

5-2022

## Investigation of Genomic Instability Induced by G-Quadruplexes in the Absence of Functional Topoisomerase 1 in Yeast

Alexandra Berroyer

Follow this and additional works at: [https://digitalcommons.library.tmc.edu/utgsbs\\_dissertations](https://digitalcommons.library.tmc.edu/utgsbs_dissertations)



Part of the [Biology Commons](#), [Medicine and Health Sciences Commons](#), and the [Microbiology Commons](#)

---

### Recommended Citation

Berroyer, Alexandra, "Investigation of Genomic Instability Induced by G-Quadruplexes in the Absence of Functional Topoisomerase 1 in Yeast" (2022). *The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences Dissertations and Theses (Open Access)*. 1165. [https://digitalcommons.library.tmc.edu/utgsbs\\_dissertations/1165](https://digitalcommons.library.tmc.edu/utgsbs_dissertations/1165)

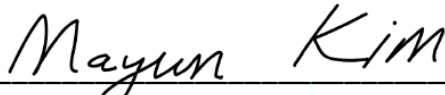
This Dissertation (PhD) is brought to you for free and open access by the The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences at DigitalCommons@TMC. It has been accepted for inclusion in The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences Dissertations and Theses (Open Access) by an authorized administrator of DigitalCommons@TMC. For more information, please contact [digitalcommons@library.tmc.edu](mailto:digitalcommons@library.tmc.edu).

**Investigation of Genomic Instability Induced by G-Quadruplexes in the Absence of  
Functional Topoisomerase 1 in Yeast**

by

*Alexandra Noel Berroyer, B. S., M. S.*

APPROVED:



Nayun Kim, Ph.D.

Advisory Professor



Ambro van Hoof, Ph.D.



Anne-Marie Krachler, Ph.D.



Margarida Albuquerque Almeida Santos, Ph.D.



Digitally signed by Grzegorz Ira  
DN: cn=Grzegorz Ira, o, ou,  
email=gira@bcm.edu, c=US  
Date: 2022.03.11 12:54:16 -06'00'

Grzegorz Ira, Ph.D.

---

APPROVED:

---

Dean, The University of Texas  
MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences

**Investigation of Genomic Instability Induced by G-Quadruplexes in the Absence of  
Functional Topoisomerase 1 in Yeast**

A

Dissertation

Presented to the Faculty of

The University of Texas

MD Anderson Cancer Center UTHealth

Graduate School of Biomedical Sciences

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

by

Alexandra Noel Berroyer, B. S., M. S.  
Houston, Texas

May 2022

To my friends and family;

Thank you for all the love and support I have received from you throughout my life.

## Acknowledgements

Completing graduate school has been a long journey, and I am grateful for many people for their support along the way as I approach the end of my voyage. First, I would like to thank my Ph.D. mentor Nayun Kim. When I first met Nayun during my Master's degree at Illinois State University (ISU) she encouraged me to apply to Ph.D. graduate programs after we had a conversation about my current research project. Further conversations with Nayun during her visit to ISU led me to confide in her that, I was nervous to apply and was questioning my abilities to complete a Ph.D. program. At that moment, Nayun told me about imposter syndrome. This was the first time I had ever heard of the term, and Nayun's encouragement and sincere willingness to support me gave me the confidence to go for it! So I applied to UT GSBS in Houston and ended up in her lab to earn a Ph.D., and I am extremely grateful for all Nayun has done for me. I would also like to acknowledge my Ph.D. advisory committee. Thanks to each one of you for the aid in completing my project and graduate education. Your advice and encouragement has helped me greatly throughout my years at the GSBS. Another huge thanks to all the Kim lab members who have helped me throughout graduate school (Norah Owiti, Chris Lopez, Shivani Singh, Kasey Stokdyk, Wahaj Zuberi, Duy Long, Lee-Ann Notice, and Preethi Saravanan). All of you have taught me something valuable to help in lab and life, and I am sincerely lucky to call you my friends.

I would also like to acknowledge my past professors and mentors at Millikin University and Illinois State University. At my undergraduate education institution, Millikin, I discovered my passion for research working with Drs. Samuel Galewsky and Marrienne Robertson. And

working with Dr. Erik Larson at Illinois State University during my master's was an amazing experience and further solidified my love for DNA.

A special thanks to my friends Gloria Alvarado, Kylee Veazey, and Ashley Melissant. You have all three been there for me when I needed encouragement and fun. I would next like to thank my family who have always believed in me even when I didn't believe in myself. I am thankful for my father, Tim Berroyer, who always made sure I had what I needed to complete my education and for always encouraging me to stay the course even when life when gets tough. I am thankful for my mother, Susan Fleming, for being my first teacher and fostering my love of learning. I will never forget the sacrifices you made so I could have a wonderful life. I am thankful for my sister, Danielle Kirby, for her support and friendship throughout my whole life. Finally, thanks to my husband and ultimate best friend Edward Stalter for everything. This journey (from Millikin to now) would have been much harder without you, and I am so lucky we get to spend the rest of our lives together. You are my number one supporter always, and I appreciate that more than I can say or express.

# Investigation of Genomic Instability Induced by G-Quadruplexes in the Absence of Functional Topoisomerase 1 in Yeast

Alexandra Noel Berroyer, B.S., M.S.

Advisory Professor: Nayun Kim, Ph.D.

Topoisomerase 1 (Top1) is an enzyme that removes transcriptionally generated negative supercoils by binding and nicking DNA. Since transcription of guanine-rich DNA leads to the formation of G-quadruplex (G4) structures, Top1's function likely suppresses G4-formation. In support of this, Top1 significantly reduces co-transcriptional G4 DNA-associated genomic instability at a model G4-motif in *Saccharomyces cerevisiae*. However, whether Top1 suppresses G4-formation on a genome-wide scale in yeast remains unexplored. Therefore, I aimed to uncover if deletion of Top1 enhances genome-wide G4-formation in *S. cerevisiae*. As an approach to quantify global G4-formation, I expressed the G4-specific antibody BG4 from a yeast vector to perform chromatin immunoprecipitation next generation sequencing (ChIP-seq) and immunofluorescence experiments. While the G4-antibody's function was verified *in vitro*, ChIP and immunofluorescence experiments failed, possibly due to localization of BG4 to the cytoplasm rather than the nucleus of yeast cells. Thus, future attempts at enumerating G4s in *TOP1*-deletion yeast cells should include the usage of expressed BG4 fused to a nuclear localization signal sequence or purified BG4 protein.

Top1 mutants arise in cancer cells treated with the Top1-targeting anticancer drug camptothecin (CPT). Here, I show that the impact on G4-induced recombination in yeast depends on the type of CPT-resistant Top1 mutant expressed. While expression of a Top1 mutant defective in duplex DNA binding results in G4-recombination levels equivalent to cells

completely lacking Top1, expression of cleavage-defective Top1 mutants has an even greater impact on G4-mediated instability. I also find that Top1 cleavage-defective mutants bind G4s *in vitro* and that the SPRTN homolog Wss1 involved in DNA/protein crosslink resolution partly suppresses G4-induced recombination in yeast cells expressing Top1 cleavage-defective mutants. Collectively, these data suggest that Top1 cleavage-defective mutants induce instability at guanine-rich DNA through G4-stabilization *in vivo*. Further, I uncovered that another G4-binding protein, Nsr1 or yeast nucleolin, contributes to G4-instability in yeast cells expressing Top1 mutants and provide additional evidence indicating that Top1 cleavage-defective mutants and Nsr1 interact when bound to G4s to form a potential replication block. Bioinformatic data revealed that cancer genomes harboring Top1 mutants predicted to be functionally defective exhibit enriched mutagenesis at G4-motifs. Yeast genetic datum showing that Top1 cleavage-defective mutants and Nsr1 have a synergistic effect on G4-instability through cooperative G4-binding taken together with the result of bioinformatic analyses suggest that CPT-resistance conferring Top1 mutants could induce mutagenesis at G4-motifs in cancer cells and complicate patient treatment. Since loss of Top1 function increases G4-instability, identifying additional protein factors that suppress or instigate G4-mediated DNA damage in the absence of functional Top1 is an attractive future direction of this work.



## Table of Contents

Approval Sheet.....	i
Title Page.....	ii
Dedication.....	iii
Acknowledgements.....	iv
Abstract.....	vi
Table of Contents.....	viii
List of Figures.....	x
List of Tables.....	xii
List of Important Abbreviations.....	xiii
<b>Chapter 1: Introduction.....</b>	<b>1</b>
1.1 Introduction to non-B DNAs and G-Quadruplex (G4) Structures.....	2
1.2 Introduction to Topoisomerase 1 (Top1).....	8
1.3 The Role of Top1 in Preventing the Formation of G4 DNA and Other Secondary Structures.....	12
1.4 Top1 binding G4 DNA.....	14
1.5 Functional Consequence of Top1-G4 DNA Binding.....	16
1.6 Top1 Mutants in Cancer Cells and G4-Induced Genomic Instability.....	27
1.7 Summary and Significance.....	33
<b>Chapter 2: Materials and Methods.....</b>	<b>35</b>
<b>Chapter 3: Quantification of G4 Structures in Yeast in the Absence of Top1.....</b>	<b>49</b>
3.1 Introduction to G4-Detection Methods.....	50
3.2 Results.....	58
3.3 Discussion.....	77
<b>Chapter 4: Cleavage-defective Topoisomerase I Mutants Sharply Increase G-quadruplex-Associated Genomic Instability.....</b>	<b>83</b>
4.1 Introduction to Top1 Mutants.....	84
4.2 Results.....	95

4.3 Discussion.....	123
<b>Chapter 5: Discussion and Future Directions.....</b>	<b>138</b>
5.1 Summary.....	139
5.2 Enumeration of G4s in the complete absence of Top1 in Yeast.....	140
5.3 Role of Cleavage-Defective Top1 Mutants in G4-Induced Genomic Instability.....	145
5.4 Yeast Chromatin Remodelers and G4-Induced Genomic Instability.....	152
5.5 Overall Conclusions.....	160
<b>Bibliography.....</b>	<b>162</b>
<b>VITA.....</b>	<b>204</b>

## List of Figures

<b>Figure 1:</b> G4s form in guanine-rich DNA when guanines interact.....	4
<b>Figure 2:</b> G4s adopt different conformations.....	5
<b>Figure 3:</b> Yeast WT Top1 and Top1Y727F bind to G4 structures.....	22
<b>Figure 4:</b> A model of genome instability induced by co-transcriptionally formed G4 DNA and the effect of Top1 activity and mutation.....	34
<b>Figure 5:</b> Schematic of <i>SμG4-GTOP</i> and <i>SμG4-GBTM</i> yeast recombination reporters.....	38
<b>Figure 6:</b> BG4 antibody purified from <i>E. coli</i> .....	59
<b>Figure 7:</b> Schematic of pGAL-BG4-FLAG vector for expression of BG4 in yeast.....	61
<b>Figure 8:</b> BG4-FLAG antibody construct is expressed from the pGAL-BG4-FLAG vector in yeast.....	62
<b>Figure 9:</b> Expression of BG4 does not affect survival of yeast.....	63
<b>Figure 10:</b> BG4 expressed in yeast binds G4s <i>in vitro</i> .....	65
<b>Figure 11:</b> Fold enrichment of BG4 at <i>SμG4</i> G4 locus relative to the non-G4 capable <i>CAN1</i> locus <i>in vivo</i> .....	67
<b>Figure 12:</b> Expected immunofluorescence microscopy results of yeast cells expressing pGAL-BG4-FLAG, the vector control, or Nsr1 FLAG.....	69
<b>Figure 13:</b> Immunofluorescence microscopy of <i>top1Δ</i> yeast cells expressing pGal-BG4-3XFLAG.....	70
<b>Figure 14:</b> Immunofluorescence microscopy of <i>top1Δ</i> yeast cells expressing the vector control.....	72
<b>Figure 15:</b> Immunofluorescence microscopy of yeast cells expressing FLAG-tagged Nsr1....	75
<b>Figure 16:</b> Yeast homologs of human Top1 mutants found in cancer cells are expressed.....	88
<b>Figure 17:</b> Yeast Top1 mutants Y740Stop and S733E are resistant to CPT.....	91
<b>Figure 18:</b> Top1 catalytic mutants bind duplex DNA substrates <i>in vivo</i> , while Top1 DNA binding mutants do not.....	94
<b>Figure 19:</b> Top1 mutants bind G4 oligos <i>in vitro</i> and increase G4-induced recombination in yeast.....	97
<b>Figure 20:</b> C-terminal 3X-FLAG tag does not impact the function of WT yTop1.....	101

<b>Figure 21:</b> The effect of transcriptional repression on the <i>GTOP</i> and <i>GBTM</i> recombination rates of Top1 mutant strains.....	102
<b>Figure 22:</b> Fold enrichment of $\gamma$ Top1Y727F-HA at the <i>S<math>\mu</math>G4 G4</i> and <i>3kb</i> loci <i>in vivo</i> .....	104
<b>Figure 23:</b> Deletion of <i>WSS1</i> increases G4-induced recombination in Top1 catalytic mutant strains.....	106
<b>Figure 24:</b> Deletion of <i>NSR1</i> and deletion of the RGG domain of Nsr1 reduces G4-induced recombination in Top1 mutant strains.....	109
<b>Figure 25:</b> Nsr1 constructs are expressed from pADH1-Nterm Nsr1, pADH1-Cterm Nsr1, and pADH1-Nsr1 in yeast.....	112
<b>Figure 26:</b> Expression of full-length <i>NSR1</i> is required to greatly exacerbate G4-induced recombination in <i>TOP1Y727F nsr1<math>\Delta</math></i> and <i>TOP1Y740Stop nsr1<math>\Delta</math></i> strains.....	113
<b>Figure 27:</b> <i>NSR1-6XHA</i> and <i>NSR1<math>\Delta</math>RGG-6XHA</i> proteins are expressed in Top1 mutant backgrounds .....	117
<b>Figure 28:</b> <i>NSR1-6XHA GTOP</i> strains and untagged <i>NSR1 GTOP</i> strains have comparable G4-induced recombination rates.....	118
<b>Figure 29:</b> Top1 mutant and Nsr1 co-immunoprecipitation (co-IP) experiments in the presence of <i>VTC4</i> .....	119
<b>Figure 30:</b> Deletion of <i>VTC4</i> does not affect G4-induced recombination in Top1 mutant strains.....	120
<b>Figure 31:</b> Quantification of Top1 mutant interactions with Nsr1.....	122
<b>Figure 32:</b> Model of co-transcriptional G4-formation and the effect of Top1 activity and mutation on G4-induced genomic instability.....	130
<b>Figure 33:</b> Somatic mutations in the <i>TOP1</i> gene are associated with high mutation rates in cancer.....	132
<b>Figure 34:</b> Expected results of proposed ddPCR replication kinetics experiment.....	148
<b>Figure 35:</b> Deletion of <i>FUN30</i> increases G4-induced recombination in <i>TOP1</i> -deletion strains.....	158

## List of Tables

<b>Table 1:</b> Sequences of DNA oligonucleotides used in yeast WT Top1 and Top1Y727F pull downs in Figure 3.....	21
<b>Table 2:</b> A selected list of human Top1 C-terminal mutations from studies, cancer cell lines, and patient samples .....	31
<b>Table 3:</b> Primers used for strain construction in this work.....	39
<b>Table 4:</b> Top1 mutants studied.....	87
<b>Table 5:</b> Catalytic activity of Top1 mutants determined by measuring the Lys+ mutation rate of yeast strains harboring a <i>pTET-lys2-AG4</i> reporter .....	92
<b>Table 6:</b> Oligonucleotides used in binding assays.....	96

## List of Important Abbreviations

<b>Abbreviations</b>	<b>Full Name</b>
CPT	camptothecin
ChIP	chromatin immunoprecipitation
dsDNA	double-stranded DNA
GBTM	pTET-lys2-GBTM
G4	G-quadruplex DNA structure
GTOP	pTET-lys2-GTOP
NCL	nucleolin
Oligo	oligonucleotide
rDNA	ribosomal DNA
NTS	non-transcribed strand
RNAP	RNA polymerase
qPCR	real time quantitative PCR
ssDNA	single-stranded DNA
SUMO	SUMOylation post-translational modification
Top1	Topoisomerase 1
Top1-cc	Topoisomerase 1 covalent cleavage complex
TS	transcribed strand
TSS	transcription start site

## Chapter 1: Introduction

Note: Portions of this chapter were derived from a previously published review in *Genes*. BERROYER, A. & KIM, N. 2020. The functional consequences of eukaryotic topoisomerase 1 interaction with G-quadruplex DNA. *Genes*, 11(2), 1-15. *Genes* is an MDPI journal. *For all articles published in MDPI journals, copyright is retained by the authors. In addition, the article may be reused and quoted provided that the original published version is cited:*

<https://www.mdpi.com/authors/rights>.

Article

available

at

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7073998/pdf/genes-11-00193.pdf>.

## 1.1 Introduction to non-B DNAs and G-Quadruplex (G4) Structures

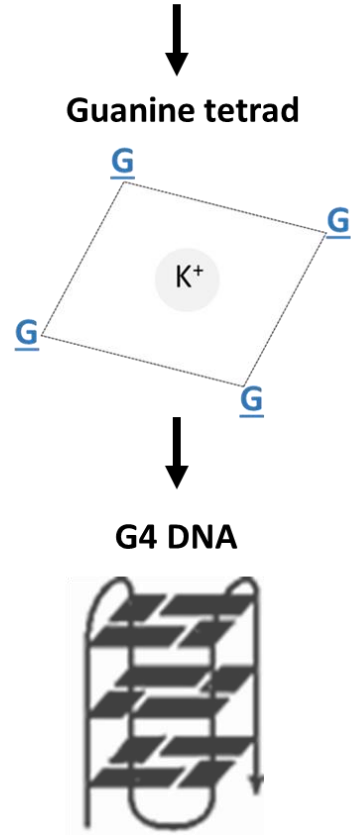
DNA can assume alternative DNA structures in addition to the canonical double-helical, B-form DNA (Wells, 2007; Mirkin, 2008; Zhao et al., 2010). Collectively termed “Non-B DNAs”, these alternative structures form during transcription, replication, and DNA repair when the duplex structure of B-form DNA is disrupted and strand separation occurs to foster non-B DNA-formation. This is because DNA strand separation greatly increases the risk for intra-strand base pairing or non-canonical base pairing to occur, which hold non-B DNA structures together. Additionally, non-B DNA formation is associated with negative helical stress as negatively supercoiled DNA can give rise to the formation of single-stranded DNA patches (Napierala et al., 2005; Sun and Hurley, 2009; Irobalieva et al., 2015).

There are many types of non-B DNAs including hairpins, cruciform, triplexes, Z-DNA, and R-loops. One of the most widely studied non-B DNAs is the G-quadruplex (G4) structure. G4s form in guanine-rich DNA when planar guanine tetrads, held together by Hoogsteen bonds among guanines, become stacked (Figure 1) (Burge et al., 2006). While non-canonical G4-motifs do exist and are structure capable (Piazza et al., 2017; Jana et al., 2021), the guanine-rich sequence of  $GGG_{N1-7}GGG_{N1-7}GGG_{N1-7}GGG$  is widely accepted in the field as the consensus, canonical G4-motif where N denotes the DNA bases between guanine triplets extruding from the structure as loops (Figure 1). G4-formation is dynamic, and a wide variety of G4 structure types have been documented. G4 DNAs can be either intermolecular where they are made up of more than one DNA strand or intramolecular where guanines from only a single DNA strand interact (Teng et al., 2021). Intramolecular G4s can adopt different structural topologies depending on if the glycosidic bonds of tetrad associated guanosines

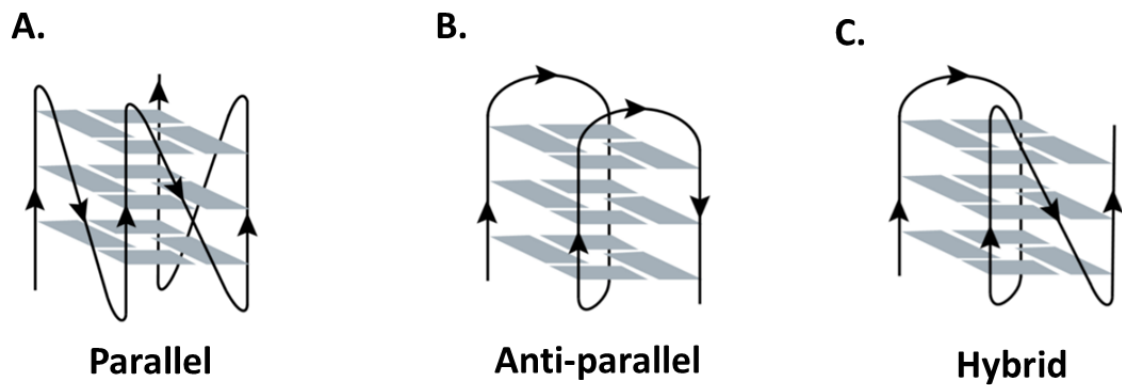


are in the *syn*- or *anti*-conformation. G4s containing guanosines all in the *anti*-conformation have a “parallel” topology where all four DNA strands of the structure have the same directionality (Figure 2A) (Burge et al., 2006; Teng et al., 2021). Conversely, G4s harboring a mix of guanosines in *anti*- and *syn*-conformations give rise to “anti-parallel” or “hybrid” G4s where the DNA strands holding the structure together are oriented in different directions (Figure 2B-C) (Burge et al., 2006; Teng et al., 2021).

