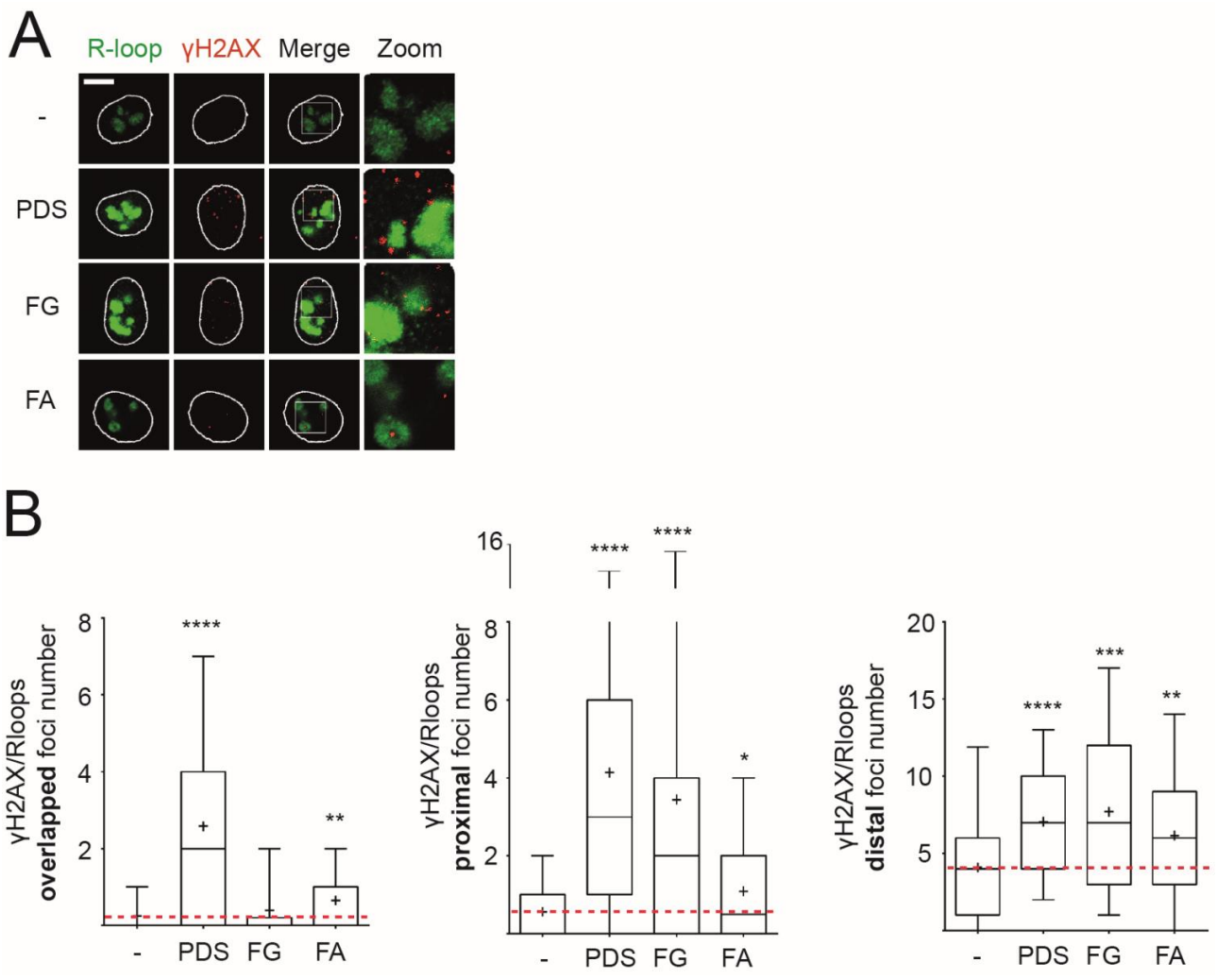


Figure 51. a) Immunofluorescence experiment on U2OS treated with 10 µM of PDS, FG and FA for 24 hours and co-labelled with S9.6 (green) and γH2AX (red) antibody; b) Each graph shows the average number of γH2AX foci that overlap, were in proximity and were distant with the S9.6 signal, respectively (from left to right). 100 nuclei were selected, 50 for each esperiment. The * indicate the significance of the variation compared to the control sample. The number of * is directly correlate with the significant increase calculated with statistical parametric tests "Kolmogorov/Smirnov".

The quantification of γH2AX foci that overlap with, were in proximity to and were distant from S9.6 signals shows that the overlap is minimal, if any, between γH2AX foci and R loops (Figure 51b) suggesting that DSBs are in different chromatin regions than R loops at 24 hours treatments. This may be consistent with a model in which DSBs are generated by the processing and removal of R loops by structure-specific endonuclease by (Sollier & Cimprich, 2015).

Next, we directly asked whether R loops are required for DSB production and γH2AX foci formation by using RNaseH1. Immunofluorescence experiments were performed in U2OS cells transfected with GFP-RNaseH1-expressing plasmid, then treated with 10 μM PDS for 24 hours and finally stained with γH2AX antibody. As a control, we transfected cells with the same plasmid containing and expressing a GFP gene only (Figure 52a).