**G4-forming sequence**  
**(NGGGNGGGNGGGNGGGN)<sub>n</sub>**  
“N” represents non-guanine bases



**Figure 1. G4s form in guanine-rich DNA when guanines interact.** Guanines from G4-forming sequences interact through Hoogsteen bonding to form planar guanine tetrads that stack upon each other to form the higher order G4 structure. Monovalent cations, like  $K^+$ , stabilize G4s through electrostatic interactions. Gray squares shown in folded G4 DNA structure represent guanine bases that comprise a tetrad.



**Figure 2. G4s adopt different conformations.** Derived from TENG, F., JIANG, Z., GUO, M., TAN, X, CHEN, F., XI, X., XU, Y. 2021. G-quadruplex DNA: a novel target for drug design. *Cell. Mol. Life Sci.*, 78, 6557–6583. <https://doi.org/10.1007/s00018-021-03921-8>. Gray squares represent guanines in tetrads. Black line represents DNA in G4-forming sequence and arrows denote directionality of the DNA strand. A. Parallel G4. B. Anti-parallel G4. C. Hybrid G4.

G4s have been shown to form in the genomes of essentially all lifeforms ranging from viruses and bacteria all the way to complex eukaryotes such as humans (Metifiot et al., 2014; Hänsel-Hertsch et al., 2017; Yadav et al., 2021). The human genome contains greater than 700,000 loci that are G4-capable under conditions that promote G4-stabilization (Chambers et al., 2015). A study investigating genome-wide *in vivo* G4-formation in cells found that G4s form in nucleosome depleted regions undergoing active transcription (Hänsel-Hertsch et al., 2016). The *Saccharomyces cerevisiae* genome contains ~1,400 G4-motifs (Todd et al., 2005). While the exact number of stable G4s that form in the yeast genome remains to be uncovered, genetic experiments suggest that as in human cells, G4-formation occurs during active transcription in yeast (Kim and Jinks-Robertson, 2011).

The non-random genomic locations of G4s within the human and *S. cerevisiae* genomes suggest that G4s are involved in certain DNA transactions. G4 DNAs occur in repetitive DNA containing tandem guanines, such as guanine-rich micro- and minisatellites (Todd et al., 2005; Capra et al., 2010; Amrane et al., 2012; Ogloblina et al., 2015). G4s are also associated with functional genetic elements including telomeres, ribosomal DNA, and promoter regions (Todd et al., 2005; Capra et al., 2010). Multiple studies have demonstrated the enrichment of G4 motifs at gene promoters suggesting that G4s play both activating and inactivating roles in transcription (Kim, 2019). Some of G4 DNA's roles in transcriptional regulation are related to the modulation of chromatin remodeling (Varizhuk et al., 2019) (Reina and Cavalieri, 2020). G4-motifs are present at meiotic and mitotic recombination hotspots, suggesting they play a role in programmed recombination (Capra et al., 2010). More recently, G4s have been shown to play a role in DNA repair, where experiments conducted in

yeast revealed that G4-formation followed by G4-binding of the Zuo1 protein directs nucleotide excision repair proteins to UV-induced DNA damage located near G4s (De Magis et al., 2020).

While there is no question that G4s play positive, modulatory roles in cells, G4-formation, if unregulated, can be extremely deleterious. Cells do have a repertoire of proteins that can resolve G4s when they are no longer needed (Mendoza et al., 2016), but G4s still are a source of genomic instability, especially when their formation or stability in cells is altered. G4s have been shown to stall transcription and both leading and lagging strand replication (Sarkies et al., 2010; Lopes et al., 2011; Sarkies et al., 2012; Schiavone et al., 2014; Dahan et al., 2018; Kim, 2019). The stalling of transcription and replication machinery will lead to DNA breaks if not correctly restarted (Gomez-Gonzalez and Aguilera, 2019). And the mutagenic capacity of G4s extends to both large-scale and small-scale genomic variations. In terms of large-scale genome changes, G4s are linked to translocations as well as the whole loss of chromosome arms (Katapadi et al., 2012; Yadav et al., 2014; Bacolla et al., 2016). In terms of small-scale genome changes, portions of the human genome harboring a high density of G4-motifs have increased levels of nucleotide deletions, nucleotide insertions, and single nucleotide polymorphisms (Williams et al., 2020).

Since G4s cause genomic instability when dysregulated, it is not surprising that aberrant G4-formation is linked to disease. Bloom syndrome is a disease caused from mutation of the gene encoding the BLM helicase (German, 1993), which is a potent G4-unwinder (Sun et al., 1998). Bloom syndrome patients have a strong pre-disposition to cancer, and a study by van Wietmarschen et al. showed that G4s contribute greatly to

genomic instability in BLM-deficient cells (van Wietmarschen et al., 2018). This indicates that aberrant G4-stability could contribute to the high rates of cancer development among Bloom syndrome patients. Further linking G4s to cancer development, oncogenic breakpoints were shown to occur at G4s (Nambiar et al., 2011; Nambiar et al., 2013; Williams et al., 2015). G4s present at oncogenic translocation break points were shown to block progression of DNA polymerases, suggesting that G4-replication conflicts could be a source of the DNA breaks that promote formation of these harmful rearrangements (Williams et al., 2015). And studies conducted with a G4-specific antibody showed that ~10,000 more G4s form in immortalized keratinocytes than in primary keratinocytes, suggesting cancer cells have a higher propensity for G4-formation and G4-induced genomic instability than healthy cells (Hänsel-Hertsch et al., 2016). In addition to cancer, G4s are linked to neurodegenerative disorders. Amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD), linked diseases marked by degeneration and loss of neurons (Lomen-Hoerth et al., 2002; Umoh et al., 2018), are commonly caused by expansion of the hexanucleotide repeat (HRE) GGGGCC<sub>n</sub> located in the *C9ORF72* gene (DeJesus-Hernandez et al., 2011; Renton et al., 2011). Because the *C9orf72* HRE is G4-capable, it was proposed that G4s contribute to ALS-FTD pathogenesis by promoting expansion of HRE (Fratta et al., 2012; Reddy et al., 2013; Haeusler et al., 2014). Further, the binding of the protein nucleolin to the *C9ORF72* G4 could stabilize the structure and lead to development of ALS-FTD (Haeusler et al., 2014).

## **1.2 Introduction to Topoisomerase 1 (Top1)**

Eukaryotic Topoisomerase I (Top1) is a type IB topoisomerase that binds to double-stranded DNA (dsDNA) and cleaves a single strand initially forming a phospho-tyrosyl bond at

the 3' end of the cleaved DNA. This intermediate is termed the Top1 cleavage complex. Self-catalyzed ligation to the 5' hydroxyl restores the intact double-stranded DNA. During Top1's catalytic cycle, helical torsional DNA stress in the form of positive or negative supercoils are removed by the swiveling of DNA strands between the cleavage and ligation steps.

As mentioned above, the first step in the catalytic cycle of topoisomerase reaction is binding to DNA. The X-ray structure of human Top1 bound to a DNA duplex shows that two lobes clamp around B-form dsDNA throughout the catalytic cycle (Stewart et al., 1998; Redinbo et al., 1998). Core domains I and II form the "cap" lobe while the core domains III, C-terminal domain, and the linker domain form the "catalytic" lobe. Although a 22-bp duplex DNA was used for the structure determination, Top1 contacts only the central 10 bp of the DNA from the position -4 to +6 (with the cleavage between the -1 and +1 positions). The essential catalytic tyrosine residue (Y723 for human Top1) is covalently attached to the phosphodiester bond 3' of the -1 position nucleotide. When this residue is mutated to a non-catalytic phenylalanine, the aromatic ring is closely positioned facing the phosphodiester bond between the -1 and +1 nucleotides. In total, 24 to 26 amino acids make direct contact with the DNA molecule, but there is only one base-specific contact indicating that the Top1–DNA interaction is not strongly sequence-dependent. Further experiments showed that, rather than sequence, superhelicity of DNA determines the strength of Top1–DNA interaction (Madden et al., 1995). Using the catalytically inactive mutant with phenylalanine at amino acid 723, it was shown that human Top1 binds preferentially to the superhelical DNA and particularly to the nodes created by the crossing of the two helical dsDNAs. In this experiment, Top1 Y723F initially bound to a relaxed DNA molecule redistributed to bind

added supercoiled DNA indicating the Top1–DNA interaction is dynamic and transient for relaxed DNA. No preference for positively or negatively supercoiled DNA was found. Although the core, the linker and the C-terminal domains participate in DNA binding, a truncated form that contains only the linker and the core domain retained the preference for binding supercoiled DNA. The linker domain was later determined to be important in the preferential binding to supercoiled DNA (Yang et al., 2009). Vaccinia Top1 is an example of a topoisomerase with a strict sequence requirement for binding; it binds and cleaves at 5'-CCCTT-3' (Sekiguchi and Shuman, 1994). For eukaryotic Topoisomerases, however, the sequence requirement is less restricted. Inferring the sites of Top1 binding preference from the sites of Top1 cleavage preference, the consensus for strong cleavage for rat and wheat germ Top1 is 5' A/T G/C T/A T (from the -4 to -1 position). However, a significant level of cleavage occurs when the -1 position is a C residue (Been et al., 1984). When the Top1 cleavage sites are analyzed by trapping the cleavage complex with the Top1 poison camptothecin (CPT), the sites of strong preference are clearly shifted from its preferential cleavage site in the absence of the intercalating drug (Siu and Pommier, 2013).

Top1's catalytic activity of removing positive and negative supercoils from DNA serves to maintain genomic stability (Kim and Jinks-Robertson, 2017). Specifically, positive supercoiled DNA is overwound while negatively supercoiled DNA is underwound. In the absence of Top1, negative and positive supercoils accumulate in cells during transcription. Top1 is recruited to genomic regions undergoing active transcription through interactions with RNA polymerase 2 and chromatin remodeling proteins so that torsional stress can be relieved (Phatnani et al., 2004; Baranello et al., 2016; Husain et al., 2016). Both positive and



negative forms of transcriptionally generated helical stress are deleterious, thus their removal by Top1 protects the genome (Ma and Wang, 2014; Ma and Wang, 2016; Kim and Jinks-Robertson, 2017; Tubbs and Nussenzweig, 2017). Negatively supercoiled DNA is linked to genomic instability as this loose DNA conformation renders DNA more prone to damage and induces the formation of non-B DNA structures (Hamperl and Cimprich, 2014). Positively supercoiled DNA can prevent DNA unwinding and causes genomic instability by blocking RNA polymerase during transcription, which can lead to dsDNA breaks (Gomez-Gonzalez and Aguilera, 2019). Further, our understanding of the function of Top1 has expanded beyond maintaining torsional homeostasis in DNA to protect the genome. In 1997, Sekiguchi and Shuman reported on the ribonuclease activity of vaccinia and human Top1 and suggested the potential role of Top1 in processing RNA (Sekiguchi and Shuman, 1997). Later, the endoribonuclease activity of eukaryotic Top1 was shown to be required in its role of removing ribonucleotides incorporated into DNA (Kim et al., 2011; Williams et al., 2013; Cho et al., 2013). In absence of a functional RNase H2 complex, which initiates error-free ribonucleotide excision repair, Top1-mediated cleavage at single ribonucleotides embedded in a DNA strand leads to mutations and genome instability (Cornelio et al., 2017; Huang et al., 2017). The role of Top1 in transcriptional regulation was also recently highlighted. Along with the type II topoisomerase Top2, Top1 functions to recruit RNA polymerase to the promoters of highly transcribed genes (Sperling et al., 2011) and is required to maintain the superhelicity of certain inducible promoters (Pedersen et al., 2012). More recently, the role of Top1 as a component of the RNA polymerase complex and as a positive regulator of promoter escape and transcription elongation was described (Baranello et al., 2016). In the following section

of this dissertation another seminal function is discussed: the role of Top1 in reducing the formation of genotoxic DNA secondary structures.

### **1.3 The Role of Top1 in Preventing the Formation of G4 DNA and Other Secondary Structures**

Non-canonical DNA structures can disrupt processes such as replication, transcription, and DNA repair. An R-loop or the long extensive hybrid between RNA and DNA strands is one such non-canonical nucleic acid structures that is a major source of endogenous genome instability (Aguilera and Gaillard, 2014). Top1 function counteracts the formation of extensive R-loops at highly transcribed regions. At the ribosomal DNA repeat loci in the yeast genome, the loss of Top1 leads to disruption of transcription and accumulation of R-loops, both of which are severely exacerbated in absence of the RNA:DNA hybrid processing nucleases RNase H1 and H2 (El Hage et al., 2010; El Hage et al., 2014). R-loop accumulation in the absence of Top1 function and the elevated genome instability as a consequence were also observed in mammalian cells (Manzo et al., 2018). Apart from the RNA–DNA hybrid or R-loops, DNA can assume structures other than the Watson-Crick double helical, B DNA structure (Mirkin, 2008). Duplex DNA, under the condition of increasing supercoils, particularly negative torsional stress, undergoes structural transformation to partially single-stranded DNA (Irobalieva et al., 2015). The exposed bases in this context can lead to intra-strand interactions promoting secondary structure formation. For example, base-pairing within a strand of DNA can occur at inverted repeats, forming a basis for a hairpin structure or a cruciform DNA. CAG trinucleotide repeat-containing DNA strands are known to assume stable hairpin structures through the intra-strand base pairing of Cs and Gs. Other unusual,

non-B structure includes the triplex or H-DNA formed at purine-rich sequences such as GAA trinucleotide repeat loci. Removal of negative torsion by Top1 prevents stretches of single-stranded DNA from folding into triplex, cruciform, and hairpin non-canonical secondary structures. The hairpin-forming CAG trinucleotide repeats are found at human genes associated with the neurological disorders Huntington's disease and spinocerebellar ataxia. The manifestation of these diseases is dependent on the expansion of the repeat sequences, which in turn is associated with the topological and structural changes at these genomic loci due to non-B DNA structure formation. In a screen of small molecules, Top1-inhibitors were identified as inducing the expansion of CAG trinucleotide repeats in a human cell culture system (Hubert et al., 2011). This was confirmed by separate work which showed that the knock-down of Top1 significantly elevated the instability of a large CAG-repeat in a human fibrocarcinoma cell line (Nakatani et al., 2015). By removing the supercoils generated during the transcription of these sequences, Top1 likely prevents the formation of pathological non-B DNA structures and helps maintain stability at these unusual repetitive sequences. Another non-B structure of interest is the four-stranded G-quadruplex DNA or G4 DNA. Held together by the Hoogsteen bonds among four guanine bases forming a ring-like structure called a G-quartet or tetrad, G4 DNA is composed of multiple runs of guanines forming multiple G-quartets stacked on top of each other (Figure 1) (Bochman et al., 2012). The significance of G4 DNA as a source of genome instability has recently become evident with the bioinformatics studies finding G4 motifs, along with other non-B DNA forming sequences, to be highly enriched at chromosomal translocation hot spots found in cancers (Katapadi et al., 2012; Bacolla et al., 2016). At BCL2 and c-MYC, two genes known to be involved in recurrent

blood cancers, G4 motifs are located close to the major chromosomal break point region (Katapadi et al., 2012). A high density of G4 motifs is also present in BCR (B-cell receptor) or the immunoglobulin heavy chain (IgH) loci, which are frequently involved in genome rearrangements and other changes in blood cancers. In mouse B lymphocytes, a decrease in the protein level of Top1 resulted in elevation in the class switch recombination (CSR) at IgH loci (Kobayashi et al., 2009; Kobayashi et al., 2011). Chromosomal translocation between IgH locus and c-Myc is also increased by the knock down of either Top1 or the chromatin remodeler SMARCA4, which is required for the efficient recruitment of Top1 to the chromatin (Husain et al., 2016). The decrease in the recruitment of Top1 to the switch regions was concurrent with the increase in the negative helicity at the same locus as measured by the incorporation of modified psoralen molecules that bind preferentially to underwound DNA. CSR initiates with the activation of transcription of the switch sequences that are unusually G/C-rich and contain multiple runs of guanines capable of folding into G4 DNA. Increased propensity for G4 DNA folding at these sequences could explain the elevation in CSR and in IgH/c-Myc translocation observed with reduced Top1 levels. Similarly, when a fragment of IgH switch region was embedded into the yeast genome, the rate of recombination occurring at this sequence was significantly elevated in the absence of Top1 (Yadav et al., 2014; Yadav et al., 2016). The importance of Top1 function in preventing the detrimental effect of G4 DNA was confirmed in human cells where Top1 was shown to protect cells from the genotoxic effect of G4-binding small molecules (Zyner et al., 2019).

#### **1.4 Top1 Binding G4 DNA**

In addition to playing a role in suppressing the formation of non-B DNA structures by preventing the accumulation of negative supercoils, a handful of studies demonstrated that Top1 can also physically bind to non-B DNA structures including G4 DNA (Thiyagarajan et al., 1998; Arimondo, 2000; Marchand et al., 2002; Shuai et al., 2010). The first strong evidence of human Top1 interaction with G4 structures is described in studies by Arimondo et al., in which an electrophoretic mobility shift assay was used to show that purified human Top1 binds to preformed intermolecular and intramolecular G4 structures (Arimondo et al., 2000). Interestingly, this same study uncovered that Top1 promotes the formation of intermolecular G4 structures. When oligonucleotides (oligos) containing stretches of five or six consecutive guanines were incubated with purified human Top1, a protease-resistant four-stranded intermolecular complex of slower mobility was formed. This G4-formation activity was specific to Top1 and not observed with Top2, histone H2A, or BSA. Another set of experiments using human Top1 further confirmed that Top1 binds to G4 DNA-forming oligos as well as G-rich DNA and RNA oligos (Marchand et al., 2002). This study also showed that the cleavage of duplex DNA by human Top1 is inhibited by the presence of intermolecular or intramolecular G4 structures, and that this inhibition is due to the binding of Top1 to G4 DNA. It was also shown that duplex DNA cleavage by Top1 is inhibited by single-stranded, non-G4 capable DNA and RNA oligos containing stretches of two or three consecutive guanines. Pre-formed G4 oligos or guanine-rich single-stranded oligos were not cleaved by Top1. More recently, an effort to find oligonucleotide aptamer inhibitors of human Top1 confirmed that a variety of DNA oligos forming G-quadruplexes or quadruplex-duplex hybrids bind Top1 and compete it away from its substrate, dsDNA (Shuai et al., 2010). Many proteins have so far been identified

as G4 DNA binding proteins. The protein-G4 DNA interactions can roughly be separated into three categories according to their effect on the stability of G4 structures; (1) promoting G4 DNA formation (yeast Rap1 and human thrombin) (2) stabilizing G4 DNA (murine nucleolin and human Ku protein) and (3) G4 DNA destabilizing (RecQ helicases BLM, WRN, and Sgs1) (Fry, 2007). More recently, a transcription factor Sub1 was identified as a potent G4 DNA binding protein that indirectly leads to destabilization of G4 via recruitment of the helicase Pif1 (Gao et al., 2015) (Lopez et al., 2017). Sub1 or its mammalian homolog PC4, previously characterized as a single-strand DNA binding protein, binds to G4 DNA with a low  $K_d$  of  $\sim 2$  nM but does not promote or disrupt G4 DNA folding (Griffin et al., 2017). Wildtype Top1, based on the experiments reported in Arimondo et al. can be considered to be in the category of proteins promoting G4 DNA formation (Arimondo et al., 2000; Fry, 2007). In the future, it will be interesting to uncover how Top1 binds to G4 DNA structures. According to the X-ray crystal structure published, Top1 forms a tight, bi-lobed clamp around duplex DNA upon binding (Redinbo et al., 1998). Since there is a significant difference between the diameters of duplex DNA and G4 DNA, which are approximately 2 nm and 2.4 – 2.8 nm, respectively (Do et al., 2011; Heddi and Phan, 2011; Amrane et al., 2012), the Top1–G4 DNA complex possibly adopts a very distinct conformation from Top1-dsDNA.