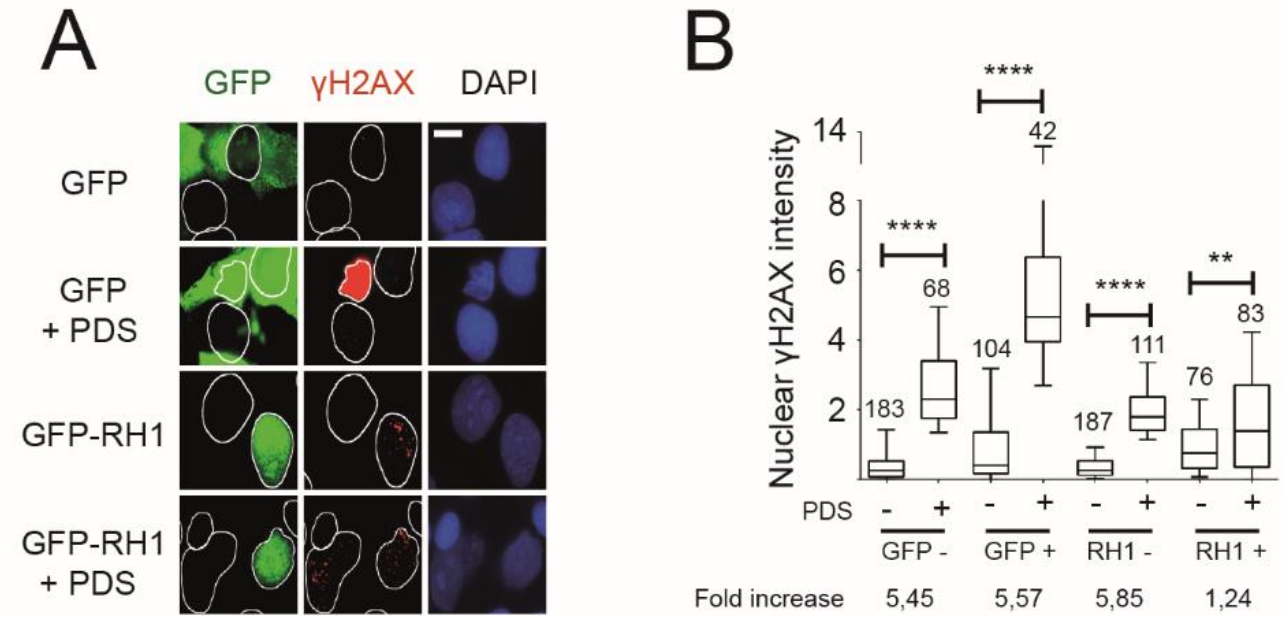


Figure 52. a) IF on U2OS transfected with transient GFP and GFP-RNaseH1, treated/untreated with 10 μM of PDS at 24 hours and labelled with γH2AX antibody; b) Quantification of γH2AX signal. The number of nuclei analyzed is reported for each sample on the bars of the graph. The * indicate the significance of the variation compared to the control sample. The number of * is directly correlate with the significant increase calculated with statistical parametric tests "Kolmogorov/Smirnov".

The results show an increase of γH2AX signal in cells positive and negative for GFP expression after PDS treatment. However, RNaseH1 positive cells showed a major reduction of γH2AX signal as compared with RNaseH1 negative cells as the fold increase from a value of 5.4-5.8 dropped to 1.24 (Figure 52b). These findings show that R loops are required for phosphorylation of H2AX and DNA damage induced by G4 ligands.

Similar experiments were then conducted with a cell line that stably carries an exogenous RNaseH1 gene under the control of a Tet promoter inducible with doxycycline (Figure 53a). The stable cell line expresses an RNaseH1 enzyme fused to an mCherry tag that allows an easy detection of the protein expression.

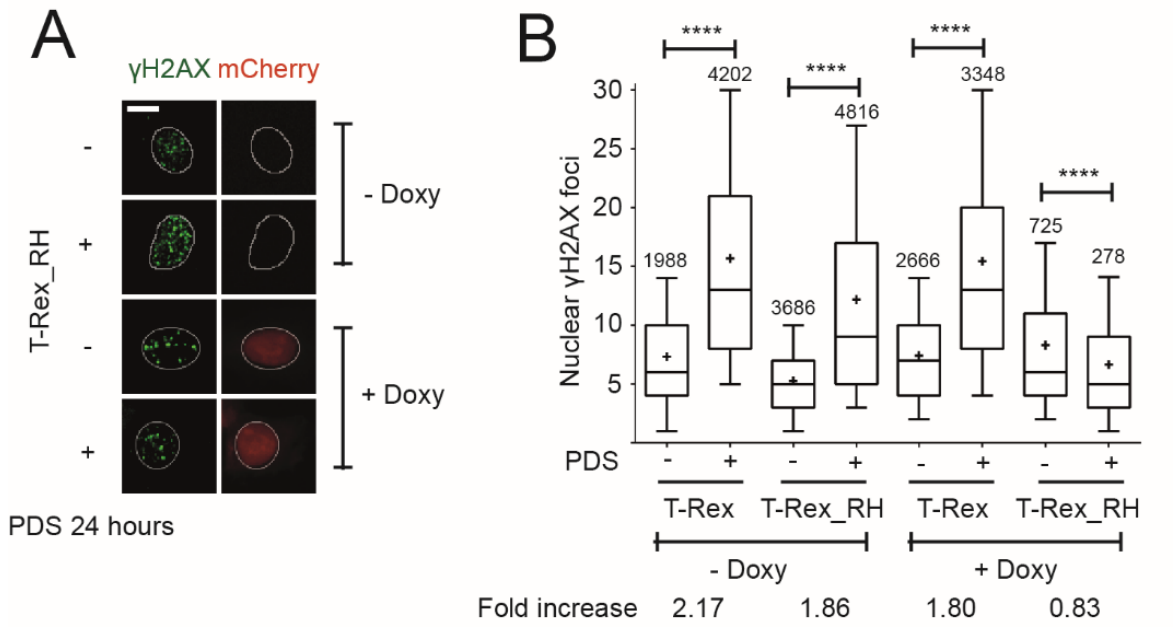


Figure 53. a) IF on U2OS_T-Rex_RH1, pre-treated/untreated with 2µg/µl of Doxycicline treated/untreated with 10 µM of PDS at 24 hours and labelled with γH2AX antibody; b) Quantification of γH2AX foci. The number of nuclei analyzed is reported for each sample on the bars of the graph. The * indicate the significance of the variation compared to the control sample. The number of * is directly correlate with the significant increase calculated with statistical parametric tests "Kolmogorov/Smirnov".

Quantification of γH2AX foci shows an increase of signal in control cells (T-Rex without RNaseH1 with and without Doxycicline) and stable RNaseH1 cells (T-Rex_RH) without Doxycicline after PDS treatment for 24 hours. T-Rex_RH cells treated with Doxycicline show a complete reduction of γH2AX signal as average foci number was comparable with that of cells untreated with PDS (Figure 53b). The T-Rex_RH cell line therefore show similar results with the RNaseH1 transient expression experiment in U2OS cells. In parallel, we assessed the expression of the RNaseH1 enzyme following doxycicline addition to the medium by following mCherry fluorescence (Figure 53a).

In these experiments, the data show lower values of foci number in the T_Rex-RH line not induced with Doxycicline as compared with the T-Rex control line. These data suggest a leaking expression activity present under these conditions. To confirm this hypothesis, an IF was performed in control cells labelled with S9.6 antibody (Figure 54a) and a Western Blot was conducted with an anti-mCherry antibody (the protein was linked to mCherry) to reveal the protein amount (Figure 54c).

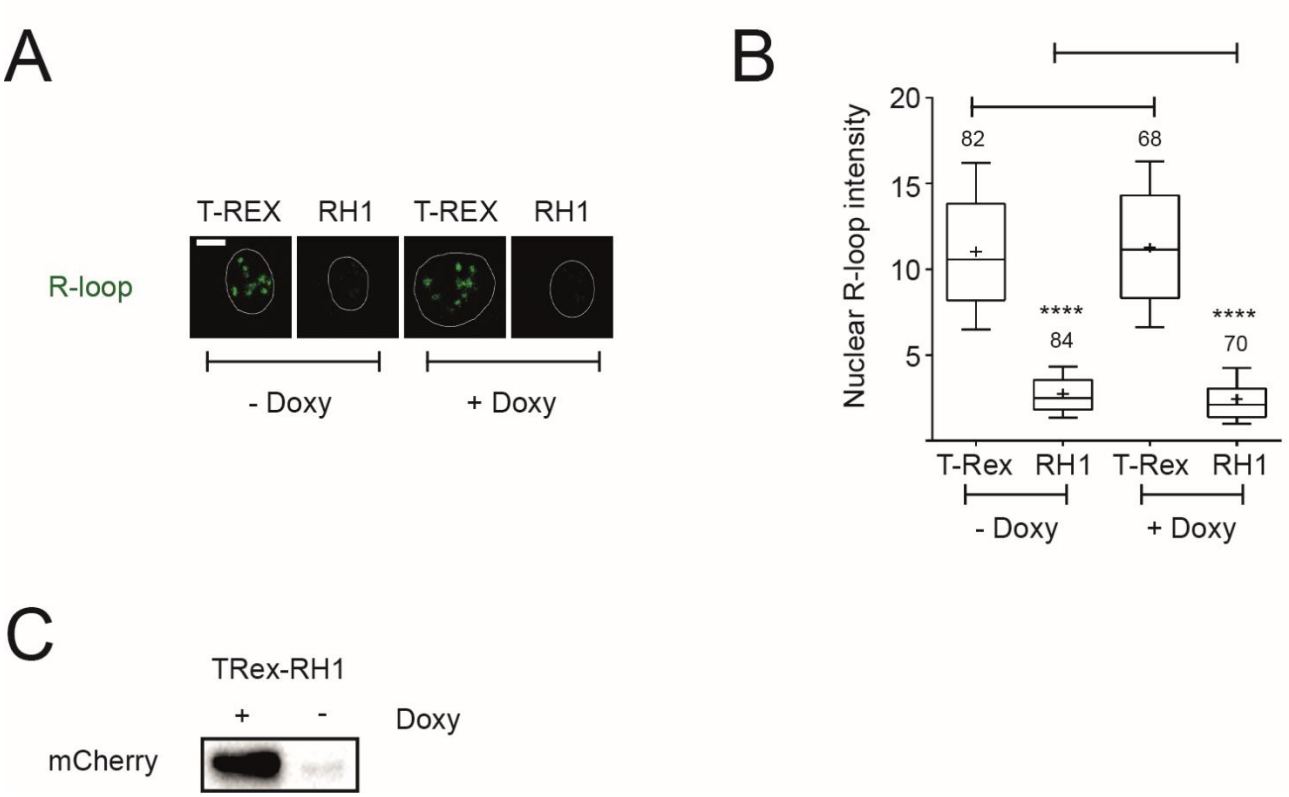


Figure 54. a) Immunofluorescence experiment on U2OS_T-rex and U2OS_T-rex_RH treated/untreated with 2 µg/µM of Doxycicline and labelled with S9.6 antibody; b) Quantification of S9.6 signal in cells induce (right side) and not induce (left side) with Doxycicline. The number of nuclei analyzed is reported for each sample on the bars of the graph. The * indicate the significance of the variation compared to the control sample. The number of * is directly correlate with the significant increase calculated with statistical parametric tests "Kolmogorov/Smirnov". c) Western Blot experiment on U2OS_T-rex_RH treated/untreated with 2 µg/µM of Doxycicline labelled with an anti-mCherry antibody.

The quantification of S9.6 signals shows that the amount of R loops in the RH1 inducible cells is much lower than the T-Rex line and almost comparable to the doxycycline-induced line (Figure 54b). At the same time, western blot results show a massive expression of mCherry in cells induced with Doxycicline and a faint band in those untreated with doxycycline (Figure 54c). The data are therefore consistent with a leaking expression activity of the promoter controlling the RNaseH1 gene that results in a decrease of R loop levels in cells at steady-state.

Altogether the findings show that reducing the R loops in cells prevents G4 ligands to induce γH2AX foci and likely DNA cleavage. This supports the hypothesis that G4 ligands stabilise not only G4, but also R loops that will then be processed by endonucleases thus generating DNA breakage and eventually leading to genome instability.

4. DISCUSSION

The results of the present PhD Thesis demonstrate that G4 ligands induce the formation of unscheduled R loop structures, along with G-quadruplex stabilization, in living cancer cells and that G4 ligand-induced DNA damage and genome instability are largely mediated by R loops. The involvement of R loops in the molecular and biological effects of the studied G4 ligands has been established by different methodological approaches such as immunofluorescence microscopy with a specific antibody, comparison with an inactive analog of a G4 ligand, overexpression of RNaseH1 in cells, and DRIP detection of R loops at specific genomic loci. All the data are consistent with a major role of R loops in biological effects of G4 ligands in human cancer cells.
Our results are in agreement with published findings in yeasts and *E.coli*. Recently, Yadav and colleagues using cassettes containing structures prone to G4 formation showed a physical-functional relationship between G4 and R-loops in yeast. They were able to determine a very close relationship between the structuring of these motifs and the topological structure of DNA. They have indeed highlighted an active role of Topoisomerase I in determining the conformational status of these structures (Yadav et al., 2014) (Yadav, Owiti, & Kim, 2016). The results of these works led to the elaboration of a model that predicts two possible ways of R loop-G4 reciprocal interaction in the genome (Yadav et al., 2014).

Moreover, our data are in agreement with a previous paper of Maizel's lab in which they show using electron microscopy that a loop region of 150-500 bp is produced by transcription of G-rich DNA in a plasmid (Duquette et al., 2004). The authors demonstrate that these co-transcriptional structures, called G-loops, contain G4s on the non template DNA strand and a stable RNA:DNA hybrid on the template strand. G-loops and G4 are efficiently formed within the plasmids when transcribed in vitro or in *E. coli* (Duquette et al., 2004).