## **1.5 Functional Consequence of Top1-G4 DNA Binding**

### **1.5.1 G4 Oligos as Top1 Inhibiting Aptamers**

One possible functional consequence of the specific, high-affinity Top1–G4 interaction is the application of G4-forming oligos or aptamers as inhibitors of Top1. Aptamers are small single-stranded DNA or RNA oligos that are selected for a high affinity

interaction with a specific target of interest (Platella et al., 2017). Aptamers targeting proteins or other biologically relevant molecules have been selected by screening libraries of DNA or RNA oligos and further studied for their potential therapeutic application. Since it was first discovered that Top1 binds and is inhibited by G4-forming oligos (Arimondo, 2000; Marchand et al., 2002) other similar results have been reported (Shuai et al., 2010; Ogloblina et al., 2015; Ogloblina et al., 2018). First, using purified Calf Thymus Top1, Shuai et al. characterized fourteen different guanine-rich oligos capable of forming either G-quadruplexes or quadruplex-duplex hybrids and showed that all of the oligos act as competitive inhibitors of Top1 catalysis (Shuai et al., 2010). The inhibitory effect of these G4 aptamers were significantly more potent when they were treated with heat-cooling in presence of KCl to fold into G4 conformations. The IC<sub>50</sub> of the G4-formed aptamers ranged from 0.1 to 2.7 μM. Top1 activity can be inhibited by small molecules such as camptothecin (CPT) that stabilize the Top1-cleavage complex (Top1-cc) and inhibit the ligation of the nicked substrate (Megonigal et al., 1997; Pommier et al., 2010). The mechanism of Top1 inhibition by the G4 aptamers, however, appeared to be mediated by the inhibition of the Top1-substrate binding. When Top1 was incubated with a mixture of a G4-forming aptamer and dsDNA substrate (i.e., pBR322 plasmid DNA), Top1 preferentially bound to the G4 aptamer (Shuai et al., 2010). The inhibitory effect of G4-forming oligos on Top1 activity was further confirmed in a study of guanine-rich microsatellite repeat sequences found in the human genome (Ogloblina et al., 2015). In this work, highly thermostable G4-forming oligos d(GGT)<sub>4</sub> and d(GGGT)<sub>4</sub> were added to HeLa cell extract to determine how each of these oligos affects the Top1-mediated relaxation of supercoiled pUC19 plasmid DNA. Compared to a random oligo of similar length,

which showed no effect in the relaxation assay, both G4 oligos d(GGT)<sub>4</sub> and d(GGGT)<sub>4</sub> were very effective in inhibiting Top1 with the IC<sub>50</sub>s of 0.63 and 0.12 μM, respectively. Another G-rich oligo d(GT)<sub>16</sub>, which is not capable of forming a thermostable G4 structure, did not inhibit Top1 activity, indicating the structure-specific nature of inhibition. When non-G4 flanking sequences were added to the effective inhibitor d(GGGT)<sub>4</sub>, the thermostability was decreased with the resulting T<sub>m</sub> of 85 and 73 °C, respectively, for d(GGGT)<sub>4</sub> and d(CACTGG-CC-(GGGT)<sub>4</sub>-TA-CCAGTG). But the longer oligo proved to be a more effective Top1 inhibitor with IC<sub>50</sub> of 0.08 μM. In another study, six different guanine-rich aptamers were tested for their inhibitory effect on human Top1 (Ogloblina et al., 2018). These aptamers were previously designed to target other oncogenic proteins, namely STAT3, SP1, VEGF, NCL, and SHP-2, and commonly formed either parallel or anti-parallel G4 DNAs that are significantly stabilized by the presence of potassium cation as expected. All of these aptamers were very effective in inhibiting the activity of Top1 in a plasmid-relaxation assay with the IC<sub>50</sub> in the low μM range. They were even more effective in inhibiting Top1 when compared to the aptamer first identified as a Top1-specific inhibitor by Shuai et al. (Shuai et al., 2010). The inhibition of Top1 by the G4 oligos was correlated with the inhibition of DNA replication, and this antiproliferative effect was specific to cancer cells (Ogloblina et al., 2018). Although the physical interaction between the G4-capable aptamers and Top1 was not determined in these studies, the specific nature of the inhibition strongly suggests that the competitive binding of the aptamers to Top1 underlies the inhibition of the catalytic activity.

### **1.5.2 The Biological Consequence of the Top1 Interaction with G4 DNA**



While Top1-binding to G4 DNA is an interesting property to exploit in the use of G4-forming aptamers to target Top1 as a therapy, whether Top1 binding to endogenous, genomic G4 structures serves a biological function is not yet clear. Chromatin immunoprecipitation experiments conducted in yeast revealed that Top1 is enriched at telomeres, which contain G4 DNA forming sequences (Lotito et al., 2008). This study also demonstrated that the expression of the yeast Top1 catalytic mutant,  $\gamma$ Top1Y727F, in a *top1 $\Delta$*  background elevates H4 K16 histone acetylation at genomic regions located proximal to telomeres. These results suggest that Top1 regulates transcription of telomere proximal genes and that the catalytic activity of Top1 is required for this function. It is possible that Top1 regulates chromatin state and expression of genes near telomeres through G4 DNA binding. Another possible function of the Top1–G4 DNA interaction is in the recruitment of G4-resolvases to the genomic G4 structures. Human Top1 was shown to interact with the Werner helicase, which can unfold G4 structures (Lebel et al., 1999; Mendoza et al., 2016), suggesting that it is possible that Top1 promotes the localization of the Werner helicase to G4 structures through its own interaction with G4 DNAs. Top1 also interacts with the SV40 T antigen, which harbors DNA helicase activity (Stahl et al., 1986). These examples of Top1 interaction with the Werner helicase and the SV40 T antigen suggest further studies should be conducted to determine whether Top1 interacts with additional DNA helicases, particularly those helicases capable of unwinding G4 DNAs.

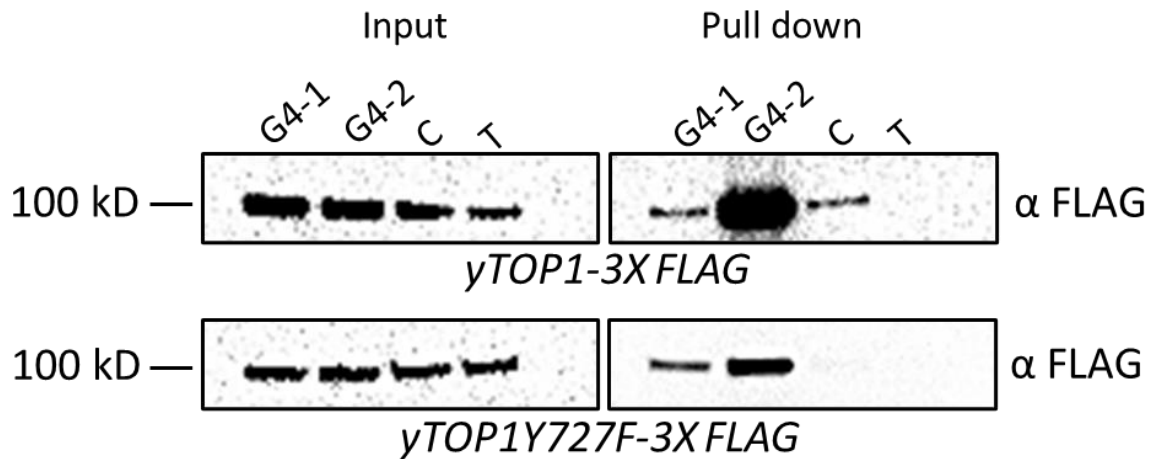
### **1.5.3 Interaction Between Mutant Top1 and G4 DNA *In Vivo***

Even though the interaction of G4 DNA with functional Top1 may result in transcriptional regulation or G4 structure resolution, another observation suggests that the

interaction of G4 DNA with Top1 catalytic mutants is deleterious. Human and *Saccharomyces cerevisiae* Top1 use amino acid residues tyrosine 723 and tyrosine 727, respectively, to undergo the nucleophilic attack of the phosphodiester DNA backbone effectively nicking the DNA (Redinbo et al., 1998). However, if either of these residues is mutated to a phenylalanine, Top1 can bind, but not nick DNA. Interestingly, expression of Top1Y727F in yeast results in exacerbated recombination at a model G4-motif (Yadav et al., 2016). This elevated G4-induced recombination observed in the presence of Top1Y727F is significantly greater than the G4-induced recombination observed in a *top1Δ* yeast strain and is dependent on transcription. The effect of Top1Y727F on G4-induced genomic instability is surprising as the level of superhelical tension accumulation is expected to be similar in a *top1Δ* strain and a Top1Y727F-expressing yeast strain. Therefore, the increase in G4-induced genomic instability observed in a Top1Y727F-expressing yeast strain compared to a *top1Δ* strain must be from another factor in addition to negative supercoil accumulation. Yeast Top1Y727F was shown to be enriched at telomeres in chromatin immunoprecipitation experiments (Lotito et al., 2008) and, *in vitro*, it preferentially binds to G4 oligos over a C-rich or a random control oligo (Table 1, Figure 3). Top1Y727F binding and stabilizing G4 structures would explain the highly elevated genomic instability at G4-motifs. Further, while WT Top1 may bind to G4 structures transiently, the lack of catalytic activity after DNA binding by yeast Top1Y727F may result in the trapping of Top1Y727F on G4 structures.

Oligo	Sequence
G4-1	5' GAGCT <u><i>GGGG</i></u> TGAGCT <u><i>GGG</i></u> CTGAGCT <u><i>GGGG</i></u> TGAGCT <u><i>GGG</i></u> CTGAGCT
G4-2	5' <u><i>AGGG</i></u> CTCTGCCTT <u><i>GGGGGGGGGG</i></u> CAGGA <u><i>AGGG</i></u> A
C	5' AGCTCAGCCCAGCTCACCCCAGCTCAGCCCAGCTCACCCCAGCTC
T	5' GCACGCGTATCTTTTTGGCGCAGGTG

**Table 1. Sequences of DNA oligonucleotides used in yeast WT Top1 and Top1Y727F pull downs in Figure 3. Guanines are underline and italicized.**



**Figure 3. Yeast WT Top1 and Top1Y727F bind to G4 structures.** Western blots of pulldowns of WT Top1-3XFLAG (top) and Top1Y727F-3XFLAG (bottom) from yeast whole cell lysates with biotinylated DNA oligonucleotides (MilliporeSigma). Biotinylated oligonucleotides G4-1, G4-2, C, and T were conjugated to Streptavidin-Coupled M-280 Dynabeads. Following the mechanical lysis of yeast cells with a Biospec Mini-bead-beater, the cell lysate was collected and sonicated. Oligo-conjugated Dynabeads were incubated at 4 °C overnight with the yeast extract, washed, and then eluted by boiling in 1XSDS-PAGE loading buffer followed by immunoblotting analysis using anti-FLAG antibody to detect 3XFlag-tagged Top1 or Top1Y727F.

Another way Top1Y727F could increase the instability of G4 DNA-forming genomic loci is by binding to G4 DNA and then preventing G4 resolvases from accessing the structure. This mechanism has been described for the human protein nucleolin (NCL) (Indig et al., 2012). NCL is an essential nuclear/nucleolar protein with multiple known functions (Tajrishi et al., 2011). Most prominently, NCL is involved in the ribosomal RNA maturation. Mammalian NCL, however, was also shown *in vitro* to be capable of high-affinity binding to G4 DNA with K<sub>d</sub> in the low nM range (Dempsey et al., 1999; Hanakahi et al., 1999). More recently, the specific interaction between NCL and the G4 DNA formed from the (GGGGCC)<sub>n</sub> repeat in the human C9orf72 gene was demonstrated *in vivo* (Haeusler et al., 2014). Expansion of the (GGGGCC)<sub>n</sub> repeats in the C9orf72 gene is associated with the neurological disorders amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD). An *in vitro* helicase assay has shown that G4 DNA, when in complex with NCL, becomes resistant to unwinding by Werner helicase, a G4 resolvase (Indig et al., 2012). If G4 structures are left unresolved, they could be potent blocks to DNA replication leading to genome rearrangements (Weitzmann et al., 1996; Sarkies et al., 2010; Paeschke et al., 2011). Similarly, it has been shown that Nsr1, the yeast homolog of NCL, binds to G4 structures *in vitro* and *in vivo* (Singh et al., 2020). Deletion of the *NSR1* gene or the expression of a truncated form of Nsr1 missing an important G4 DNA-binding domain in a *nsr1Δ* background significantly reduces recombination at a model G4-motif. This indicates that Nsr1, like NCL, increases G4-induced instability through G4 binding. Of note, *NSR1*-deletion also reduces G4-induced recombination in a *TOP1Y727F* background, however, not to wild type levels as observed in *top1Δ nsr1Δ* strain (Singh et al., 2020). Because Nsr1 and Top1 physically interact (Azevedo et al., 2015), it is possible that these two

G4 DNA binding proteins form a higher-order complex to increase G4-induced instability in a synergistic manner.

#### **1.5.4 DNA-Protein Complexes as DNA Replication Barriers**

Replication forks can stall and collapse at DNA lesions leading to genome instability. Such genotoxic impediment to replication can include DNA–protein complexes (Gadaleta and Noguchi, 2017). Especially, proteins covalently trapped on DNA are known to be natural replication blocks. During the removal of helical tension, Top1 forms a phospho-tyrosyl bond with the 3' end of nicked DNA, and this covalent DNA–protein complex is termed a Top1-cleavage complex (Top1-cc) (Redinbo et al., 1998). After strand swiveling has occurred, the 5' OH of the nicked strand attacks the phospho-tyrosyl bond connecting Top1 to DNA which results in the ligation of DNA and release of the enzyme. Top1-ccs, such as those induced by the Top1 poison camptothecin (CPT) and its derivatives, block DNA replication and induce DNA strand breaks and recombination (Pommier et al., 2010). On the other hand, there are examples of non-covalent proteins-DNA complexes that block DNA replication machinery in both prokaryotes and eukaryotes. In *Escherichia coli*, a protein called Tus binds to Ter sites within DNA, and this DNA–protein interaction terminates replication in a unidirectional manner to prevent the collision of replication forks approaching each other in a head-on orientation (Hill et al., 1987; Hill et al., 1989; Kamada et al., 1996). The replication fork barrier created by the Tus-Ter interaction functions to ensure proper replication termination of the circular chromosome (Mohanty et al., 1996). Another example of non-covalent DNA–protein complex that can block replication in *E. coli* is the array of lacI repressor molecules bound to Lac operon (Payne et al., 2006). In eukaryotes, tight DNA–protein interactions that inhibit

DNA replication occur at the highly transcribed genomic loci that encode ribosomal RNA. In *S. cerevisiae*, similar polar Replication Fork Blocks (RFBs) are generated at each of the 150–200 rDNA tandem repeats located at Chromosome XII (Brewer and Fangman, 1988). Fob1 binding at the rDNA array on chromosome XII is an example of the DNA–protein complex serving as replication fork block (Kobayashi and Horiuchi, 1996). In yeast cells defective for the DNA helicase Rrm3, replication was blocked at discrete genomic loci including known heterochromatin regions such as centromeres and silent mating type loci but also at tRNA genes, inactive replication origins, and transcriptional silencers (Ivessa et al., 2003). These regions are characterized by the non-histone DNA–protein complexes, indicating that high-affinity DNA–protein complexes can form replication blocks or barriers.

DNA secondary structures such as G4 DNA also form potent replication barriers (Gadaleta and Noguchi, 2017). G4 DNA in the template strand stalls the replicative polymerases *in vitro* and translesion DNA polymerases such as the mammalian Pol  $\epsilon$  and  $\kappa$  and bacterial Pol IV are necessary for continued synthesis (Woodford et al., 1994; Sarkies et al., 2010; Eddy et al., 2015; Eddy et al., 2016; Berroyer et al., 2019). Replication in yeast is impeded at G4 motifs when the G4 DNA helicase Pif1 is disrupted (Ribeyre et al., 2009; Paeschke et al., 2011; Paeschke et al., 2013). G4 DNA stabilized by small molecule ligands such as pyridostatin (PDS) or PhenDC3 increase the instability at G4-forming sequences in eukaryotic genomes indicating that the stability of G4 DNA correlates with its efficacy as replication block (Piazza et al., 2010; Rodriguez et al., 2012; Piazza et al., 2015; Moruno-Manchon et al., 2017). Similarly, a more problematic obstacle to replication is expected when G4 DNA is in complex with a high-affinity binding protein such as the Top1

catalytic mutant (Top1Y727F), which binds to G4 DNA to form a stable complex (Table 1, Figure 3). When the Top1Y727F mutant was expressed in *top1Δ* yeast cells, there was a significant and G4-specific elevation in the recombination rate (Yadav et al., 2016). On the other hand, expression of the Top1T722A mutant, which can competently cleave DNA but is defective in the re-ligation step leading to the formation of Top1-ccs, resulted in a modest elevation in recombination independent of the sequence. CPT-treatment also led to a non-specific elevation of recombination. Together, these data indicate that the G4-specific elevation of recombination upon expressing Top1Y727F is not due to inhibition of Top1 catalytic activity or due to increased formation of Top1-ccs. The preferential binding of the mutant protein to G4 DNA is then one mechanism possibly underlying the sharp increase in G4-specific recombination following the expression of Top1Y727F. That is, Top1 mutants bound tightly to G4 DNA structures could be strongly disruptive to replication and recombinogenic in a manner similar to those covalent and non-covalent DNA–protein complexes that function as replication barrier sites.

While proteins that are covalently trapped or are tightly bound to DNA are known to block replication, they can be resolved to prevent genomic instability in several different manners. In yeast, the protease named weak suppressor of *SMT3* protein 1 or Wss1, degrades proteins trapped on DNA, including Top1-ccs (Stingle et al., 2014). Top1-ccs become SUMOylated in yeast, and the SUMOylation of proteins trapped on DNA by DNA-bound SUMO ligases directs and enhances the recruitment of Wss1, which contains SUMO-interacting motifs. Wss1 degrades DNA-trapped proteins almost entirely, leaving behind small peptide remnants that are further processed by the proteasome or are bypassed during replication



by translesion polymerases. The human homolog of Wss1 is a protein named Spartan; the loss of Spartan leads to an accumulation of unrepaired Top1-ccs in human cells (Stingele et al., 2016; Vaz et al., 2016; Maskey et al., 2017). However, instead of SUMO-interacting motifs as in yeast Wss1, the recruitment and activity of Spartan is regulated by its ubiquitin-binding domain (Stingele et al., 2017). In the future, it will be interesting to explore if Top1 mutants trapped on G4 DNAs are modified by SUMOylation and/or ubiquitylation and subsequently processed by Wss1 and Spartan in *S. cerevisiae* and humans, respectively.

### **1.6 Top1 Mutants in Cancer Cells and G4-Induced Genomic Instability**

Top1 has long been the molecular target of chemotherapeutics. CPT and CPT derivatives are widely used to treat many cancers, and work by stabilizing Top1-ccs as mentioned above. Cancer cells can become resistant to CPT through multiple different mechanisms (Beretta et al., 2013). First, consistent with cell culture experiments demonstrating that the reduced levels of Top1 confer resistance to CPT, cancers relapsing after chemotherapy spontaneously become resistant to CPT through reduced expression of Top1 (Burgess et al., 2008; Zuco et al., 2010). Cancer cells can also become impervious to CPT-treatment through mutations of Top1 that reduce the ability either to bind to duplex DNA or to cleave duplex DNA following binding. A study conducted in yeast found intragenic suppressors of the yeast Top1T722A mutation (Hann et al., 1998), which mimics the cytotoxic activity of CPT. One such suppressor mutant, yeast Top1G369D, has a critically reduced duplex DNA binding ability. Another study found that T729 is mutated in a human lung cancer cell line that is resistant to the CPT analog irinotecan (Kubota et al., 1992). Further investigation demonstrated that expression of the human Top1T729K and Top1T729E

mutants in yeast confers resistance to CPT, and that these Top1 mutants are defective in duplex DNA binding (Losasso et al., 2008). Other studies have demonstrated that Top1 catalytic mutants, which harbor reduced duplex DNA cleavage abilities, exist in cancer cell lines and cancer patients. The homozygous Top1G365S mutant, which was found in a colon cancer cell line that is resistant to the active metabolite of irinotecan, displays 50% reduced catalytic activity (Arakawa et al., 2006). Another mutant, Top1W736Stop, was found in a non-small cell lung cancer patient treated with irinotecan and is predicted to have reduced catalytic activity (Tsurutani et al., 2002). A large number of Top1 mutations identified in CPT-resistant cancer cells and patient samples remain uncharacterized. Some of the mutations located in the catalytic C-terminal domain of human Top1 (amino acids 713–765) are listed in Table 2. Although these are missense or nonsense mutations located in the catalytically critical domain of the protein, the residues in the C-terminal domain also participate in the contact with dsDNA (Stewart et al., 1998; Redinbo et al., 1998). For a majority of those listed mutations, whether each of the mutations results in the loss of DNA cleavage or DNA binding activity of Top1 has not been studied. Of note, additional mutations not listed that are located in the core domain of Top1 can also result in reduced enzyme catalysis, like the G365S mutation found in a CPT-resistant colorectal cancer cell line as described above (Arakawa et al., 2006). Top1 is required to suppress recombination and instability at a highly transcribed G4-motif in yeast and suppresses the genotoxic effect of G4-ligands in human cells (Yadav et al., 2014; Yadav et al., 2016; Zyner et al., 2019). siRNA-mediated knock-down of Top1 in mouse cells elevates CSR and chromosomal translocations involving the G4 DNA-forming IgH switch region sequences (Kobayashi et al., 2011; Husain et al., 2016). The reduced levels of

Top1 in the CPT-resistant human cancer cells is then expected to disrupt its normal cellular role of preventing excess levels of negative supercoiling at highly transcribed loci and to thereby elevate the accumulation of non-B DNA formation including G4 DNA. The reduced Top1 levels and also the reduced Top1 function in cells with DNA-binding defective mutants (e.g., Top1T729K and Top1T729E) very likely manifest in elevated genome instability associated with the formation of G4 DNA and other non-B structures. However, in those CPT-resistant cancers with Top1 mutants that are catalytically defective but competent for DNA binding, additional consideration must be made according to an interesting piece of preliminary data from yeast studies. Expression of the yeast Top1 mutant Y727F further exacerbates the instability at G4 DNA-forming genomic loci beyond the loss of Top1 function by additionally binding to G4 DNA and other G4 DNA-binding proteins (Azevedo et al., 2015; Yadav et al., 2016; Singh et al., 2020). It needs to be examined whether a similar G4-specific deleterious effect is produced with the expression of human Top1 mutants such as Top1G365S and Top1W736Stop mutants, which can bind but not cleave DNA. Incidences of secondary cancers following treatment with the CPT-derivative irinotecan have been documented. In one case, a patient with X-linked agammaglobulinemia and metastatic colorectal cancer experienced an increase in cancer progression and severe hypocalcemia after treatment with irinotecan (Li et al., 2019). Another study revealed that a colon cancer patient developed secondary acute promyelocytic leukemia following treatment with irinotecan and oxaliplatin (Merrouche et al., 2006). While both Top1 mutations and secondary genomic rearrangements have been discovered in cancer patients treated with CPT or CPT derivatives, a link between these documented Top1 mutations and G4-induced

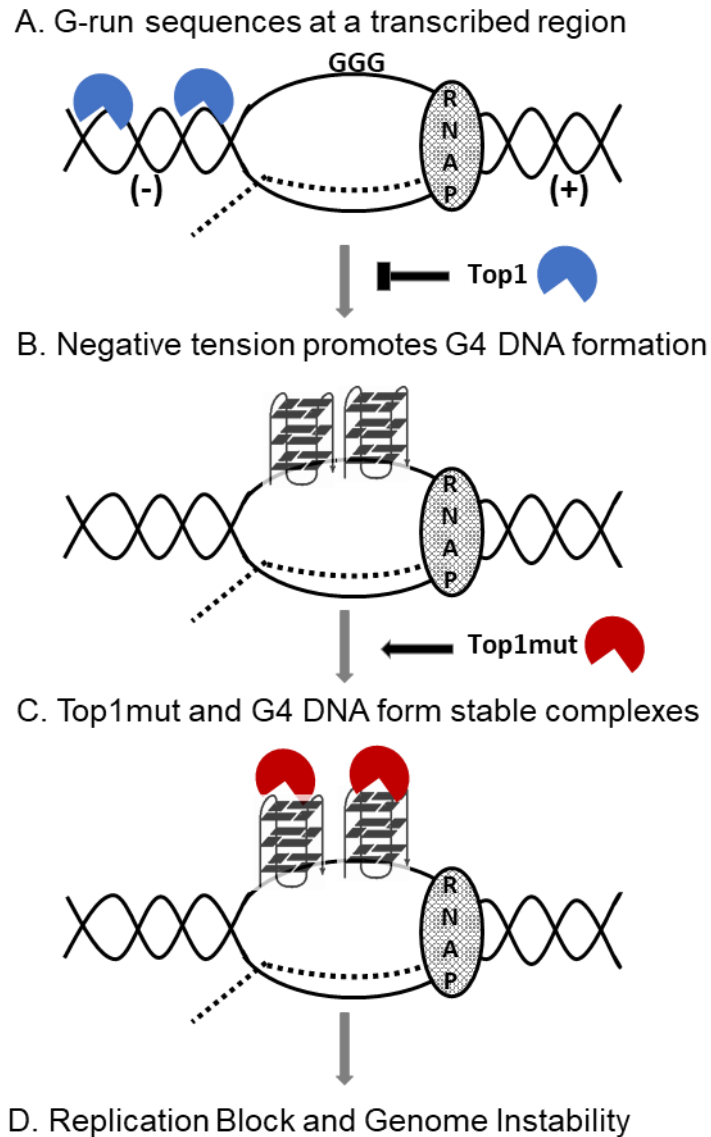
secondary genomic rearrangements has not been studied. In the future, it will be important to uncover if cancer cells from patients treated with CPT or CPT derivatives have increased secondary genomic rearrangements at G4-forming loci. The increased potential for G4-induced genomic instability due to cleavage-defective Top1 catalytic mutants could render cancers more complicated to treat and potentially worsen patient outcomes.

<b>Top1 Mutant</b>	<b>Origin</b>	<b>Reference and Mutation ID</b>
Q713H	breast cancer tissue tumor sample	ICGC(BRCA-US); COSMIC Genomic Mutation ID COSV63696584
I714T	human colorectal cancer cell line	98 COSMIC Genomic Mutation ID COSV63693067
R727S	urinary tract cancer tumor sample	99 COSMIC Genomic Mutation ID COSV99057653
T729A	irinotecan-resistant human lung cancer cell line	94
W736C	urinary tract cancer tumor sample	ICGC/GDC(BLCA-US); COSMIC Genomic Mutation ID COSV63695286
W736STOP	non-small cell lung cancer patient treated with irinotecan	97
E741STOP	urinary tract cancer tumor sample	100 COSMIC Genomic Mutation ID COSV63696859
T747P	prostate cancer tumor sample	101 COSMIC Genomic Mutation ID COSV63695236
R749W	soft tissue/smooth muscle tumor sample	100 COSMIC Genomic Mutation ID COSV63696078
R749Q	large intestine carcinoma sample	ICGC(COAD-US); COSMIC Genomic Mutation ID COSV63696579
A753S	liver cancer tumor sample	ICGC(LICA-CN); COSMIC Genomic Mutation ID COSV63692735
A759T	large intestine carcinoma tumor sample	ICGC(COCA-CN) COSMIC Genomic Mutation ID COSV63694655

**Table 2. A selected list of human Top1 C-terminal mutations from studies, cancer cell lines, and patient samples.** ICGC, International Cancer Genome Consortium (<https://icgc.org>); COSMIC, Catalogue of Somatic Mutations in Cancer (see reference Tate et al., 2018 and [cancer.sanger.ac.uk](https://cancer.sanger.ac.uk)).

## 1.7 Summary and Significance

Given the significance of G4 DNA-forming sequences in the development of cancer-associated genome rearrangements, it is very important to study how guanine-run containing sequences are converted into hotspots of genome instability. As shown in yeast and mammalian systems, Top1 plays an essential role in preventing genome instability at G4 DNA-forming loci by removing transcription-associated superhelical tension and maintaining a proper topological conformation (Figure 4 A-B). Furthermore, newly emerging evidence indicate that Top1 can specifically and preferentially interact with G4 DNA. The interaction between Top1 mutants and G4 DNA, however, could lead to the formation of pathological complexes that interfere with efficient DNA replication (Figure 4C-D). Based on a preliminary observation obtained from a yeast model system linking the expression of a Top1 mutant with acutely elevated G4 DNA-associated genome instability, further characterization of human Top1 mutants, particularly those arising in CPT-treated cancer cells, could shed light on the significance of the interaction between mutant proteins and non-B DNA structures. On the other hand, the significance of the Top1–G4 DNA interaction can be viewed from a very different perspective. Multiple recent studies suggest that G4 DNA-forming oligos can preferentially bind and interfere with the activity of Top1, which has been a very important target of anti-cancer therapy. This property can be exploited to develop a new class of drugs targeting Top1 that can serve as an alternative and a complement to the Top1-poisoning CPT and CPT derivatives.



**Figure 4. A model of genome instability induced by co-transcriptionally formed G4 DNA and the effect of Top1 activity and mutation.** RNAP—RNA polymerase complex. Dotted line—the nascent transcript. (-)—negative tension behind the transcription complex. (+)—positive tension ahead of the transcription complex. Top1mut—Top1 mutant.



## Chapter 2: Materials and Methods

Note: Portions of this chapter were derived from a manuscript accepted for publication in *Microbial Cell*. BERROYER, A., BACOLLA, A., TAINER, J. A., and KIM N. 2022. Cleavage-defective Topoisomerase I mutants sharply increase G-quadruplex-associated genomic instability. *Microbial Cell*, Accepted January 19<sup>th</sup> 2022. *microbialcell.com* is the property of the Shared Science Publishers OG. Unless otherwise stated, all articles and corresponding materials accompanying them that are published by Shared Science Publishers OG on this Web Site are licensed by their respective authors for your use and distribution provided you cite the original source as specified in the corresponding Creative Commons Attribution License. Creative Commons Attribution License weblink: <https://creativecommons.org/licenses/by/4.0/>. Article available at <http://microbialcell.com/researcharticles/2022a-berroyer-microbial-cell/>.

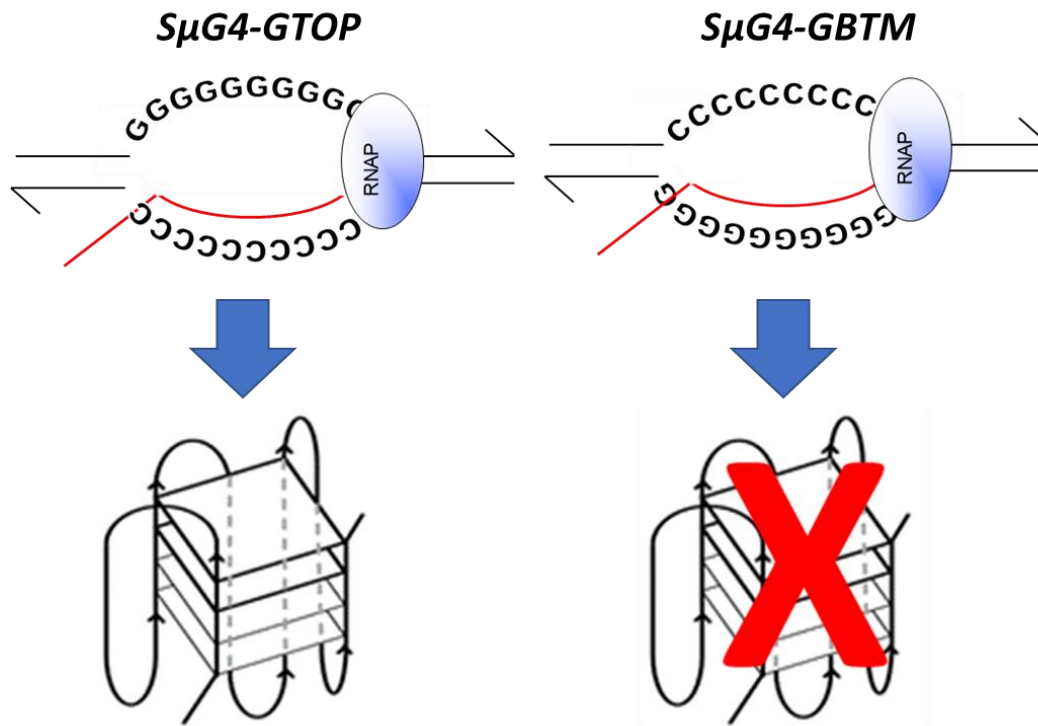
## 2.1 Protein Purification

HIS- and FLAG-tagged BG4 antibody was expressed in *E. coli* BL121 (DE3) cells from the pSANG10-3F-BG4 plasmid (Addgene; Cat #55756). BG4-HIS-FLAG protein was purified using the protocol from Moruno-Manchon et al., 2017. Briefly, one liter of *E. coli* cells harboring pSANG10-3F-BG4 were grown shaking at 37 °C until an OD<sub>600</sub> of 0.8-0.9 was reached. Then IPTG was added to the culture to a final concentration of 1 mM and cells were grown shaking at 30 °C for 15-17 hours. Next, cells were collected by centrifugation and pellets were frozen at -80 °C. Cell pellets were resuspended in NP40 buffer (25 mM Tris-HCl pH 8.0, 100 mM NaCl, 10% glycerol, 1% NP40, 10 mM imidazole, and 1X Protease Inhibitor Cocktail II (GenDEPOT; Cat #P5101-005)) then lysed via sonication (QSONICA sonicator; 2 sec ON, sec OFF, 6 min. at 30% amplitude output). Following sonication, lysates were clarified by centrifugation at 4°C for an hour at 16,000 rcf. Then, 2 ml of HisPur™ Ni-NTA Resin (ThermoFisher) was equilibrated with NP40 buffer, added to the clarified lysate, and incubated with the clarified lysate overnight rotating at 4 °C. The next day, the lysate was passed through a column to collect the HisPur™ Ni-NTA Resin. The HisPur™ Ni-NTA Resin was washed with 50 ml of wash buffer (NP40 buffer + 40 mM imidazole). HIS tagged BG4-FLAG was eluted from the HisPur™ Ni-NTA Resin by the passing of 10 ml of elution buffer (NP40 buffer + 250 mM imidazole) through the Ni resin packed column. Purified BG4 elutions were analyzed by resolving elution samples via SDS PAGE followed by Coomassie gel staining.

## 2.2 Yeast Strain and Plasmid Construction

Yeast strains used in this study were derived from YPH45 (*MATa*, *ura3-52 ade2-101 trp1Δ1*) and the construction of the *pTET-lys2-GTOP* (*SμG4-GTOP*) and *-GBTM* (*SμG4-GBTM*)

reporter containing strains was previously described (Kim and Jinks-Robertson, 2011). A schematic of the *S $\mu$ G4-GTOP* and *S $\mu$ G4-GBTM* recombination reporter constructs is in Figure 5. Unless noted otherwise, all yeast gene knock out and Top1 and Nsr1 epitope-tagged strains were constructed by one-step allele replacement where parental strains were transformed by the LiOAc method (Ito et al., 1983) with PCR products containing selectable marker cassettes. All PCR primers used in strain construction and allele replacement are listed in Table 3. All tagged and mutated strains were confirmed by Sanger sequencing. The construction of *yTOP1Y727-3XFLAG* strains was previously described (Yadav et al., 2016). The *yTOP1S733E-3XFLAG* mutant strains were created using the “delitto perfetto” method (Storici and Resnick, 2006); *yTOP1-3XFLAG* strains were first transformed with PCR product created with the Top1S733E F and Top1S733E R primers to insert the *URA3* marker near the S733 codon. A second transformation with a duplex oligo containing the S733E mutant codon (primers Top1S733E and Top1S733E RC annealed) was performed, resulting in the loss of the *URA3* marker and 5-Fluoroorotic acid resistance. The *NSR1-6XHA* strains were constructed by transformation with PCR product consisting of the pHYG-AID\*-6HA plasmid (Morawska and Ulrich, 2013) amplified with the Nsr1-6XHA F and Nsr1-6XHA R primers. The *Nsr1 $\Delta$ RGG-6XHA* strains were constructed by transformation with PCR product consisting of the pHYG-AID\*-6HA plasmid amplified with the Nsr1 $\Delta$ RGG-6XHA F and Nsr1 $\Delta$ RGG-6XHA R primers.



**Figure 5. Schematic of *SμG4-GTOP* and *SμG4-GBTM* yeast recombination reporters.** The guanines of *SμG4* are located on the non-transcribed strand in *GTOP*, while the guanines of *SμG4* are located on the transcribed strand in *GBTM*. *GTOP* is G4-capable during transcription while *GBTM* is not. The *GTOP* and *GBTM* constructs were inserted into a *LYS2* gene that is controlled by a doxycycline repressible promoter and block production of lysine. If a strand break occurs in *GTOP* or *GBTM*, a truncated version of the *LYS2* gene is used as a template for recombination and lysine will be produced. Recombination rates at *GTOP* and *GBTM* are determined by scoring *LYS+* recombinants. RNAP = RNA polymerase complex. Nascent mRNA is denoted by red line.

<b>Primer</b>	<b>Sequence</b>
<b>Top1Y740Stop F</b>	5' CCAGGTTTCACTGGGCACTTCCAAAATCAATTATATAGACCCTA GACTTTCTGTGGTATTTGCAAAAAGAGGGAACAAAAGCTGG AG
<b>Top1Y740Stop R</b>	5' CCGATTCTATGGCCATTGAATTTTCTCTTAGGGTTTTGTA AAAATCTTTCAATCGGAACATCATACTATAGGGCGAATTGGG T
<b>Top1S733E F (Allele replacement step 1)</b>	5' GTTGAAAGATAAAGAGGAAAACCTCCAGGTTTCACTGGGCAC TTCCAAAATCAATTATATAGACCCTAGACTTTCTCAGCTGAAGC TTCGTACG
<b>Top1S733E R (Allele replacement step 1)</b>	5' GAATTTTTCTCTTAGGGTTTTGTA AAAATCTTTCAATCGGAA CATCATACTTTTGCAAAATACCACAGGCCACTAGTGGATCTG
<b>Top1S733E (Allele replacement step 2)</b>	5' GGAAAACCTCCAGGTTTCACTGGGCACTTCCAAAATCAATTAT ATAGACCCTAGACTTGAAGTGGTATTTGCAAAAAGTATGATG TTCCGATTGAAAAGATTTTACAAAACCTAAG
<b>Top1S733E RC (Allele replacement step 2)</b>	5' CTTAGGGTTTTGTA AAAATCTTTCAATCGGAACATCATACT TTTGCAAAATACCACTTCAAGTCTAGGGTCTATATAATTGATT TGGAAGTGCCAGTGAAACCTGGGAGTTTTCC
<b>Nsr1-6XHA F</b>	5' TGCTAATACTGCTCCATTGGGCAGATCAAGAAATACCGCTTCTT TCGCTGGTTCAAAGAAAACATTTGATCCCGGAAATACCCATA C
<b>Nsr1-6XHA R</b>	5' TCAAAAATGGAATATAGTAATTAACGTAAAAGAGAAAAAATT GAAATTGAAATTCATTTCTCAGTTAAAGCCTTCGAGC GTCCC
<b>Nsr1ΔRGG-6XHA F</b>	5' GGTGAATACATTGACAACAGACCAGTTAGATTAGACTTCTCTT CTCCAAGACCAACAACGATCCCGGAAATACCCATAC
<b>Nsr1ΔRGG-6XHA R</b>	5' AATGGAATATAGTAATTAACGTAAAAGAGAAAAAATTGAAAT TGAAATTCATTTCTCAGTTAAAGCCTTCGAGCGTCCC

**Table 3. Primers used for strain construction in this work.**

The pGAL-BG4-FLAG yeast expression plasmid was created by amplifying the coding sequence of BG4 harboring a C-terminal 3X FLAG-tag from the pSANG10-3F-BG4 vector (Addgene Cat #55756) via PCR. The BG4-3XFLAG PCR product was then inserted into the pRS426 2 micron vector (Christianson et al., 1992) containing a galactose inducible promoter (pGAL) and a *URA3* marker using *Spe1* and *EcoR1* restriction sites. The Top1Y727F-HA yeast expression plasmid (pAR7-HA) (Lotito et al., 2008) used for ChIP RT-qPCR experiments was a gift from the Capranico lab at the University of Bologna in Bologna, Italy. The pADH1-Nterm Nsr1 (N-term Nsr1; amino acids 1-171) and pADH1-Cterm Nsr1 (C-term Nsr1; amino acids 172-414) plasmids were gifts from A. Saiardi from the University College London, UK (Azevedo et al., 2015). The pADH1-Nsr1 (Nsr1 FL; amino acids 1-414) plasmid was constructed by replacing the Nsr1 N-terminus DNA sequence from pADH1-Nterm Nsr1 with the full length Nsr1 open reading frame that was PCR-amplified from the pRS316-derived Nsr1-expression *CEN* vector that has been previously described (Singh et al., 2020).

### **2.3 Western Blotting**

Yeast whole cell lysates were prepared for western blotting as previously described (Kushnirov, 2000). Whole cell lysate samples were centrifuged and resuspended in 2X SDS sample buffer and boiled for 10 min at 95 °C before being resolved on 4-20% SDS-PAGE gels (Bio-Rad; Cat# 456-1093). Proteins were transferred to PVDF membrane using a Trans-Blot® SD cell machine (Bio-Rad; Cat# 170-3940) and then probed with either an  $\alpha$ -FLAG antibody conjugated to horse radish peroxidase (HRP) (Sigma; Cat# A8592), an HRP-conjugated  $\alpha$ -HA antibody (Sigma; Cat# H6533), an  $\alpha$ -HRP-conjugated GAPDH antibody (Invitrogen; Cat# MA5-15738-HRP), or a primary  $\alpha$ -GST antibody (Invitrogen; Cat# MA4-004) followed by incubation

with a secondary HRP-conjugated  $\alpha$ -mouse Ig antibody (R&D Systems; Cat# HAF007) as indicated in the respective figures. Blots were visualized by treatment with GenDEPOT West-Q Femto ECL (Cat# W3680-010) and a Bio-Rad ChemiDoc™ MP imaging system. Quantification of Top1 and Nsr1 protein levels was performed using Image Lab software. The pixel volumes of FLAG- or GST-bands after subtracting background were divided by GAPDH loading control pixel volumes. Averages and standard deviations from at least 3 independent experiments were calculated and a student's T-test (GraphPad Prism) was used to assess statistical differences where indicated.