In my Thesis work, a first step was the selection of molecules able to stabilize G4s *in vivo*. Two commercial molecules were chosen for their selectivity as G4 binders, such as Pyridostatin (PDS) and Braco-19, and some molecules synthesized in the Rambaldi's laboratory in Bologna (FG, FA and Compound 3). Immunofluorescence experiments were performed with a specific antibody against G4 structures. The antibody, called BG4, is a single chain antibody, purified from an expression plasmid kindly provided by S. Balasubramanian, Cambridge, UK (Biffi, Tannahill, et al., 2013), which has allowed the visualization of G4 structures in immunofluorescence microscopy. The results obtained from this assay showed a strong increase of G4s *in vivo* by PDS, and to a lesser extent by FG and Braco-19. FA, did not show any increase in the specific signal, confirming the data obtained with the

physical and chemical analyzes (Sparapani et al., 2010) (Amato et al., 2016) and making it an excellent negative compound for further *in vivo* studies. As far as compound 3 is concerned, it has a good G4 stabilization activity at low concentrations *in vivo*. However, its notable cytotoxicity makes it not comparable with PDS, B-19 and FG, which are markedly less cytotoxic in cultured cells. Cytotoxicity tests showed that FA has a good cytotoxic activity, but comparable to FG, most likely due to the presence of two aldehyde groups that can easily react with cellular macromolecules.

In a side project, in collaboration with A. Randazzo's group from University of Naples, we showed that the B3F10 compound can stabilize G4 in living cells. Physical and chemical analyses have also shown the capacity of the compound to stabilize both G-Triplex and G-Quadruplex structures (Amato et al., 2017). Our IF data showed a mild, but significant, stabilization of G4s in nuclei of cultured cancer cells. However, it was not possible to test the ability to stabilize G-Triplex structures *in vivo* due to the absence of an effective tool (Amato et al., 2017).

To investigate the role of R loops, IF experiments were performed with the R loop-specific monoclonal antibody called S9.6 that recognize DNA:RNA hybrid duplexes. Cells treated 24 hours with PDS, Braco-19 and FG showed an increase of S9.6 signals, whereas FA was fully ineffective. Furthermore, co-labeling with a nucleolus-specific marker, nucleolin, has shown that most of the increase is localized in the nucleoplasmic compartment, where the staining patterns show a well-spotted signal, not associated with nucleolar RNA, but with newly formed hybrids. Braco-19 after 24 hours produced a large fraction of apoptotic cells (not shown), and increased R loops were less evident and detected at 18 hours. This result can be interpreted with the specific ability of Braco-19 to stabilize G4s at the telomeric level where R loop might be less frequent. A strong increase in R loop signal in the nucleoplasm after 24 hours has been detected with 50 and 100 μM of Braco-19, but these concentrations are highly cytotoxic and therefore the result may be due to a high number of apoptotic cells.

Then, kinetics of formation of non-canonical DNA conformations were performed at short times, and the results obtained showed a bell-shaped trend in cells treated with PDS, FG and Braco-19. Again, FA did not have any detectable effect.The curves of R loops and G4s show the same trend and were in parallel to each other suggesting a contextual mechanism of formation/stabilization. The observations were particularly consistent when considering the nucleoplasm compartment only.

In addition to immunofluorescence experiments, DRIP was performed to measure R loop levels at selected loci with quantitative PCR (qPCR). The selected genomic regions were known to have significant levels of R loops (RPL13A, EGR1, EIF1α) or to lack any R loop (SNRPN, α-Sat) in

untreated cells. The results showed an increase of R loops at positive loci in cells treated with 10 μM of PDS and FG at short and long time of treatment. As IF and DRIP techniques have shown consistent findings, we propose that G4 stabilizers induce an increase of S9.6 signals in cultured cells. However, to assess the specificity of the Ab S9.6 signal, we asked the question of whether the fluorescence increase was affected by RNase H1. Cells transfected with a plasmid expressing an RNaseH1-GFP fusion protein showed a decrease of S9.6 signals in comparison to untreated cells. Therefore, we can affirm that the increase in the S9.6 signal is due to a stabilization of R loops by the tested G4 binders. Then, a further supporting finding was obtained by developing a co-labeling protocol with the two antibodies, S9.6 and BG4. Co-labeling results showed that a significant fraction of G4 spots ovelapped with R loop signals after treatment with PDS and FG at short and long times. Thus, these data indicate that R loops and G4s stabilized by the G4 ligand were close to each other, suggesting that the stabilized G4 can favor R loop formation in the same chromatin domain.

Published reports indicate that primary normal cell lines are less prone to form G4 structures than cancer cells (Rodriguez et al., 2012). In agreement with these results, human WI-38 fibroblasts did not show an increase of G4 foci and R-loop signals in cells treated with G4 stabilizers. Therefore, under our experimental conditions, the G4 stabilizers do not stabilize G4s, and consequently R-loops, in normal WI-38 fibroblasts.

The biological role of G4s is widely discussed in the literature. Here, we have asked ourselves the question of whether DNA damage are somehow mediated by R loop formation. The starting point was the confirmation that G4 ligands induce an increase of DNA damage, as detected in immunofluorescence experiments and western blots with anti γH2AX antibody. After 24 hours treatment with 10 μM of PDS, we detected an increase of γH2AX foci in agreement with previous reports (Rodriguez et al., 2012). Cells treated 24 hours with 10 μM of FG showed a moderate increase in the damage, compared with PDS. FA, despite of a higher cytotoxicity, did not show a strong effect on DNA damage.

The analysis of DNA damage response markers did not stop here, indeed other proteins involved in these processes were tested. Cells treated 24 hours with 10 μM of PDS and FG show a substantial increase in RAD51 foci, about 3 times compared to control. Furthermore, cells treated 24 hours with 10 μM of PDS show an increase of pATM and 53BP1 in the phosphorylated and non-phosphorylated forms, and the perfect co-localization between γH2AX foci and pATM/p53BP1 foci. In contrast, we detected only a modest but significant increase in pATM foci in 24-hour treated cells with 10 μM FG, confirmed by Western Blot analysis. No substantial increase of DDR markers was observed in cells treated with FA.

As R loops were induced immediately upon addition of the G4 ligand to the medium, we wondered whether DNA damage was generated at shorter times than 24 hours. Cells treated with 10 μM of PDS and 15 μM of Braco-19 has shown an increase in γH2AX foci starting at one hour and then increasing until the maximum peak at 24 hours. However, for FG the major increase was detected between 1 and 4 hours. For FA, only a slight increase was detected likely due to a cell killing mechanism different from G4/R loop stabilization. The observed difference in G4 ligand action seems to be very strong, in particular FG seems to have completely different effects as compared with PDS and B-19 compounds. Levels of γH2AX did not increase in fibroblast WI38 cells treated with 10 μM of PDS, FG, FA and 15 μM of Braco-19. These data further confirmed that the compounds induced DNA damage via stabilization of G4 and R loops. This is also strongly supported by RNase H1 expression that can rescue DNA damage in U2OS cells after PDS treatments.

Lastly, the analysis of the cell cycle revealed the importance of the cell replication state for DNA damage and checkpoint activation by FG. Firstly, cytofluorimetry experiments have shown an increase of the G2 phase in cells treated 24 hours with all compounds, in agreement with published findings. Furthermore, EdU staining has revealed a consistent increase of the three markers analyzed, (γH2AX, pATM and p53BP1) in all cell cycle phases for PDS whereas the increase was specific to G2M phase and less evident in the S phase for FG.

In conclusion, altogether these data represent the first demonstration in cell of a correlation between these two non-B DNA structures and of the role of R loops in DNA damage induced by G4 binders. We also showed for the first time that G4 stabilizers trigger DNA damage already after 1 hour of treatment immediately following R loop reduction. If DNA damage at 24 hours may be linked to G4 ligand activity as well as to activated cellular mechanisms, DNA damage at shorter times may be more directly related to stabilization of G4s and R loops. Furthermore, it also seems evident that the action of the studied compounds depends on the state of DNA replication and cell cycle phase.

At molecular levels, we propose a model of G4 ligand-induction of R loops in which the stabilization of the G4 on the non-template strand stabilizes an R loop on the template strand. The R loop may then eventually lead to DNA cleavage by processing factors (Figure 55).

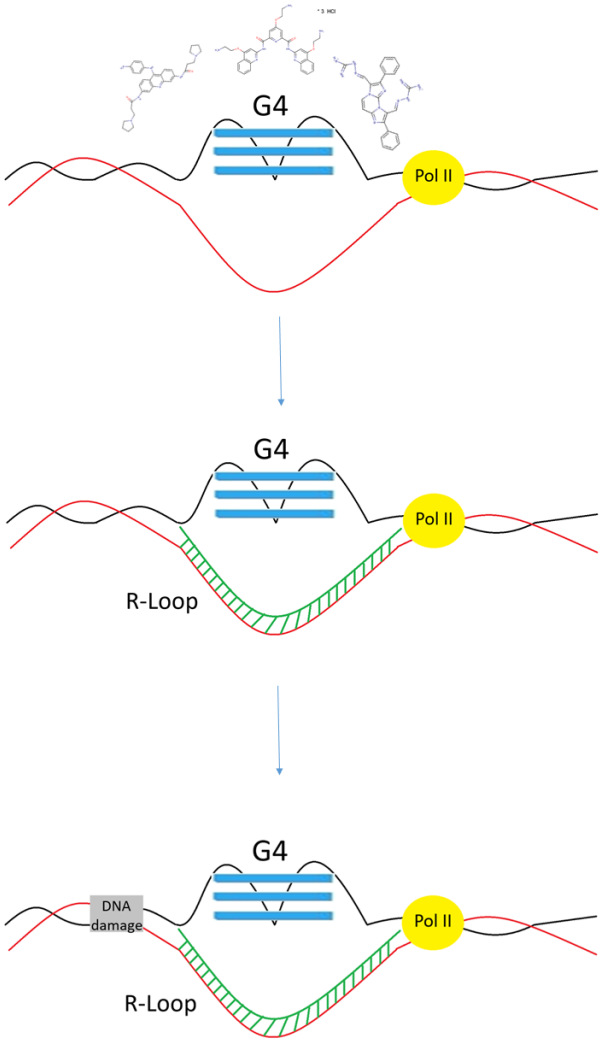


Figure 55. Schematic representation of correlation between G4 and R-Loop. During transcription treatment by G4 binders induced stabilization of G-Quadruplex, that induce the hybrid formation on the other strend. This structure thermodynamically more stable induce genomic instability and DNA damage.

5. BIBLIOGRAPHY

Aguilera, A., & Gómez-González, B. (2008). Genome instability: A mechanistic view of its causes and consequences. *Nature Reviews Genetics*, *9*(3), 204–217. https://doi.org/10.1038/nrg2268

Amato, J., Morigi, R., Pagano, B., Pagano, A., Ohnmacht, S., De Magis, A., … Randazzo, A. (2016). Toward the Development of Specific G-Quadruplex Binders: Synthesis, Biophysical, and Biological Studies of New Hydrazone Derivatives. *Journal of Medicinal Chemistry*, *59*(12), 5706–5720. https://doi.org/10.1021/acs.jmedchem.6b00129

Amato, J., Pagano, A., Cosconati, S., Amendola, G., Fotticchia, I., Iaccarino, N., … Randazzo, A. (2017). Discovery of the first dual G-triplex/G-quadruplex stabilizing compound: a new opportunity in the targeting of G-rich DNA structures? *Biochimica et Biophysica Acta - General Subjects*, *1861*(5), 1271–1280. https://doi.org/10.1016/j.bbagen.2016.11.008

Baaklini, I., Hraiky, C., Rallu, F., Tse-Dinh, Y. C., & Drolet, M. (2004). RNase HI overproduction is required for efficient full-length RNA synthesis in the absence of topoisomerase I in Escherichia coli. *Molecular Microbiology*, *54*(1), 198–211. https://doi.org/10.1111/j.1365-2958.2004.04258.x

Bedrat, A., Lacroix, L., & Mergny, J. L. (2016). Re-evaluation of G-quadruplex propensity with G4Hunter. *Nucleic Acids Research*, *44*(4), 1746–1759. https://doi.org/10.1093/nar/gkw006

Beletskii, A., & 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*, *93*(24), 13919–13924. https://doi.org/10.1073/pnas.93.24.13919

Bernstein, N. K., Hammel, M., Mani, R. S., Weinfeld, M., Pelikan, M., Tainer, J. A., & Glover, J. N. M. (2009). Mechanism of DNA substrate recognition by the mammalian DNA repair enzyme, Polynucleotide Kinase. *Nucleic Acids Research*, *37*(18), 6161–6173. https://doi.org/10.1093/nar/gkp597

Biffi, G., Di Antonio, M., Tannahill, D., & Balasubramanian, S. (2013). Visualization and selective chemical targeting of RNA G-quadruplex structures in the cytoplasm of human cells. *Nature Chemistry*, *6*(1), 75–80. https://doi.org/10.1038/nchem.1805

Biffi, G., Tannahill, D., McCafferty, J., & Balasubramanian, S. (2013). Quantitative visualization of DNA G-quadruplex structures in human cells. *Nature Chemistry*, *5*(3), 182–186.