#### **2.4 Cell Survival Spot Assays**

Cell survival spotting assays were performed by first growing yeast cells transformed with either pGAL-BG4-FLAG or the empty vector control (2 $\mu$  vector with pGAL promoter and *URA3* selectable marker) in 5 mls synthetic complete dextrose media lacking uracil (SCD-ura) overnight at 30 °C. The next day, cultures were diluted to an OD<sub>600</sub> of 0.3 in 5 mls SCD-ura media and were grown until the culture reached an OD<sub>600</sub> of 1. Then, a total of 0.3 OD<sub>600</sub> of cells were taken from each culture and were resuspended in 100  $\mu$ l of water. The cells were then diluted 1:10 serially 5 times and 10  $\mu$ l of each dilution was spotted on plates containing either 2% glucose, 2% raffinose, 2% raffinose + 0.01% galactose, 2% raffinose + 0.5% galactose, 2% raffinose + 1% galactose, or 2% raffinose + 2% galactose. Spot assay plates were imaged every day for 3 days after growth at 30 °C using a Bio-Rad ChemiDoc™ MP imaging system.

Cell survival spot assays of Top1 mutant yeast cells plated on CPT-containing plates was conducted as listed above with the following modifications. *MUS81*-deletion cells were

grown in 5 mls of liquid YPD media overnight at 30°C shaking. The next day, cultures were diluted to an OD<sub>600</sub> of 0.15 in 5 mls of liquid YPD media and were grown at 30°C until the cultures reached an OD<sub>600</sub> of 0.6-0.7. A total of 0.3 OD<sub>600</sub> of cells were collected from each culture and were resuspended in 100 µl of sterile water. Then, 5 µl of serial 1:10 dilutions of cells were spotted onto YPD plates containing either 0.35% DMSO, 5 µM CPT + 0.35% DMSO, or 20 µM CPT + 0.35% DMSO. Spot assay plates were imaged every day for 2 days after growth at 30 °C using a Bio-Rad ChemiDoc™ MP imaging system. CPT used was purchased from Sigma (Cat# 208925).

## **2.5 *In vitro* DNA Binding Assays**

*In vitro* DNA binding assays were performed as previously described with some modifications (Gao et al., 2015). Oligos with 5'- and 3'-biotin attachments used in pull downs were purchased from Sigma. For each sample, 25 pmol of biotinylated oligos were folded by boiling for 5 min at 95°C in a heat block followed by slow cooling to room temperature in 100 µL of 10 mM Tris pH 7.5 + 100 mM KCl. Folded oligos were incubated while rotating with 6.25 µL of Streptavidin-Coupled M-280 Dynabeads (Invitrogen; Cat# 11205D) that were washed twice and resuspended in 10 mM Tris HCl pH 7.5 + 1 mM EDTA + 300 mM KCl. After 1 h incubation at room temperature, oligo-bound beads were washed twice with and resuspended in 5 mM Tris HCl pH 7.5 + 0.5 mM EDTA + 150 mM KCl and kept at 4°C until further use. Oligo-bound beads were rinsed once with 1 mL lysis buffer B (50 mM Hepes pH 7.5, 1 mM EDTA pH 8.0, 300 mM KCL, 10% glycerol, 0.05% NP40, 1mM PMSF, 1 mM DTT, and fungal protease inhibitor cocktail (Sigma; Cat# P8215) (50 µl/ 10 ml)) and were resuspended in 100 µL lysis buffer B right before being added to yeast whole cell lysates.



For yeast whole cell lysate preparation, a 5 mL YPD overnight culture was used to inoculate a 500 mL YPD culture and cells were grown shaking at 30°C until the culture reached an OD<sub>600</sub> of 1.5 – 2.0. A total of 410 OD<sub>600</sub> of cells were collected via centrifugation at 4°C after washing twice with H<sub>2</sub>O and once with lysis buffer A (50 mM HEPES pH 7.5, 1 mM EDTA pH 8.0, 300 mM KCL, and 10% glycerol) followed by freezing at -80°C. Frozen pellets were resuspended in lysis buffer B and lysates were prepared by 4 rounds of bead beating with acid washed glass beads for 30 sec followed by 5 min of incubation on ice. After bead beating, lysates were sonicated for 10 cycles of 20 sec ON/40 sec OFF at low amplitude with a Bioruptor (Diagenode) at 4 °C. Lysates were clarified by centrifugation at 4°C and oligo-conjugated streptavidin magnetic beads were added. A magnet was used to pull down beads after overnight incubation rotating at 4°C followed by four 1 mL washes in lysis buffer B and elution in 50 µL 2X SDS sample buffer. Subsequent western blot analyses of elutions and input samples (clarified lysate) were carried out as described above. Pull down/Input was calculated by dividing the pull-down pixel volumes by the input pixel volumes. Averages and standard deviations of ratios from 3 independent experiments were calculated and a student's T-test (GraphPad Prism) was used to assess statistical differences where indicated.

## **2.6 Chromatin Immunoprecipitation**

ChIP was performed as previously described (Lopez et al., 2017) (Singh et al., 2020). In short, yeast cells transformed with pGAL-BG4-FLAG, the control vector pRS426, or the Top1Y727F-HA yeast expression plasmid (pAR7-HA) were crosslinked with 1% formaldehyde followed by quenching with 136 mM glycine. Cells were lysed by bead beating and sonication was used to shear the chromatin into ~500 bp fragments using a Bioruptor (Diagenode) with

the following settings at 4 °C: 5 cycles of 30 sec ON/30 sec OFF at high amplitude. Samples were incubated with 10 µl of Anti-FLAG® M2 Magnetic Beads (Sigma) or Protein G-Dynabeads (Life Technologies) conjugated to an α-HA antibody overnight at 4 °C, followed by washing and cross-link reversal by incubation with proteinase K at 42 °C for 4 hours and 65 °C for 12 hours. DNA was purified with the MiniElute PCR Purification Kit (Qiagen) followed by quantitative (q) PCR with SensiFAST SYBR no-ROX Mastermix (Bioline) and a CFX Connect instrument (Biorad). Twelve and a half ng of input or ChIP DNA was used as a template and 0.4 µM of primers were used in each PCR reaction with cycling conditions of 95 °C for 3 min. followed by 40 cycles of 95 °C for 5 s, 60 °C for 10 s, and 72 °C for 10s. The qPCR primers used have been previously described (Lopez et al., 2017) (Singh et al., 2020) and target either a locus that is 100 bp from the *SµG4* insertion site (“*G4*”), a negative control locus ~3 kb from *SµG4* (“*3 kb*”), or another negative control, non-G4 capable locus (“*CAN1*”). Fold enrichment of BG4 and HA-Top1Y727F at the *G4* locus relative to the *CAN1* locus was calculated for each sample using Cq values obtained from qPCR and the following equation:  $2^n$  ( $n = (G4 \text{ input Cq average} - CAN1 \text{ input Cq average}) - (G4 \text{ IP Cq average} - CAN1 \text{ IP Cq average})$ ). Fold enrichment of HA-Top1Y727F at the *3 kb* locus relative to the *CAN1* locus was calculated for each sample using the same enrichment equation from above, replacing *G4* Cq averages with *3 kb* Cq averages.

## 2.7 Immunofluorescence

Immunofluorescence with BG4 was performed as previously described (Pringle et al., 1991). WT and *top1Δ* cells transformed with pGAL-BG4-FLAG were grown at 30°C until an OD<sub>600</sub> of 1. Formaldehyde (4%) was added to the culture flask, and cells were crosslinked for

10 min. at 30°C shaking. After that, cells were collected by centrifugation, resuspended in 40mM KPO<sub>4</sub> (pH6.5), 500µM MgCl<sub>2</sub>, and 4% formaldehyde, and were incubated shaking at 30 °C for another hour. Then, formaldehyde was quenched by addition of 100mM glycine. Cells were washed twice with 40mM KPO<sub>4</sub> (pH6.5) and 500µM MgCl<sub>2</sub> and once with 40mM KPO<sub>4</sub> (pH6.5), 500µM MgCl<sub>2</sub>, and 1.2M sorbitol. Cells were then resuspended in 40mM KPO<sub>4</sub> (pH6.5), 500µM MgCl<sub>2</sub>, and 1.2M sorbitol and were made into spheroplasts by incubation at 37°C for 30 min with 0.6 mg/ml Zymolase 20T (USBiological). Next, spheroplasts were washed once with 40mM KPO<sub>4</sub> (pH6.5), 500µM MgCl<sub>2</sub>, and 1.2M sorbitol and resuspended in 1X PBS followed by permeabilization with 0.5% Triton-X for 10 min at room temperature. After permeabilization, spheroplasts were washed three times with 1X PBS and blocked with 1% BSA, 1X PBS, and 0.5% Tween-20 for 50 min. at room temperature. Fluorophore conjugated Dylight 488 α-FLAG antibody (ThermoFisher Scientific) (2 µg/ml) was added to spheroplasts and incubated at room temperature for 1 hr. Then spheroplasts were washed three times with 1X PBS + 0.5% Tween-20 and resuspended in 1X PBS. DAPI dihydrochloride (Invitrogen) (300 nM) was added and incubated with the spheroplasts for 10 min at room temperature. Spheroplasts were washed five times with 1X PBS, then mounted onto a polylysine coated slide with Prolong™ Diamond Antifade Mountant (Invitrogen). Prepared slides were be visualized with an Olympus Fluoview FV3000 confocal microscope by Dr. Karan Kaval. Immunofluorescence was also performed with yeast expressing FLAG tagged Nsr1, the most abundant protein in the nucleolus of yeast cells (Ginisty et al., 1999) as a positive control and with yeast transformed with an empty vector as a negative control.

## **2.8 Determination of Recombination and Mutation Rates**

Fluctuation analysis was performed as previously described (Kim and Jinks-Robertson, 2011). The *SμG4-GTOP* and *SμG4-GBTM* recombination reporter constructs are pictured in Figure 5. Briefly, 12-36 individual 1-mL cultures of each strain were used for fluctuation analyses and the recombination rates were calculated using the method of median as described previously (Spell and Jinks-Robertson, 2004). Recombination rates are considered significantly different if their 95% confidence intervals indicated with error bars do not overlap. For fluctuation analysis of *NSR1*-deletion strains expressing different Nsr1 protein constructs, cells were transformed with pADH1-Nterm Nsr1, pADH1-Cterm Nsr1, pADH1-Nsr1, or pRS426 as a vector control. Twelve individual colonies per strain were used to inoculate 1 mL cultures in synthetic complete media lacking uracil and containing 2% glucose (SCD-Ura) and grown at 30°C. After 4 days, cultures were washed and diluted appropriately and then plated on agar plates containing either SCD-Ura for determination of total CFU or SCD-Ura/-Lys for determination of Lys+ recombinants. Recombination rates were determined as described above. Where applicable, doxycycline hyclate (Sigma) was added to the growth media to the concentration of 2 μg/ml.

## **2.9 Top1-SUMO Pull Down**

pGPD2-His-SMT3 was constructed by cloning the PCR-amplified His-SMT3 sequence from the pYlplac211-ADH-His-SMT3 plasmid (a gift from Stefan Jentsch lab) into the pGPD2 vector (Addgene Cat# 43972) using BamH1 and Xma1 restriction sites. pGPD2-His-SMT3 was transformed into *wss1Δ* strains harboring the indicated *TOP1* alleles. Three to five colonies from SCD-ura selection plates containing 2% glucose were used to inoculate 5 mL of SCD-ura liquid media containing 2% glucose and were grown overnight at 30 °C. The next day, 2.5 mL

of overnight culture was added to 50 mL SCD-ura liquid media containing 2% glucose and cultures were grown at 30 °C shaking until they reached an OD<sub>600</sub> of ~1. Then, a total of 50 OD<sub>600</sub> of cells were collected via centrifugation at 4 °C and froze at -80 °C after washing twice with H<sub>2</sub>O. The pull down of SUMO-conjugated proteins from whole cell lysates using HisPur Ni-NTA Resin (Thermo Scientific; Cat# 88221) was performed as previously described (Ohkuni et al., 2015). Input and pull down samples were resolved by SDS PAGE and transferred to PVDF membrane with a Trans-BlotSD cell machine (Bio-Rad; Cat# 170-3940). Blots were probed with an α-FLAG-HRP antibody (Sigma; Cat# A8592) and visualized by treatment with ECL substrate (GenDEPOT; Cat# W3680-010) and a Bio-Rad ChemiDoc MP imaging system.

### **2.10 Co-Immunoprecipitation (co-IP) Experiments.**

For each sample, 500 µL of saturated overnight culture was added to 50 mL of fresh YPD media and grown shaking at 30 °C until they reached an OD<sub>600</sub> of 1-1.5. A total of 44 OD<sub>600</sub> of cells were collected by centrifugation at 4 °C and washed once in lysis buffer C (50 mM Tris pH 8, 150 mM NaCl, and 7 mM EDTA) before pellets were frozen at -80 °C. Pellets were lysed by the addition of 500 µL of lysis buffer D (lysis buffer C with 5 mM DTT, 1 mM PMSF, and fungal protease inhibitor cocktail (Sigma; Cat# P8215) (50 µL/ 10 mL)) and acid washed glass beads followed by 4 rounds of bead beating for 30 sec followed by 5 min of incubation on ice. Next, lysates were clarified by centrifugation at 4 °C and 500 µL of lysis buffer D was added to each sample. Fifty µL of equilibrated anti-Flag M2 Affinity Gel Beads (Sigma; Cat# A2220) were added to each lysate and samples were incubated for 1 hr rotating at 4 °C. Anti-Flag M2 Affinity Gel Beads were then collected by centrifugation at 8,000 x g for 30 sec and washed 4 times. Each wash consisted of 1) centrifugation at 8,000 x g for 30 sec, 2) removal of

supernatant, 3) addition of 1 mL wash buffer E (lysis buffer D with 0.75% Triton X-100), and then rotation at 4 °C for 25 min. After the last wash, beads were resuspended in 20  $\mu$ L of lysis buffer D. Proteins were eluted from the anti-Flag M2 Affinity Gel Beads by adding 20  $\mu$ L of 200  $\mu$ g/mL 3X-FLAG peptide (Sigma; Cat# F4799) and then incubating for 30 min at room temperature while rotating. After elution, samples were centrifuged and 30  $\mu$ L of supernatant was collected and placed into a fresh 1.5 mL epi tube. Fifteen  $\mu$ L of 4X SDS sample buffer was added to the collected supernatants and samples were boiled for 10 min at 95 °C before being resolved on a 4-20% SDS PAGE gel (Bio-Rad; Cat# 4561093) and subjected to western blotting. Quantitative and statistical analyses of the co-IP were carried out as indicated for the oligo pull-down procedure above.

# **Chapter 3: Quantification of G4 Structures in Yeast in the Absence of Top1**

### 3.1 Introduction to G4-Detection Methods

Non-B form DNA structures called G4s are known to form in certain repetitive DNA sequences (Burge et al., 2006). Formation of G4s usually requires some amount of duplex melting to allow for guanines to base pair with one another through Hoogsteen bonding. Because of this, levels of helical torsion within DNA are thought to regulate G4-formation, where increased levels of negatively supercoiled DNA yield conditions that favor the formation of non-B DNA structures (Napierala et al., 2005; Sun and Hurley, 2009). Transcription is an essential DNA transaction that generates negative helical stress, and thus, non-B DNA structure formation is correlated with transcription levels (Wang and Vasquez, 2014). As mentioned in Chapter 1, yeast experiments demonstrated that the elevated recombination observed at a model G4-capable motif (*S $\mu$ G4-GTOP*) relative to a guanine-rich motif not prone to G4-formation (*S $\mu$ G4-GBTM*) is completely dependent on transcription (*S $\mu$ G4* recombination reporters depicted in Figure 5) (Kim and Jinks-Robertson, 2011). This suggests G4-formation does occur during transcription. Bioinformatics predicts there are > 1,400 G4-motifs present in the yeast genome (Todd et al., 2005), but it remains unknown how many G4 structures stably form in yeast *in vivo*. Other data listed in Chapter 1 support that activity of the enzyme Top1 modulates G4-formation during transcription by removing DNA negative supercoils (Kim and Jinks-Robertson, 2011; Husain et al., 2016). These data allude to the model where significantly more G4s form in highly-transcribed genomic loci in yeast cells lacking Top1 than in normal cells. However, this postulate remains to be tested on a genome-wide scale. Therefore, the goal of the work described in this chapter was to uncover if a



greater number of G4s form in the yeast genome in the absence of Top1 by scoring stably-folded G4s *in vivo*.

There are many different techniques used to detect G4-formation both *in vitro* and *in vivo*. Gellert et al., 1962 first observed guanine tetrad formation in a gel-like substance produced from cooling a high concentration guanosine solution (Gellert et al., 1962). X-ray crystallography is a technique that solves 3D structures using diffraction patterns of X-rays passed through crystals to map the electron density of molecules (Smyth and Martin, 2000). When Gellert et al., 1962 used this technique to analyze the guanosine gel, a helical structure containing 4 units per turn was revealed (Gellert et al., 1962). Since Jerry Donohue had uncovered that 2 guanines can base-pair with each other, Gellert et al., 1962 proposed that 4 guanines interact to form planar guanine tetrads (Donohue, 1956; Gellert et al., 1962). Since the discovery of the guanine tetrad, X-ray crystallography has been used to solve the 3D structures of folded-G4s formed in G-rich oligos (Parkinson and Collie, 2019). Another *in vitro* technique called circular dichroism (CD) is commonly used to discern the structural conformations of G4-capable DNA oligo nucleotides (Kypr et al., 2009). In CD experiments, a single wavelength of circularly polarized light is passed through a solution containing DNA fragments that absorb different amounts of right circularly polarized light and left circularly polarized light. Measuring the differential absorption of left and right circularly polarized light, called molar ellipticity, across different wavelengths allows for the secondary structures of nucleic acids to be inferred in solutions. For example, guanine-rich DNA oligos that form parallel G4s exhibit CD spectrum results with a positive peak in molar ellipticity at 260 nm and a negative peak at 240 nm (Balagurumoorthy et al., 1992). Another technique used to

determine if guanine-rich DNA oligos adopt G4-conformations *in vitro* is the polymerase stop assay (Weitzmann et al., 1996; Han et al., 1999). This method is based on the finding that stable G4-structures block progression of polymerases during *in vitro* DNA synthesis (Weitzmann et al., 1996). In polymerase stop assays, a primer labeled at its 5'-end is annealed to a DNA template followed by the addition of a purified DNA polymerase enzyme. After the primer-extension reaction has occurred, the 5'-end labeled extension products are resolved on a urea-containing denaturing PAGE. If G4-forms in the DNA template strand, the labeled primer will only be extended up to the G4-structure, resulting in detection of an extension product that is smaller than the expected full-length extension product. G4-formation *in vitro* can be further confirmed by performing CD and polymerase stop assay experiments in the presence of  $K^+$  and  $Li^+$  ions, where  $K^+$  cations are known to support G4-formation where  $Li^+$  cations are known to have a destabilizing effect on G4s (Konig et al., 2010).

Like CD and polymerase stop assays, dimethyl sulfate (DMS) foot printing is another method to detect G4-formation. DMS methylates the N7 atom of guanines in B-form DNA, but in folded G4s, the N7 atom of guanines participate in Hoogsteen bonding and thus, are protected from DMS methylation (Dabrowiak et al., 1997). Following DMS treatment of DNA, piperidine is used to cleave the DNA at sites of methylated purines and cleavage products can be analyzed by resolution on a 16% polyacrylamide denaturing gel to determine G4-formation and -stability. Notably, while DMS footprinting can be done to probe *in vitro* G4-formation of oligos and plasmids, it can also measure the *in vivo* formation of G4s at targeted loci via ligation-mediated PCR, a method where a DNA linker with a known DNA sequence is

annealed to the 5' ends of target DNA fragments (Drouin et al., 2001). One study treated human lung epithelial carcinoma cells with DMS to demonstrate *in vivo* G4-formation at the *VEGF* promoter (Sun et al., 2011). Following DMS treatment, the cells were lysed and extracted DNA was treated with piperidine. The genomic DNA cleavage products were then subjected to ligation-mediated PCR, allowing for amplification of *VEGF* promoter specific DNA fragments that were resolved by electrophoresis and analyzed. While *in vivo* DMS-foot printing experiments have led to findings in the G4-field of study, like the aforementioned validation of *in vivo* *VEGF* promoter G4-formation, this technique yields results that require lengthy interpretation times. Further, *in vivo* DMS foot printing can only reasonably be used to investigate individual genomic G4-loci rather than G4-formation on a genome-wide scale. Therefore, newer *in vivo* methods of G4-detection have focused on analyzing global genomic G4-formation in cells.