https://doi.org/10.1038/nchem.1548

Bochman, M. L., Paeschke, K., & Zakian, V. A. (2012). DNA secondary structures: stability and function of G-quadruplex structures. *Nature Reviews Genetics*, *13*(11), 770–780. https://doi.org/10.1038/nrg3296

Boguslawski, S. J., Smith, D. E., Michalak, M. A., Mickelson, K. E., Yehle, C. O., Patterson, W. L., & Carrico, R. J. (1986). Characterization of monoclonal antibody to DNA · RNA and its application to immunodetection of hybrids. *Journal of Immunological Methods*, *89*(1), 123–130. https://doi.org/10.1016/0022-1759(86)90040-2

Burge, S., Parkinson, G. N., Hazel, P., Todd, A. K., & Neidle, S. (2006). Quadruplex DNA: Sequence, topology and structure. *Nucleic Acids Research*, *34*(19), 5402–5415. https://doi.org/10.1093/nar/gkl655

Capra, J. A., Paeschke, K., Singh, M., & Zakian, V. A. (2010). G-quadruplex DNA sequences are evolutionarily conserved and associated with distinct genomic features in Saccharomyces cerevisiae. *PLoS Computational Biology*, *6*(7), 9. https://doi.org/10.1371/journal.pcbi.1000861

Chakraborty, P., & Grosse, F. (2011). Human DHX9 helicase preferentially unwinds RNA-containing displacement loops (R-loops) and G-quadruplexes. *DNA Repair*, *10*(6), 654–665. https://doi.org/10.1016/j.dnarep.2011.04.013

Chambers, V. S., Marsico, G., Boutell, J. M., Di Antonio, M., Smith, G. P., & Balasubramanian, S. (2015). High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nature Biotechnology*, *33*(8), 877–881. https://doi.org/10.1038/nbt.3295

Chaudhuri, J., & Alt, F. W. (2004). ERRATUM: Class-switch recombination: interplay of transcription, DNA deamination and DNA repair. *Nature Reviews Immunology*, *4*(8), 655–655. https://doi.org/10.1038/nri1425

Cirincione, R., Di Maggio, F. M., Forte, G. I., Minafra, L., Bravatà, V., Castiglia, L., … Cammarata, F. P. (2017). High-Intensity Focused Ultrasound– and Radiation Therapy–Induced Immuno-Modulation: Comparison and Potential Opportunities. *Ultrasound in Medicine and Biology*, *43*(2), 398–411. https://doi.org/10.1016/j.ultrasmedbio.2016.09.020

Cottarel, J., Frit, P., Bombarde, O., Salles, B., Négrel, A., Bernard, S., … Calsou, P. (2013). A noncatalytic function of the ligation complex during nonhomologous end joining. *Journal of Cell Biology*, *200*(2), 173–186. https://doi.org/10.1083/jcb.201203128

Domínguez-Sánchez, M. S., Barroso, S., Gómez-González, B., Luna, R., & Aguilera, A. (2011). Genome instability and transcription elongation impairment in human cells depleted of THO/TREX. *PLoS Genetics*, *7*(12), 19–22. https://doi.org/10.1371/journal.pgen.1002386

Drolet, M., Phoenix, P., Menzel, R., Massé, E., Liu, L. F., & Crouch, R. J. (1995). Overexpression of RNase H partially complements the growth defect of an Escherichia coli delta topA mutant: R-loop formation is a major problem in the absence of DNA topoisomerase I. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(8), 3526–30. https://doi.org/10.1073/pnas.92.8.3526

Dson, J. P. R. I. A. (1975). Attachment of Nascent R N A Molecules to SuperheHcal D N A Indiana University, 566–579.

Duquette, M. L., Handa, P., Vincent, J. A., Taylor, A. F., & Maizels, N. (2004). Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. *Genes and Development*, *18*(13), 1618–1629. https://doi.org/10.1101/gad.1200804

El Hage, A., French, S. L., Beyer, A. L., & Tollervey, D. (2010). Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis. *Genes and Development*, *24*(14), 1546–1558. https://doi.org/10.1101/gad.573310

French, S. L., Sikes, M. L., Hontz, R. D., Osheim, Y. N., Lambert, T. E., El Hage, A., … Beyer, A. L. (2011). Distinguishing the Roles of Topoisomerases I and II in Relief of Transcription-Induced Torsional Stress in Yeast rRNA Genes. *Molecular and Cellular Biology*, *31*(3), 482–494. https://doi.org/10.1128/MCB.00589-10

Gan, W., Guan, Z., Liu, J., Gui, T., Shen, K., Manley, J. L., & Li, X. (2011). R-loop-mediated genomic instability is caused by impairment of replication fork progression. *Genes and Development*, *25*(19), 2041–2056. https://doi.org/10.1101/gad.17010011

Gellert, M., Lipsett, M. N., & Davies, D. R. (1962). Helix Formation By Guanylic Acid. *Proceedings of the National Academy of Sciences*, *48*(12), 2013–2018. https://doi.org/10.1073/pnas.48.12.2013

Gessner, R. V., Frederick, C. A., Quigley, G. J., Rich, A., & Wang, A. H. J. (1989). The molecular structure of the left-handed Z-DNA double helix at 1.0-?? atomic resolution. Geometry, conformation, and ionic interactions of d(CGCGCG). *Journal of Biological Chemistry*, *264*(14), 7921–7935.

Ginno, P. A., Lim, Y. W., Lott, P. L., Korf, I., & Chédin, F. (2013). GC skew at the 5’ and 3’ ends of human genes links R-loop formation to epigenetic regulation and transcription termination. TL - 23. *Genome Research*, *23 VN-r*(10), 1590–1600. https://doi.org/10.1101/gr.158436.113

Ginno, P. A., Lott, P. L., Christensen, H. C., Korf, I., & Chédin, F. (2012). R-Loop Formation Is a Distinctive Characteristic of Unmethylated Human CpG Island Promoters. *Molecular Cell*, *45*(6), 814–825. https://doi.org/10.1016/j.molcel.2012.01.017

Gomez-Gonzalez, B., & Aguilera, A. (2007). Activation-induced cytidine deaminase action is strongly stimulated by mutations of the THO complex. *Proceedings of the National Academy of Sciences*, *104*(20), 8409–8414. https://doi.org/10.1073/pnas.0702836104

Gray, L. T., Vallur, A. C., Eddy, J., & Maizels, N. (2014). G quadruplexes are genomewide targets of transcriptional helicases XPB and XPD. *Nature Chemical Biology*, *10*(4), 313–318. https://doi.org/10.1038/nchembio.1475