For the visualization of G4s in cells, fluorescent G4-probes have been developed (Zhang et al., 2018; Chilka et al., 2019). One small molecular G4-probe called Naphtho- template-assembled synthetic G-quartet (N-TASQ) is both a smart fluorophore and G4-binding ligand (Laguerre et al., 2014). In solution, the N-TASQ molecule exists in an “open conformation”, where the molecule’s synthetic guanine attachments are extended outward from the fluorescent-capable center of N-TASQ, preventing fluorescence (Laguerre et al., 2016). Upon G4-binding, N-TASQ adopts a “closed conformation” where the synthetic planar guanine tetrads of NTASQ undergo pi-stacking with the planar guanine tetrads of folded G4s. In this closed conformation, the N-TASQ molecule is able to fluoresce. This G4-specific “light-up probe” has been used in multiple studies to show the *in vivo* formation of G4s in nuclei of

different cell types including murine melanoma, human breast cancer, and human osteosarcoma cells (Zhang et al., 2018). Notably, N-TASQ is also being pursued as a potential diagnostic and site-specific therapeutic delivery tool in cancers or other diseases marked by increased potential for G4-formation (Chilka et al., 2019).

In addition to small fluorescent G4-probes, G4-specific antibodies have been developed to assess *in vivo* G4-formation as well. One of the first G4-specific antibodies, Sty49, was used to visualize telomeric G4-formation in the macronuclei of the ciliated protozoan *Stylonychia lemnae* (Schaffitzel et al., 2001). While these initial experiments demonstrated that G4s do form in the chromosomes of organisms, the specificity of Sty49 to the telomeric G4s of *S. lemnae* prevents its usage in investigating genome-wide G4-formation in other species. Since the development of Sty49, G4 antibodies that bind more than just telomeric G4s in specific organisms have been engineered. Currently, the most widely-used G4-specific antibody is BG4 (Biffi et al., 2013), a monoclonal single-chain variable fragment (ScFV) antibody where the variable region of an immunoglobulin heavy chain is fused to the variable region of an immunoglobulin light chain via a short peptide linker (Huston et al., 1988). BG4 was identified in a phage-display screen using a library of  $2.3 \times 10^{10}$  ScFV antibodies (Schofield et al., 2007) aiming to find a G4-specific antibody that could interact with a broad range of G4-structure types (Biffi et al., 2013). After BG4 was shown to interact *in vitro* with a panel of intramolecular G4-oligos in a G4-selection screen, the BG4 coding sequence was then cloned into the pSANG10 *E. coli* expression vector for antibody purification. Enzyme-linked immunosorbent assays (ELISAs) conducted with G4-oligos immobilized on a 96-well plate and purified BG4 revealed that BG4 binds intramolecular and

intermolecular G4s, exhibiting Kds of 0.5-1.6 nM and ~2 nM, respectively. Importantly, binding of BG4 to a duplex DNA oligo, a single-stranded DNA oligo, or an RNA hairpin was not detected in ELISA experiments, demonstrating BG4 is highly specific for G4 DNA. Further ELISA analyses revealed that BG4 is capable of high affinity interactions with G4s of different structural conformations; i.e. parallel, anti-parallel, and hybrid G4s. Thus, it was concluded that BG4 is a highly specific antibody that binds to a wide-range of G4 structures.

Since its discovery, BG4 has been used to visualize G4s in the nuclei of cells in immunofluorescence experiments. Specifically, Biffi et al 2013 first used FLAG-tagged BG4 and a secondary fluorescent  $\alpha$ -FLAG antibody to visualize G4-puncta in the nuclei of fixed human osteosarcoma cells. It was confirmed that the puncta recognized by BG4 were G4-DNA structures since incubation of the cancer cells with either G4-capable oligonucleotides or DNase I abrogated the detection of nuclear BG4 foci. Quantitative BG4-immunofluorescence of synchronized human mammary adenocarcinoma cells revealed that G4s form in all phases of the cell cycle, with the most BG4-G4 foci visible in S-phase. Further demonstrating that BG4-immunofluorescence can be quantitative, treatment of osteosarcoma cells with the G4-stabilizing ligand pyridostatin (PDS) increased the number of BG4-G4 foci present in nuclei by 2.9-fold.

While immunofluorescence experiments with BG4 have allowed for quantification of G4s in cells treated with G4-stabilizing ligands and in differing cell cycle stages, these experiments do not provide much detailed information about the locations of stably folded G4s in the genome outside of well-studied genetic elements like telomeres (Biffi et al., 2013). Multiple bioinformatic studies have mapped the locations of potential G4-forming sequences

(PQS) or G4-motifs in genomes and uncovered G4-motifs are enriched in G-rich micro- and mini-satellites, recombination hot spots, TSSs, promoters, telomeric DNA, and ribosomal DNA (Todd et al., 2005; Marsico et al., 2019; Rawal et al., 2006; Eddy and Maizels, 2008). In the human genome, one computational study estimated the presence of >375,000 G4-motifs comprised of the sequence GGGN<sub>1-7</sub>GGGN<sub>1-7</sub>GGGN<sub>1-7</sub>GGG (Huppert and Balasubramanian, 2005). This study also uncovered that there are limited numbers of G4-motifs in exonic coding strands, suggesting G4-formation is disfavored in RNA-forming DNA sequences in humans. A separate bioinformatic study that took advantage of variations of next-generation sequencing quality at unstable G4-capable loci found 736,689 G4-motifs in primary human B lymphocytes (Tu et al., 2021). This higher number of > 700,000 G4-motifs relative to the ~ 375,000 G4-motifs found in the previous study scanning the human genome for the GGGN<sub>1-7</sub>GGGN<sub>1-7</sub>GGGN<sub>1-7</sub>GGG sequence, considered to be the canonical G4-motif (Davis and Maizels, 2011; Bochman et al., 2012), was the result of identification of non-canonical G4-motifs containing long loops, bulges, and two tetrads. Around 90% of the G4-motifs identified in this modified sequencing analysis method were confirmed to form G4s *in vitro* by G4-seq, a method where next generation sequencing is performed after high-throughput polymerase stop assays conducted with purified genomic DNA (Chambers et al., 2015). As mentioned above, bioinformatics estimates 1,400 G4-motifs in the *S. cerevisiae* genome (Todd et al., 2005) that are enriched in gene promoters (Capra et al., 2010). However, when G4-seq was applied to *S. cerevisiae*, only 143 polymerase stalling G4-sites were identified (Marsico et al., 2019).

While the studies described above yielded useful results that expanded the definition of G4-motifs, the bioinformatic information regarding locations of G4-motifs in genomes does not prove *in vivo* G4-formation, especially since transcriptional state of chromatin affects the ability of G-rich DNA sequences to convert to G4-conformations (Kim et al., 2011; Kim, 2019). Therefore, ChIP-seq experiments conducted with G4-specific antibodies provided valuable insights into the genomic locations where *bona fide* G4-formation occurs in cells. Hänsel-Hertsch et al., 2013 used purified FLAG-tagged BG4 in ChIP-seq experiments to show that ~10,000 G4s form *in vivo* in immortalized human keratinocytes. The locations of stably-folded G4s overlapped with nucleosome-depleted genomic regions identified by next-generation sequencing following formaldehyde-assisted isolation of regulatory elements, demonstrating that *in vivo* G4-formation relies on an open chromatin state. G4s were also enriched in promoters and 5' UTRs of genes, especially in cancer-associated genes, in line with other findings showing that G4s have roles in transcriptional regulation (reviewed in Kim, 2019). This work also provided two pieces of evidence supporting that G4-formation is dependent on transcription. First, when cells were treated with the histone deacetylase inhibitor entinostat, which leads to constitutive acetylation of histone H3K27 and thus sustained transcriptional activity (Saito et al., 1999; Tomazou et al., 2015), ~ 4,000 intensified or new BG4-ChIP-seq peaks were detected. Second, BG4 ChIP-seq experiments in non-immortalized human keratinocytes showed ~ 10-fold less G4s relative to immortalized cells with abnormally elevated transcriptional activity. In summary, Hänsel-Hertsch et al., 2013 not only uncovered important information regarding *in vivo* G4-formation, but also showed that

BG4 ChIP-seq can be used to quantify and compare the number of G4s that form in genomes under different cellular conditions.

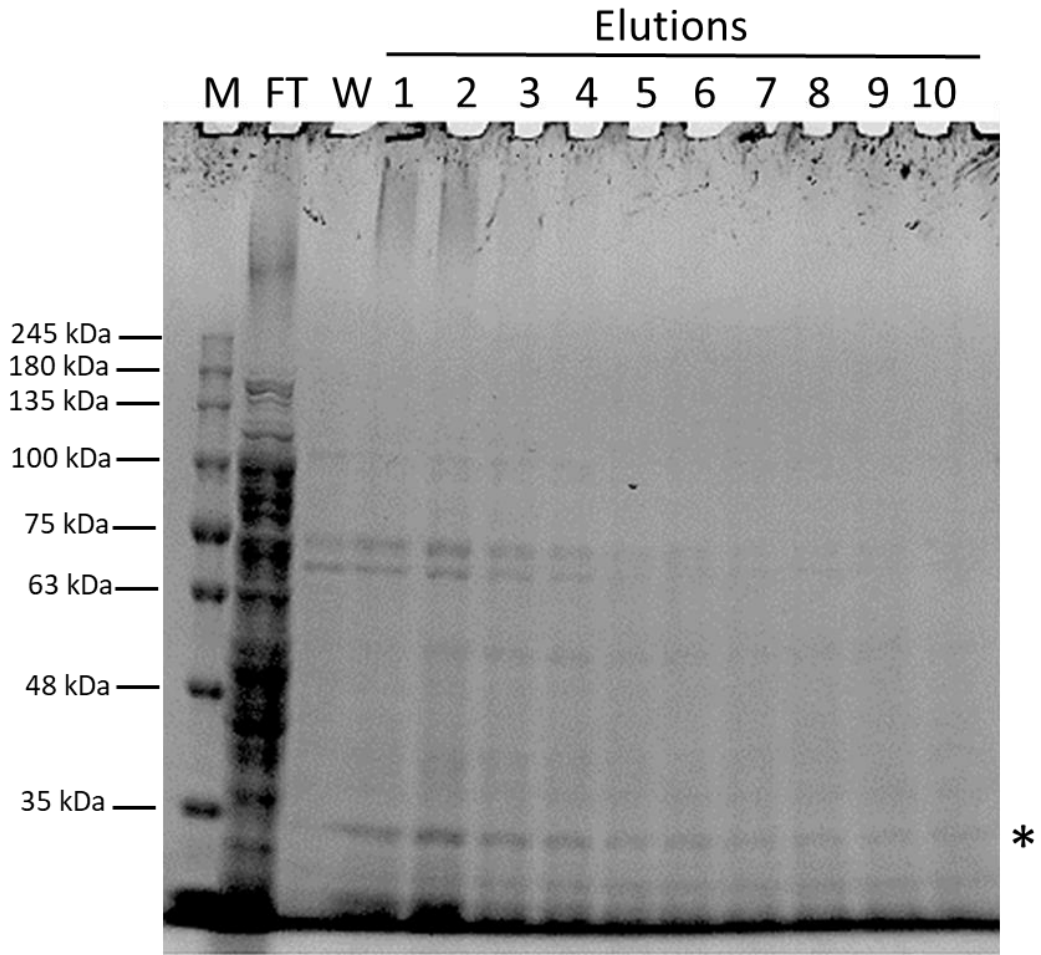
Based on previous results from Kim lab showing that *TOP1*-deletion enhances G4-induced genomic instability (Kim and Jinks-Robertson, 2011; Yadav et al., 2014; Yadav et al., 2016), I set out to quantify the number of G4s that form in the genomes of *WT* and *top1Δ* yeast cells through expression of the G4-specific antibody BG4 for immunofluorescence and ChIP-seq experiments. I hypothesized that significantly more G4 structures would be able to be detected *in vivo* in *top1Δ* yeast cells relative to *WT* yeast cells.

## **3.2 Results**

### **3.2.1 Construction and Validation of Yeast BG4 Expression Plasmid.**

I initially set out to use purified BG4 in immunofluorescence and ChIP experiments to visualize and quantify the number of G4s present in yeast. However, my protein preps of His-tagged BG4 expressed in and purified from *Escherichia coli* BL121 (DE3) cells yielded low amounts of BG4 (Figure 6). Additionally, while my purification elutions did contain a ~ 30 Kda protein that corresponds with the expected size of BG4, they also contained multiple other contaminating proteins indicating that the BG4 protein purification protocol needs optimization (Figure 6).





**Figure 6. BG4 antibody purified from *E. coli*.** Coomassie stained SDS PAGE gel of His-tagged BG4 purified from BL21 (DE3) *E. coli* cells using HisPur™ Ni-NTA Resin (ThermoFisher). M = protein marker. FT = flow through. W = wash. \* denotes location of BG4 band.

Therefore, rather than using purified BG4, I tried out a new approach of BG4-G4 detection by expressing the BG4 antibody in yeast cells. To do this, I created a BG4 expression plasmid called pGAL-BG4-FLAG. The pGAL-BG4-FLAG construct consists of the C-terminally His- and FLAG-tagged BG4 protein under control of a galactose inducible promoter (pGAL) on a 2 $\mu$  yeast expression plasmid with a *URA3* selectable marker (Figure 7). After plasmid construction, I transformed yeast with pGAL-BG4-FLAG and showed that BG4 is expressed in *WT* and *top1 $\Delta$*  cells in the presence of galactose through western blotting (Figure 8). In order to test whether overexpression of BG4 is toxic to cell growth in yeast, I performed a spot assay to measure the survival of *WT* and *top1 $\Delta$*  cells transformed with pGAL-BG4-FLAG or a vector control on synthetic complete -URA (SCD-URA) plates containing either glucose or galactose (Figure 9). I found that both *WT* and *top1 $\Delta$*  cells transformed with pGAL-BG4-FLAG grew as well as the cells with the vector control on plates containing galactose ranging in concentrations from 0.01% to 2%. I also did not observe any differences in the survival of cells transformed with pGAL-BG4-FLAG between the plates containing glucose where BG4 expression is suppressed and the plates containing galactose where BG4 expression is induced.

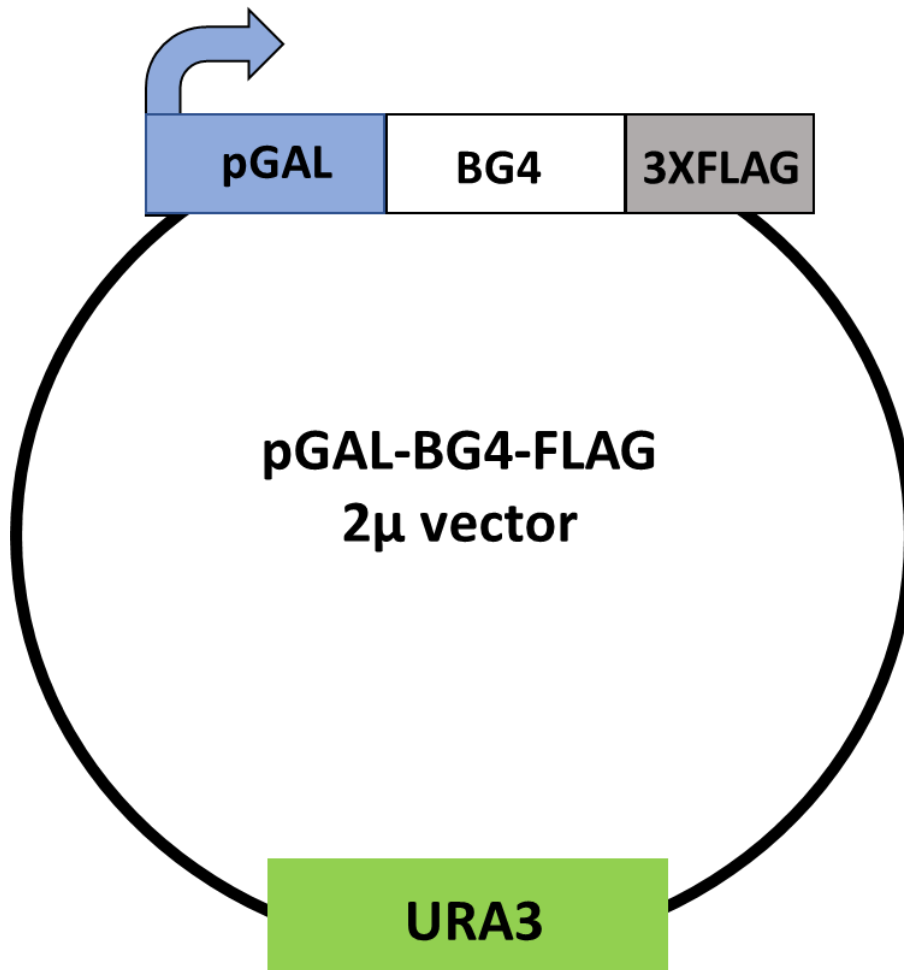
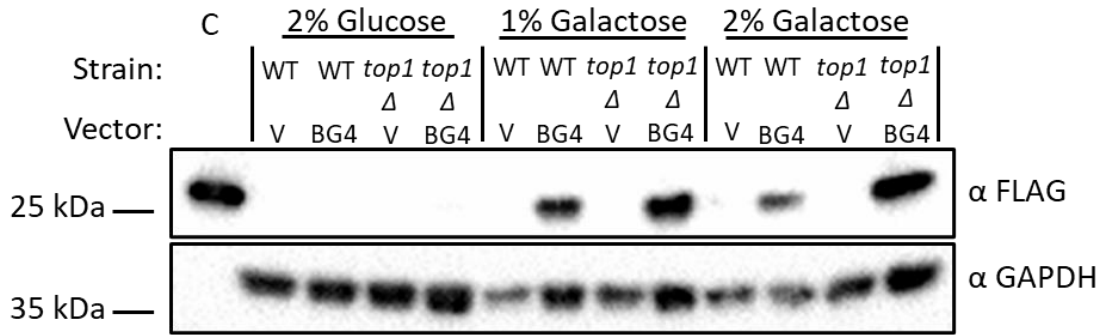
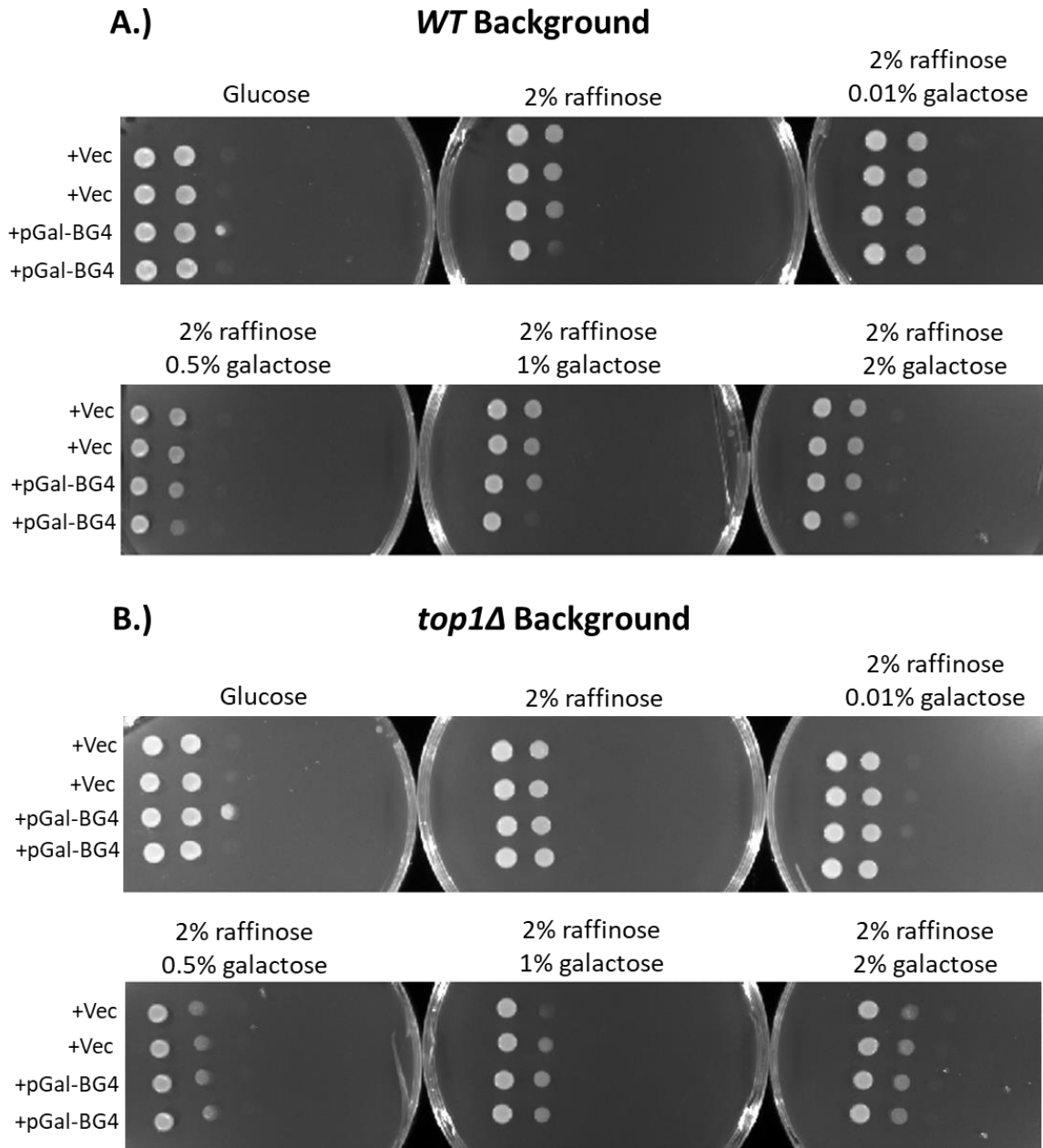


Figure 7. Schematic of pGAL-BG4-FLAG vector for expression of BG4 in yeast.



**Figure 8. BG4-FLAG antibody construct is expressed from the pGAL-BG4-FLAG vector in yeast.** Western blot of lysates prepared from *WT* and *top1* $\Delta$  cells transformed with control and pGAL-BG4-FLAG vectors. Blots were probed with an  $\alpha$ -FLAG antibody to detect BG4 or an  $\alpha$ -GAPDH antibody for a loading control. C = purified BG4 (size control). V = vector control. BG4 = pGAL-BG4-FLAG.



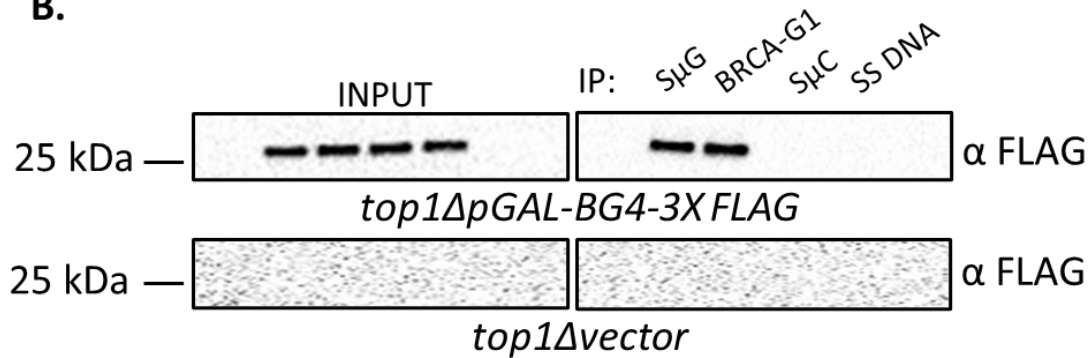
**Figure 9. Expression of BG4 does not affect survival of yeast.** Survival assays of yeast cells transformed with pGAL-BG4-FLAG (pGal-BG4) or vector control (Vec) in media with different carbon sources. Cells were diluted 1:10 serially before spotting on plates. Images shown were taken 3 days after growth at 30 °C. A. Survival of *WT* cells. B. Survival of *top1Δ* cells.

Next, in order to verify if expressed BG4 was functional and able to bind to G4 structures, I performed an *in vitro* oligo binding assay. First, either biotinylated G4-capable oligonucleotides (oligos) (S $\mu$ G and BRCA-G1) or biotinylated non-G4 control oligos (S $\mu$ C and ssDNA) conjugated to streptavidin magnetic beads were incubated with yeast whole cell extracts prepared from cells transformed with pGAL-BG4-FLAG (Figure 10A). After pull down of the magnetic beads, western blotting of pull down samples revealed that BG4 bound the S $\mu$ G and BRCA1-G1 G4 oligos but not the M1 control oligo (Figure 10B), demonstrating that the BG4 construct expressed in yeast cells is functional and G4-specific.

**A.**

Oligonucleotide	Sequence
S $\mu$ G	5' GAGCT <b><u>GGGG</u></b> TGAGCT <b><u>GGG</u></b> CTGAGCT <b><u>GGGG</u></b> TGAGCT <b><u>GGG</u></b> CTGAGCT
BRCA-G1	5' A <b><u>GGG</u></b> CTCTGCCTT <b><u>GGGGGGGGG</u></b> CAGGA <b><u>GGGA</u></b>
S $\mu$ C	5' AGCTCAGCCAGCTC <b><u>ACCCC</u></b> AGCTCAGCCAGCTC <b><u>ACCCC</u></b> AGCTC
ssDNA	5' GCACGCGTATCTTTTGGCGCAGGTG

**B.**



**Figure 10. BG4 expressed in yeast binds G4s *in vitro*.** A. Table of oligos used in biotin oligo pull down in Figure 10B. S $\mu$ G and BRCA-G1 oligos form G4s while S $\mu$ C and ssDNA do not. Guanines that participate in G4-formation are in bold, underlined, and italicized. B. Western blot from biotinylated oligo pull down assay performed with streptavidin magnetic beads and yeast whole cell lysates expressing BG4-FLAG from pGAL-BG4-FLAG or the vector control. Blots were probed with an  $\alpha$ -FLAG-HRP (Sigma) antibody.

### 3.2.2 Measuring Expressed BG4-G4 Structure Binding *in vivo*.

Next, in order to determine the number and locations of G4s that form in the yeast genome, I aimed to perform chromatin immunoprecipitation followed by next generation sequencing (ChIP-seq) in *WT* and *top1Δ* cells transformed with pGal-BG4-FLAG. I expected to observe a higher number of BG4-FLAG enrichment peak locations in cells lacking Top1 compared to *WT* cells throughout the yeast genome. Additionally, I expected to uncover a positive correlation between stably-folded *in vivo* G4s and highly transcribed genes, where lowly transcribed genes harboring G4-motifs would exhibit less *in vivo* G4-formation than highly transcribed genes with G4-motifs.

Before performing ChIP-seq in cells transformed with pGAL-BG4-FLAG, I first wanted to validate that BG4 is enriched at the model G4-forming *SμG4-GTOP* locus in *top1Δ* cells. Thus, I transformed *top1Δ* yeast cells containing the *SμG4-GTOP* reporter with either pGAL-BG4-FLAG or the vector control, and performed ChIP-qPCR using primers that target the *SμG4* locus and a negative control *CAN1* locus. I expected to observe increased fold enrichment (*G4* enrichment/ *CAN1* enrichment) of BG4 at the *SμG4-GTOP* locus compared to vector control, which would indicate BG4-G4 DNA *in vivo* binding. However, ChIP-qPCR results did not show increased enrichment of BG4 at the *G4* locus relative to the vector control (Figure 11). Therefore, I did not move on to perform BG4 ChIP-seq with the expressed pGAL-BG4-FLAG construct.



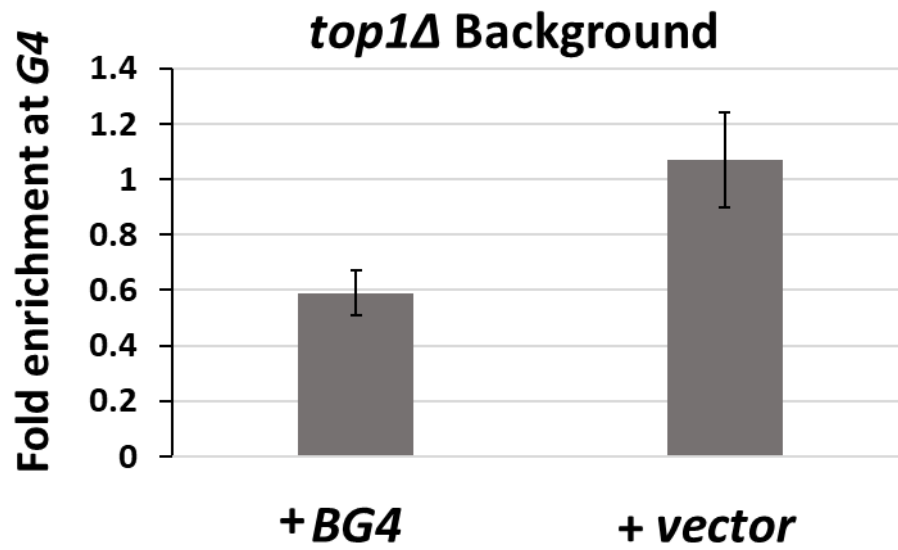


Figure 11. Fold enrichment of BG4 at *SμG4 G4* locus relative to the non-G4 capable *CAN1* locus *in vivo*. Averages and standard deviations of BG4 and vector control fold enrichments at *G4* locus in a *top1Δ GTOP* background (n=2).

### 3.2.3 Visualizing G4 Structures in Yeast *in vivo*.

As an alternate approach to quantify the number of G4 structures that form in *WT* and *top1Δ* yeast cells, I performed immunofluorescence experiments with yeast cells transformed with either pGal-BG4-FLAG or the vector control. With the help of Dr. Karan Kaval in the Garsin lab at UTHealth, a GFP-conjugated  $\alpha$ -FLAG antibody (Dylight 488  $\alpha$ -FLAG antibody; ThermoFisher Scientific) was used to detect  $\alpha$ -FLAG/BG4 foci in *top1Δ* cells. We expected to detect multiple  $\alpha$ -FLAG/BG4 foci within the DAPI-stained nuclei of cells expressing pGal-BG4-FLAG (expected immunofluorescence results are depicted in Figure 12). However, we were not able to detect  $\alpha$ -FLAG/BG4 foci within the nuclei of pGAL-BG4 transformed cells (Figure 13). In fact, the pGal-BG4-FLAG cells did not look different from the negative control cells transformed with the empty vector, where no green foci were expected to be visualized (Figures 13 and 14). In both pGal-BG4-FLAG- and vector control-transformed cells, there were high levels of green fluorescent background signal (Figures 13C and 14C), indicating that my immunofluorescence protocol needs to be optimized to reduce background signal. Further, some cells in both pGAL-BG4-FLAG- and vector control-transformed samples seemed to have  $\alpha$ -FLAG puncta localized to the outside of the cell (Figures 13C and 14C), suggesting that the  $\alpha$ -FLAG antibody may not be efficiently getting into the cells and that cell permeabilization conditions should be optimized.

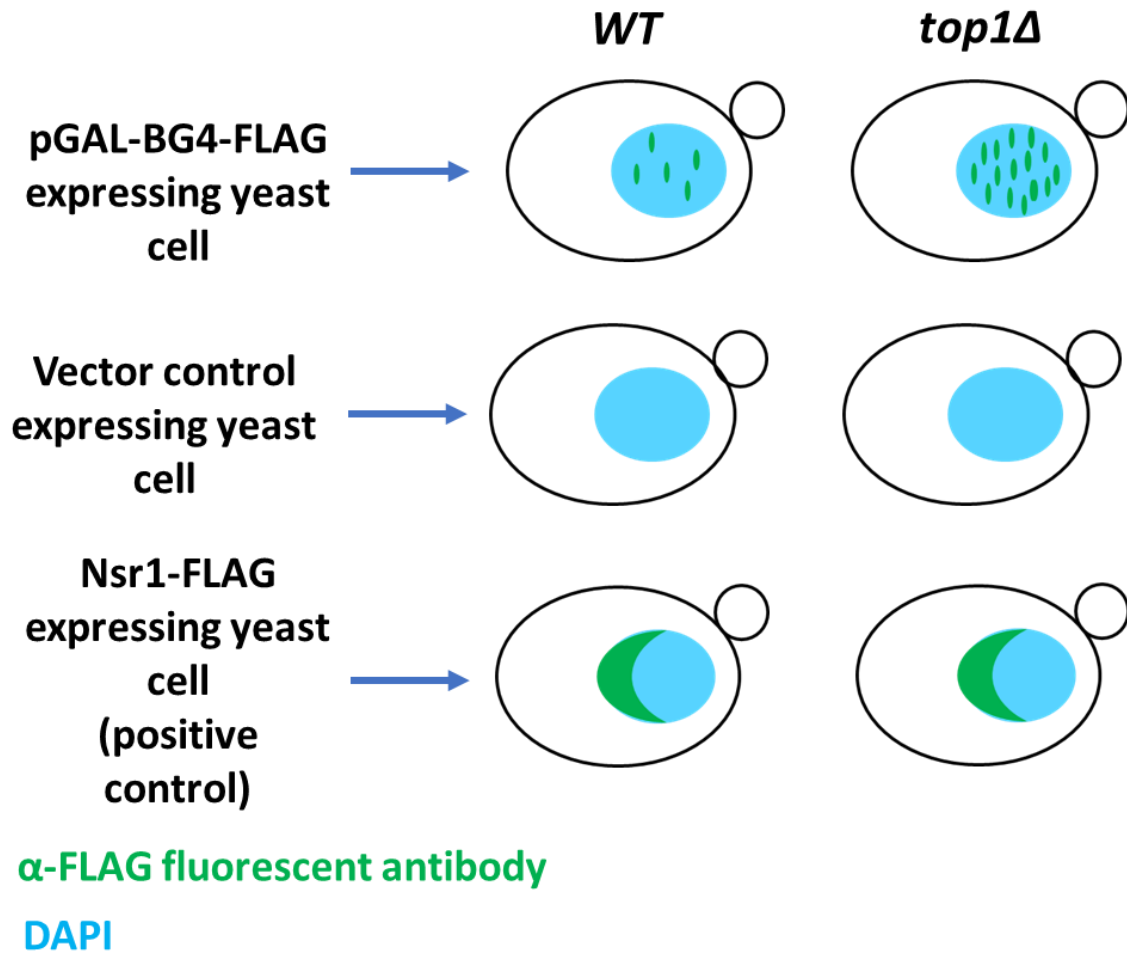
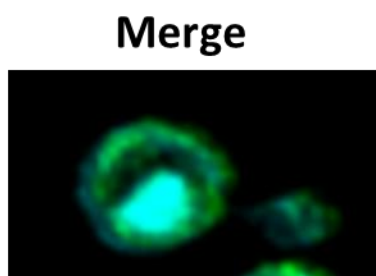
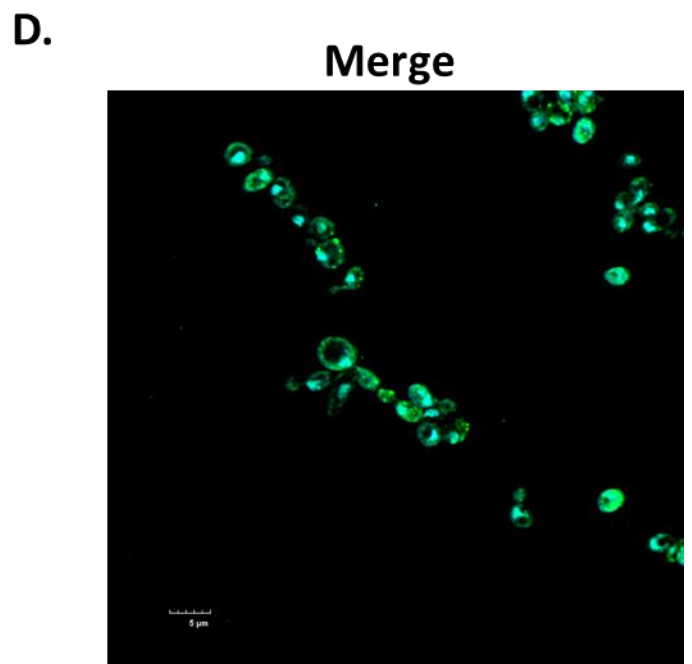
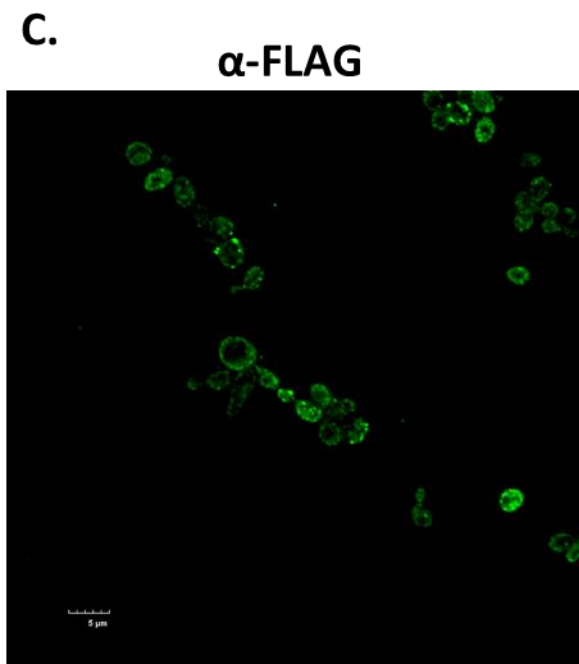
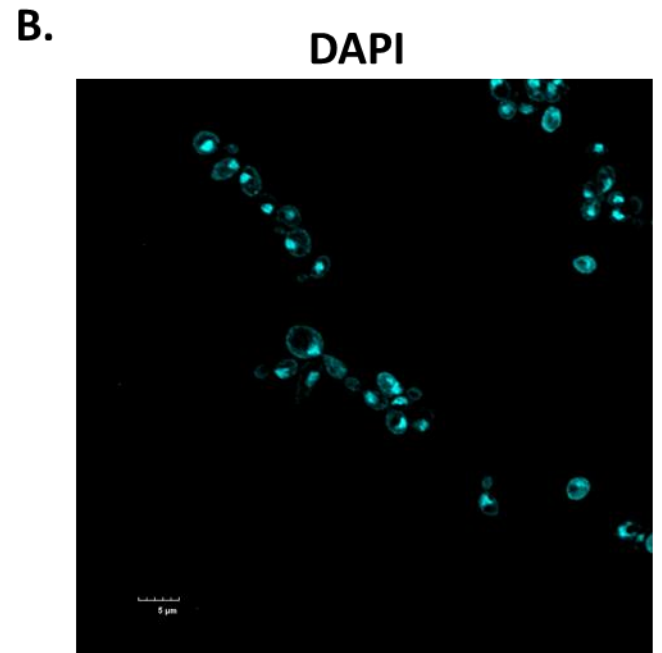
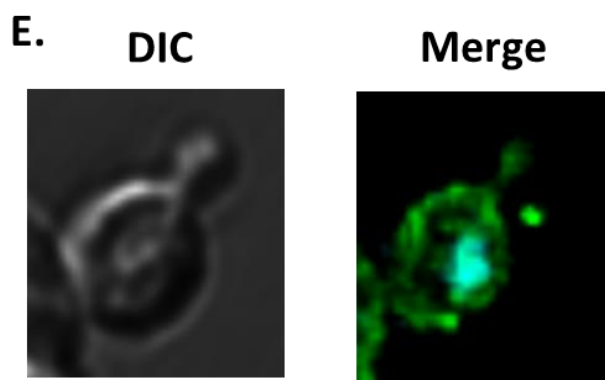
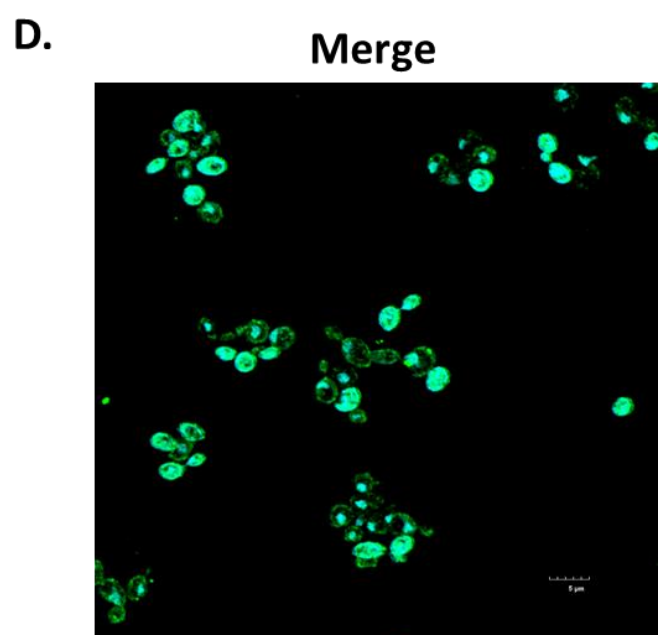
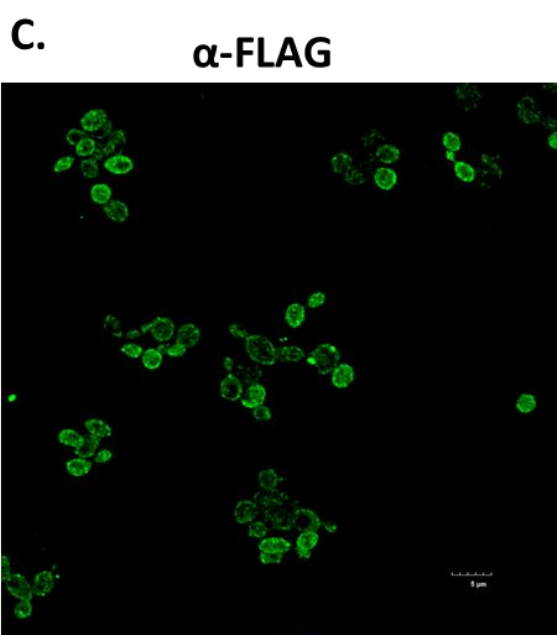
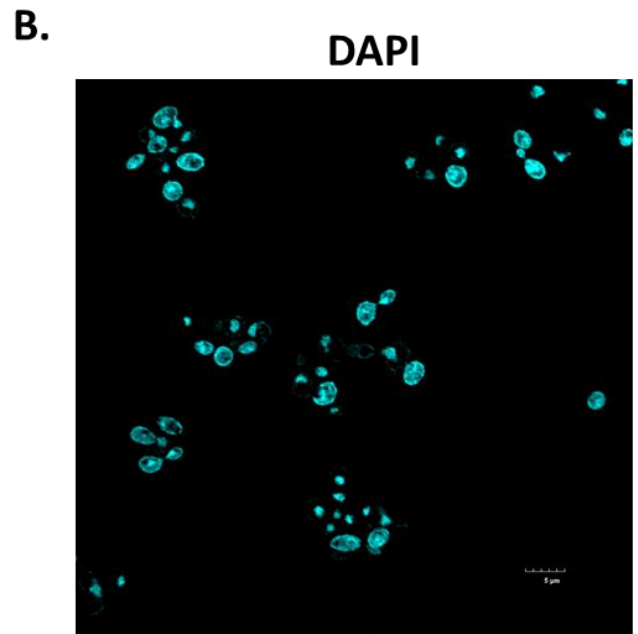
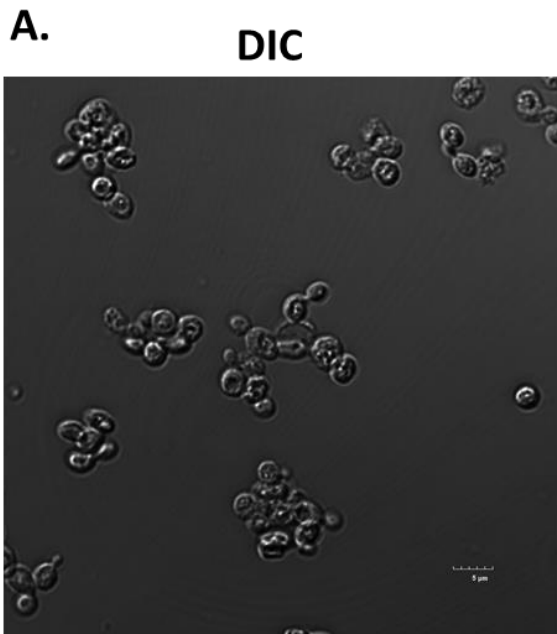


Figure 12. Expected immunofluorescence microscopy results of yeast cells expressing pGAL-BG4-FLAG, the vector control, or Nsr1-FLAG (positive control).