Hamperl, S., Bocek, M. J., Saldivar, J. C., Swigut, T., & Cimprich, K. A. (2017). Transcription-Replication Conflict Orientation Modulates R-Loop Levels and Activates Distinct DNA Damage Responses. *Cell*, *170*(4), 774–786.e19. https://doi.org/10.1016/j.cell.2017.07.043

Hänsel-Hertsch, R., Beraldi, D., Lensing, S. V, Marsico, G., Zyner, K., Parry, A., … Balasubramanian, S. (2016). G-quadruplex structures mark human regulatory chromatin. *Nature Genetics*, *48*(10), 1267–1272. https://doi.org/10.1038/ng.3662

Hartono, S. R., Korf, I. F., & Chédin, F. (2015). GC skew is a conserved property of unmethylated CpG island promoters across vertebrates. *Nucleic Acids Research*, *43*(20), 9729–9741. https://doi.org/10.1093/nar/gkv811

Huertas, P., & Aguilera, A. (2003). Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. *Molecular Cell*, *12*(3), 711–721. https://doi.org/10.1016/j.molcel.2003.08.010

Huppert, J. L., & Balasubramanian, S. (2007). G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Research*, *35*(2), 406–413. https://doi.org/10.1093/nar/gkl1057

Jeggo, P. A. (n.d.). in DNA Double-Strand Break Repair, *2*, 317–318.

Kakarougkas, A., & Jeggo, P. A. (2014). DNA DSB repair pathway choice: An orchestrated handover mechanism. *British Journal of Radiology*, *87*(1035). https://doi.org/10.1259/bjr.20130685

Massé, E., Phoenix, P., & Drolet, M. (1997). DNA topoisomerases regulate R-loop formation during transcription of the rrnB operon in Escherichia coli. *Journal of Biological Chemistry*, *272*(19), 12816–12823. https://doi.org/10.1074/jbc.272.19.12816

Mischo, H. E., Gómez-González, B., Grzechnik, P., Rondón, A. G., Wei, W., Steinmetz, L., … Proudfoot, N. J. (2011). Yeast Sen1 helicase protects the genome from transcription-associated instability. *Molecular Cell*, *41*(1), 21–32. https://doi.org/10.1016/j.molcel.2010.12.007

Muers, M. (2011). Mutation: the perils of transcription. *Nature Reviews. Genetics*, *12*(3), 156. https://doi.org/10.1038/nrg2960

Mueser, T. C., Hinerman, J. M., Devos, J. M., Boyer, R. A., & Williams, K. J. (2010). Structural analysis of bacteriophage T4 DNA replication: a review in the Virology Journal series on bacteriophage T4 and its relatives. *Virology Journal*, *7*(1), 359. https://doi.org/10.1186/1743-422X-7-359

Neidle, S. (2015). A Personal History of Quadruplex-Small Molecule Targeting. *Chemical Record*, *15*(4), 691–710. https://doi.org/10.1002/tcr.201500011

Paleček, E. (1991). Local Supercoil-Stabilized DNA Structure. *Critical Reviews in Biochemistry and Molecular Biology*, *26*(2), 151–226. https://doi.org/10.3109/10409239109081126

Pellegrini, L., Yu, D. S., Lo, T., Anand, S., Lee, M., Blundell, T. L., & Venkitaraman, A. R. (2002). Insights into DNA recombination from the structure of a RAD51–BRCA2 complex. *Nature*, *420*(6913), 287–293. https://doi.org/10.1038/nature01230

Petersen-Mahrt, S. K., Harris, R. S., & Neuberger, M. S. (2002). AID mutates E. coli suggesting a DNA deamination mechanism for antibody diversification. *Nature*, *418*(6893), 99–103. https://doi.org/10.1038/nature00862

Phoenix, P., Raymond, M. A., Massé, É., & Drolet, M. (1997). Roles of DNA topoisomerases in the regulation of R-loop formation in vitro. *Journal of Biological Chemistry*, *272*(3), 1473–1479. https://doi.org/10.1074/jbc.272.3.1473

Reaban, M. E., Lebowitz, J., & Griffin, J. A. (1994). Transcription induces the formation of a stable RNA·DNA hybrid in the immunoglobulin α switch region. *Journal of Biological Chemistry*, *269*(34), 21850–21857.

Read, M., Harrison, R. J., Romagnoli, B., Tanious, F. A., Gowan, S. H., Reszka, A. P., … Neidle, S. (2001). Structure-based design of selective and potent G quadruplex-mediated telomerase

inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(9), 4844–9. https://doi.org/10.1073/pnas.081560598

Riballo, E., Kühne, M., Rief, N., Doherty, A., Smith, G. C. M., Recio, M. J., … Löbrich, M. (2004). A pathway of double-strand break rejoining dependent upon ATM, Artemis, and proteins locating to ??-H2AX foci. *Molecular Cell*, *16*(5), 715–724. https://doi.org/10.1016/j.molcel.2004.10.029

Roberts, R. W., & Crothers, D. M. (1992). Stability and Properties of Double and Triple Helices : Dramatic Effects of RNA or DNA Backbone Composition Author (s): Richard W . Roberts and Donald M . Crothers Published by : American Association for the Advancement of Science Stable URL : http://ww, *258*(5087), 1463–1466.

Rodriguez, R., Miller, K. M. K., Forment, J. V, Bradshaw, C. R., Nikan, M., Britton, S., … Jackson, S. P. (2012). Small molecule-induced DNA damage identifies alternative DNA structures in human genes. *Nature Chemical …*, *8*(3), 301–310. https://doi.org/10.1038/nchembio.780.Small

Rodriguez, R., Müller, S., Yeoman, J. A., Trentesaux, C., Riou, J. F., & Balasubramanian, S. (2008). A novel small molecule that alters shelterin integrity and triggers a DNA-damage response at telomeres. *Journal of the American Chemical Society*, *130*(47), 15758–15759. https://doi.org/10.1021/ja805615w

Sahakyan, A. B., Chambers, V. S., Marsico, G., Santner, T., Di Antonio, M., & Balasubramanian, S. (2017). Machine learning model for sequence-driven DNA G-quadruplex formation. *Scientific Reports*, *7*(1), 1–11. https://doi.org/10.1038/s41598-017-14017-4

Sauer, M., & Paeschke, K. (2017). G-quadruplex unwinding helicases and their function *in vivo*. *Biochemical Society Transactions*, BST20170097. https://doi.org/10.1042/BST20170097

Schiavone, D., Jozwiakowski, S. K., Romanello, M., Guilbaud, G., Guilliam, T. A., Bailey, L. J., … Doherty, A. J. (2016). PrimPol Is Required for Replicative Tolerance of G Quadruplexes in Vertebrate Cells. *Molecular Cell*, *61*(1), 161–169. https://doi.org/10.1016/j.molcel.2015.10.038

Skourti-Stathaki, K., Kamieniarz-Gdula, K., & Proudfoot, N. J. (2014). R-loops induce repressive chromatin marks over mammalian gene terminators. *Nature*, *516*(7531), 436–439. https://doi.org/10.1038/nature13787

Skourti-Stathaki, K., & Proudfoot, N. J. (2014). A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression. *Genes and Development*, *28*(13), 1384–1396. https://doi.org/10.1101/gad.242990.114

Skourti-Stathaki, K., Proudfoot, N. J., & Gromak, N. (2011). Human Senataxin Resolves RNA/DNA Hybrids Formed at Transcriptional Pause Sites to Promote Xrn2-Dependent Termination. *Molecular Cell*, *42*(6), 794–805. https://doi.org/10.1016/j.molcel.2011.04.026

Sollier, J., & Cimprich, K. A. (2015). Breaking bad: R-loops and genome integrity. *Trends in Cell Biology*, *25*(9), 514–522. https://doi.org/10.1016/j.tcb.2015.05.003

Sollier, J., Stork, C. T., García-Rubio, M. L., Paulsen, R. D., Aguilera, A., & Cimprich, K. A. (2014). Transcription-Coupled Nucleotide Excision Repair Factors Promote R-Loop-Induced Genome Instability. *Molecular Cell*, *56*(6), 777–785. https://doi.org/10.1016/j.molcel.2014.10.020

Sparapani, S., Bellini, S., Gunaratnam, M., Haider, S. M., Andreani, A., Rambaldi, M., … Neidle, S. (2010). Bis-guanylhydrazone diimidazo[1,2-a:1,2-c]pyrimidine as a novel and specific G-quadruplex binding motif. *Chemical Communications*, *46*(31), 5680–5682. https://doi.org/10.1039/c0cc00020e

Sun, Q., Csorba, T., Skourti-Stathaki, K., Proudfoot, N. J., & Dean, C. (2013). R-Loop Stabilization Represses Antisense Transcription at the Arabidopsis FLC Locus. *Science*, *340*(6132), 619–621. https://doi.org/10.1126/science.1234848

Symington, L. S., & Gautier, J. (2011). Double-Strand Break End Resection and Repair Pathway Choice. *Annual Review of Genetics*, *45*(1), 247–271. https://doi.org/10.1146/annurev-genet-110410-132435

Thomas, M., White, R. L., & Davis, R. W. (1976). Hybridization of RNA to double-stranded DNA: formation of R-loops. *Proc Natl Acad Sci U S A*, *73*(7), 2294–2298. https://doi.org/10.1073/pnas.73.7.2294

Todd, A. K., Johnston, M., & Neidle, S. (2005). Highly prevalent putative quadruplex sequence motifs in human DNA. *Nucleic Acids Research*, *33*(9), 2901–2907. https://doi.org/10.1093/nar/gki553

Tuduri, S., Crabbé, L., Conti, C., Tourrière, H., Holtgreve-Grez, H., Jauch, A., … Pasero, P. (2009). Topoisomerase I suppresses genomic instability by preventing interference between replication

and transcription. *Nature Cell Biology*, *11*(11), 1315–1324. https://doi.org/10.1038/ncb1984

Wahba, L., Amon, J. D., Koshland, D., & Vuica-ross, M. (2011). Article RNase H and Multiple RNA Biogenesis Factors Cooperate to Prevent RNA : DNA Hybrids from Generating Genome Instability. *Molecular Cell*, *44*(6), 978–988. https://doi.org/10.1016/j.molcel.2011.10.017

Watson, J. D., & Crick, F. H. C. (1953). Molecular structure of nucleic acids. *Nature*. https://doi.org/10.1097/BLO.0b013e3181468780

Woodford, K. J., Howell, R. M., & Usdin, K. (1994). A novel K+-dependent DNA synthesis arrest site in a commonly occurring sequence motif in eukaryotes. *Journal of Biological Chemistry*, *269*(43), 27029–27035. https://doi.org/10.4061/2010/564946

Xiao, S., Zhang, J. Y., Zheng, K. W., Hao, Y. H., & Tan, Z. (2013). Bioinformatic analysis reveals an evolutional selection for DNA:RNA hybrid G-quadruplex structures as putative transcription regulatory elements in warm-blooded animals. *Nucleic Acids Research*, *41*(22), 10379–10390. https://doi.org/10.1093/nar/gkt781

Xu1, B., & Clayton2, D. A. (1996). RNA-DNA hybrid formation at the human mitochondrial heavy-strand origin ceases at replication start sites: an implication for RNA-DNA hybrids serving as primers. *The EMBO Journal*, *15*(12), 3135–3143.

Yadav, P., Harcy, V., Argueso, J. L., Dominska, M., Jinks-Robertson, S., & Kim, N. (2014). Topoisomerase I Plays a Critical Role in Suppressing Genome Instability at a Highly Transcribed G-Quadruplex-Forming Sequence. *PLoS Genetics*, *10*(12). https://doi.org/10.1371/journal.pgen.1004839

Yadav, P., Owiti, N., & Kim, N. (2016). The role of topoisomerase i in suppressing genome instability associated with a highly transcribed guanine-rich sequence is not restricted to preventing RNA:DNA hybrid accumulation. *Nucleic Acids Research*, *44*(2), 718–729. https://doi.org/10.1093/nar/gkv1152

Yu, K., Chedin, F., Hsieh, C.-L., Wilson, T. E., & Lieber, M. R. (2003). R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nature Immunology*, *4*(5), 442–451. https://doi.org/10.1038/ni919

Zimmer, J., Tacconi, E. M. C., Folio, C., Badie, S., Porru, M., Klare, K., … Tarsounas, M. (2016). Targeting BRCA1 and BRCA2 Deficiencies with G-Quadruplex-Interacting Compounds. *Molecular Cell*, *61*(3), 449–460. https://doi.org/10.1016/j.molcel.2015.12.004

Zimmermann, M., & De Lange, T. (2014). 53BP1: Pro choice in DNA repair. *Trends in Cell Biology*, *24*(2), 108–117. https://doi.org/10.1016/j.tcb.2013.09.003