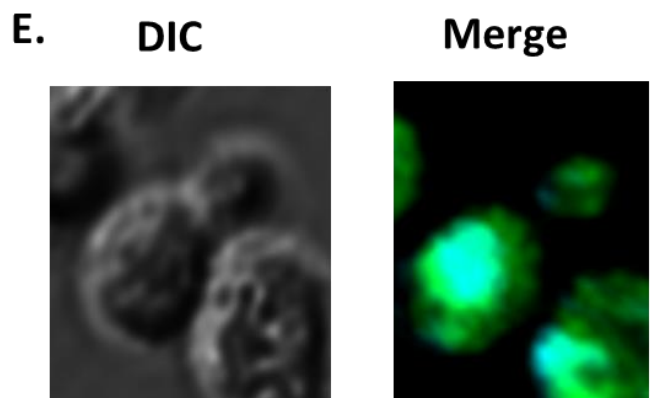
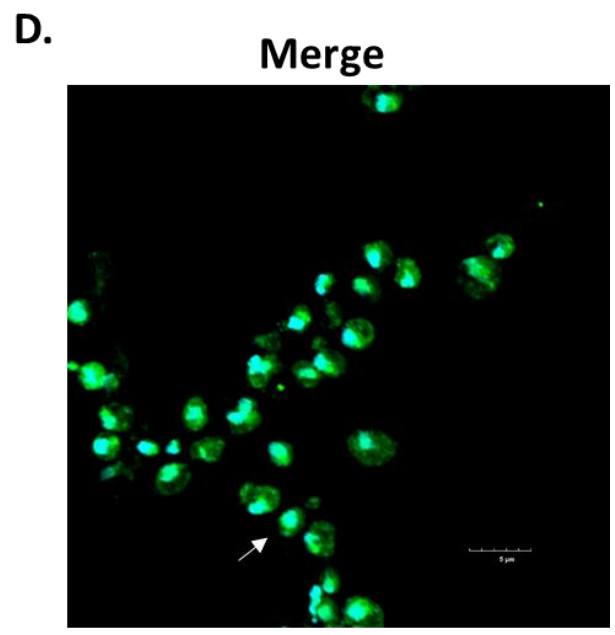
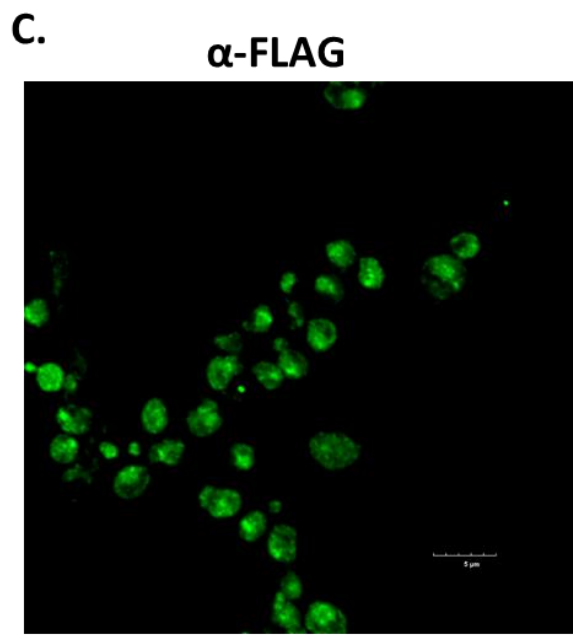
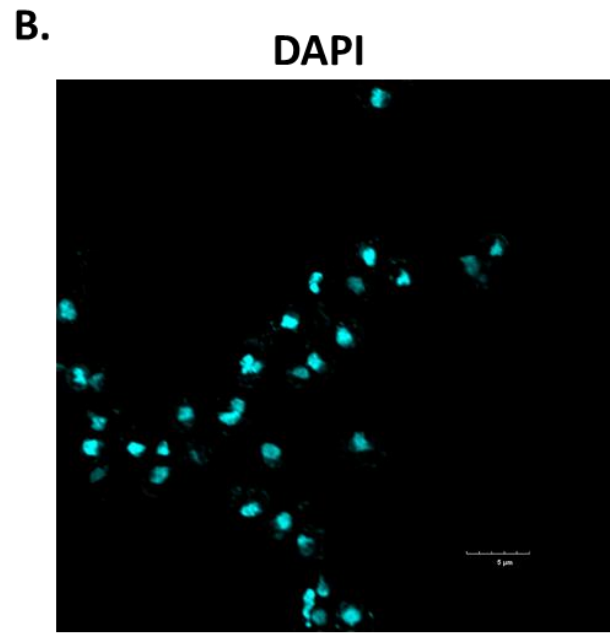
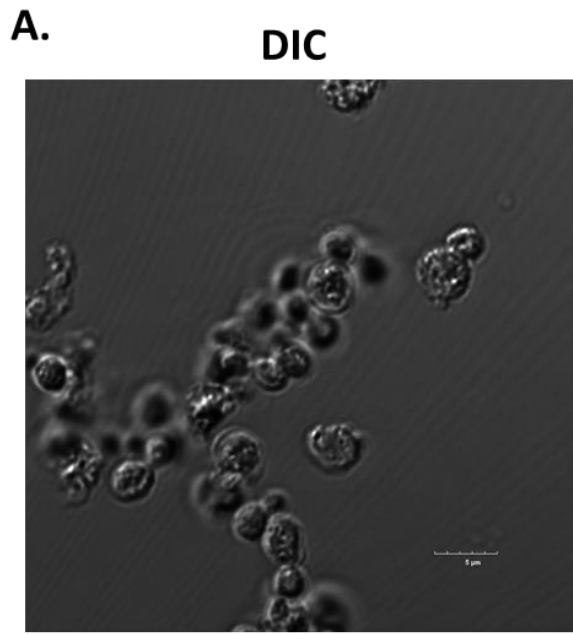
**Figure 13. Immunofluorescence microscopy of *top1Δ* yeast cells expressing pGAL-BG4-3XFLAG.** Slides were prepared by Alexandra Berroyer and were imaged by Dr. Karan Kaval with an Olympus Fluoview FV3000 confocal microscope. All scale bars are 5  $\mu\text{m}$ . *A.* Differential interference contrast (DIC) image. *B.* DAPI stain fluorescent image. *C.* Image of cells stained with an  $\alpha$ -FLAG Dylight 488 antibody (ThermoFisher). *D.* Merge of *B.* and *C.* *E.* Zoom in on yeast cell from *A* and *D.*



**Figure 14. Immunofluorescence microscopy of *top1Δ* yeast cells expressing the vector control.** Slides were prepared by Alexandra Berroyer and were imaged by Dr. Karan Kaval with an Olympus Fluoview FV3000 confocal microscope. All scale bars are 5  $\mu\text{m}$ . *A.* Differential interference contrast (DIC) image. *B.* DAPI stain fluorescent image. *C.* Image of cells stained with an  $\alpha$ -FLAG Dylight 488 antibody (ThermoFisher). *D.* Merge of *B.* and *C.* *E.* Zoom in on yeast cell from *A* and *D.*

To further check my immunofluorescence staining technique, I performed the protocol with positive control yeast cells that expressed a FLAG-tagged version of the nucleolar protein Nsr1 (Figure 15) (Ginisty et al., 1999). The yeast nucleolus occupies around one-third of the total volume of the nucleus and is visualized as a crescent shaped structure (Aris and Blobel, 1988; Matos-Perdomo and Machin, 2019). Thus, we expected to see colocalization of a crescent-shaped Nsr1-FLAG focus within the DAPI stained nucleus (Figure 12). However, except for one cell denoted by an arrow in Figure 15D, we did not observe colocalization of Nsr1-FLAG with DAPI and the background associated with the fluorescent FLAG antibody was very high. Overall, I concluded my immunofluorescence protocol needs to be modified to reduce background fluorescence, improve cell permeability, and increase antibody specificity.





**Figure 15. Immunofluorescence microscopy of yeast cells expressing FLAG-tagged Nsr1.**

Slides were prepared by Alexandra Berroyer and were imaged by Dr. Karan Kaval with an Olympus Fluoview FV3000 confocal microscope. All scale bars are 5  $\mu\text{m}$ . *A.* Differential interference contrast (DIC) image. *B.* DAPI stain fluorescent image. *C.* Image of cells stained with  $\alpha$ -FLAG Dylight 488 antibody (ThermoFisher). *D.* Merge of *B.* and *C.* Arrow denotes cell exhibiting DAPI and Nsr1-3XFLAG colocalization. *E.* Zoom in on cell denoted by arrow in *C.*

### 3.3 Discussion

Helical torsion in DNA causes complications during cellular DNA transactions (Gomez-Gonzalez and Aguilera, 2019). One source of torsional stress is transcription, which yields positively supercoiled DNA and negatively supercoiled DNA located in front of and behind RNA polymerase 2, respectively (Ma and Wang, 2014; Ma and Wang, 2016; Kim and Jinks-Robertson, 2017). While normal, unstressed DNA contains  $\sim 10.5$  bp per turn, positively supercoiled DNA is overwound and contains  $< 10.5$  bp per turn and negatively supercoiled DNA is underwound and contains  $> 10.5$  bp per turn. Both positively and negatively supercoiled DNA are problematic for cells. Extreme levels of positive supercoils can block the progression of RNA polymerase during transcription, resulting in genomic instability and accumulation of truncated messenger RNA molecules. Conversely, the accumulation of negative supercoils during transcription is deleterious because it leads to the formation of single-stranded patches in DNA and mutagenic R-loop formation. Single-stranded DNA, in either an R-loop or in a portion of melted duplex DNA, is more prone to DNA damage and base modification than duplex DNA (Lindahl, 1993; Beletskii and Bhagwat, 1996; Kim and Jinks-Robertson, 2012; Hamperl and Cimprich, 2014). Further, the formation of replication-blocking non-B DNA structures, like G4s, usually requires for duplex melting to occur which can favor DNA base interactions within a single strand of DNA (Sun and Hurley, 2009).

G4s normally form in single-stranded DNA sequences containing a GGGN1-7GGGN1-7GGGN1-7GGGN1-7 motif when guanines interact with one another through Hoogsteen bonding (Bochman et al., 2012; Burge et al., 2006). G4-formation happens during transcription and is linked to negatively supercoiled DNA (Sun and Hurley, 2009; Kim, 2019).

Top1 removes transcription-associated negative supercoils from DNA, and the genomic instability induced by multiple non-B DNA structures, including G4s, is repressed by Top1 functioning (Drolet, 2006; Kim and Jinks-Robertson, 2011; Hubert et al., 2011; Yadav et al., 2014; Yadav et al., 2016). This supports a model where Top1 suppresses G4-induced genomic instability in cells by preventing the accumulation of transcriptionally derived negative supercoils and consequential formation of G4s. Here, I set out to test this model by two approaches: ChIP-seq and immunofluorescence with the G4-specific antibody BG4.

Hänsel-Hertsch et al., 2016 successfully used BG4-ChIP-seq to quantify the number of G4s in the genomes of immortalized and non-immortalized human keratinocytes (Hänsel-Hertsch et al., 2016). While purified BG4-FLAG was used for their experiments, I tried to perform ChIP experiments by expressing BG4-FLAG in yeast from the pGAL-BG4-FLAG vector as a novel approach (Figure 7). Although my BG4-FLAG construct expressed from pGAL-BG4-FLAG was able to bind G4s *in vitro* (Figure 10), I did not detect enrichment of BG4-FLAG at the G4 locus compared to a vector control in a *TOP1*-deletion background via ChIP-qPCR (Figure 11). This indicates that expressed BG4 was not able to bind to G4s *in vivo* and that further work needs to be done to enable the usage of expressed BG4 in the genome-wide quantification of G4s in yeast. One potential reason expressed BG4 was not enriched at *S $\mu$ G4-GTOP* could be that the BG4 construct does not contain a nuclear localization sequence. Since the import of most proteins into the nucleus is mediated via importin proteins that recognize nuclear localization sequences (Mattaj and Englmeier, 1998), it is possible that expressed BG4 is only localized to the cytoplasm of yeast cells. Continuous cytoplasmic localization would prevent expressed BG4-FLAG from binding G4 DNAs in the nucleus. Therefore, a nuclear

localization amino acid signal sequence, such as the widely used SV40 large T-antigen nuclear localization signal sequence (CTPPKKKRKV) (Kalderon et al., 1984), should be fused to the BG4-FLAG amino acid sequence in pGAL-BG4-FLAG. Alternatively, ChIP-seq experiments could be performed with purified BG4 antibody as was done in Hänsel-Hertsch et al., 2016. In fact, purified BG4-FLAG has recently become commercially available from Millipore Sigma (Catalog # MABE917), bypassing the requirement to purify BG4 from *E. coli* myself. Besides usage of BG4, ChIP-seq experiments could be performed with other yeast G4-binding proteins, like Sub1 or Nsr1, that have been shown to be enriched at the *G4* locus significantly more than at a non-G4 capable control locus *in vivo* via ChIP-qPCR (Lopez et al., 2017; Singh et al., 2020).

Although I didn't see any increased enrichment of expressed BG4 at the model *G4* locus relative to the vector control in *top1Δ* in ChIP-qPCR experiments, I still tried performing immunofluorescence in *top1Δ* cells transformed with pGAL-BG4-FLAG. I reasoned that visualizing the localization of our expressed BG4 construct within *top1* yeast cells may provide information that could be used to modify and improve ChIP-qPCR experiments. However, I did not observe any BG4-foci in the nuclei of the *top1Δ* cells transformed with pGAL-BG4-FLAG. One issue with the immunofluorescence experiments performed was that a high amount of background green fluorescence was present in cells stained with the GFP-conjugated  $\alpha$ -FLAG antibody (Figures 13 and 14). In the future, the green fluorescent background should be reduced by increased blocking time and more extensive washing as well as treatment with different antibody dilutions to find an optimal antibody concentration. Another issue was that  $\alpha$ -FLAG puncta seemed to be localized outside of some cells (Figures

13 and 14). This indicates that the Dylight 488  $\alpha$ -FLAG antibody may not be getting into all cells efficiently, and binding of the fluorescent antibody outside of cells may also explain the high levels of background fluorescence observed. The problem with permeability of the  $\alpha$ -FLAG antibody was also apparent in immunofluorescence experiments using the Nsr1-FLAG positive control depicted in Figures 15D and 15E, where only one cell seemed to exhibit colocalization of Nsr1 with the yeast nucleus. In the future, this protocol should be optimized and different permeabilization conditions should be tested such as treatment with different detergent concentrations and treatment times, and usage of various detergents altogether to try and increase the entrance of antibody into cells. Also, the spheroplasting protocol should be improved to find an optimum condition that makes the yeast cell wall more permeable. Since the expressed BG4 construct does not harbor a nuclear localization signal sequence, it is unclear if expressed BG4 is localized in the cytoplasm only or if it also enters the nucleus. Thus, either adding a nuclear localization signal sequence to the expressed BG4 construct or performing the immunofluorescence experiments with purified BG4 added to yeast cells after crosslinking and permeabilization could significantly improve the visualization of G4s in the yeast nucleus. Further, small molecules that fluoresce upon G4-binding, such as NaphthoTASQ, have been used to visualize G4s via microscopy in human cells (Laguerre et al., 2016). These small fluorescent G4-binding ligands may enter the yeast cell more readily than an antibody, making them an attractive alternative to immunofluorescence in the visualization of G4s. Additionally it could be that the small size of the yeast nucleus ( $\sim 2 \mu\text{m}$  in diameter) makes visualizing G4s in cells technically challenging (Taddei and Gasser, 2012). Therefore, chromosome spreading may be a better approach to visualize G4s that

form in the yeast genome. Chromosome spreading is a technique where whole chromosomes are extracted from cells and are spread on a glass microscope slide (Rockmill, 2009). Once spread on the slide, the yeast chromosomes could be stained with BG4-FLAG or a fluorescent G4 ligand. In fact, chromosome spreading coupled with immunofluorescence was used to visualize R-loops in yeast chromosomes (Lafuente-Barquero et al., 2017).

In conclusion, I was not able to quantify the number of G4s in the yeast genome *in vivo* by either ChIP-seq or immunofluorescence approaches performed with my BG4 construct expressed from pGAL-BG4-FLAG. Future approaches should include either a modified form of the expressed BG4 construct containing a nuclear localization signal sequence, purified BG4, other yeast G4-binding proteins, or small fluorescent G4 ligands. Also, the experimental approach called DMS footprinting mentioned in section 3.1 could be employed to measure G4-formation/stability *in vivo* in *WT* and *top1Δ* cells at specific G4-capable genomic loci. If *TOP1*-deletion results in an increase of G4-formation at *SμG4-GTOP* *in vivo*, then more DMS protection would be observed in *SμG4-GTOP* DNA isolated from *top1Δ* cells than in *SμG4-GTOP* DNA isolated from *WT* cells.

Determining if the absence of Top1 significantly affects the extent of G4 DNA formations *in vivo* is clinically relevant given the contribution of G4s to the development and progression of diseases like cancer (Bacolla et al., 2016; Bacolla et al., 2019). Additionally, Top1 is a major cancer chemotherapeutic target and is frequently mutated or present at reduced levels in cancer cells as a consequence of treatment (Beretta et al., 2013). Therefore, understanding how lack of functional Top1 impacts G4-formation on a genome-wide scale

could be relevant to the development of novel therapeutic strategies for cancer patients displaying resistance to Top1-targeting anticancer drugs.



## Chapter 4: Cleavage-Defective Topoisomerase I Mutants Sharply Increase G-quadruplex-Associated Genomic Instability

Note: Portions of this chapter were derived from a manuscript accepted for publication in *Microbial Cell*. BERROYER, A., BACOLLA, A., TAINER, J. A., and KIM N. 2022. Cleavage-defective Topoisomerase I mutants sharply increase G-quadruplex-associated genomic instability. *Microbial Cell*, Accepted January 19<sup>th</sup> 2022. *microbialcell.com* is the property of the Shared Science Publishers OG. Unless otherwise stated, all articles and corresponding materials accompanying them that are published by Shared Science Publishers OG on this Web Site are licensed by their respective authors for your use and distribution provided you cite the original source as specified in the corresponding Creative Commons Attribution License. Creative Commons Attribution License weblink: <https://creativecommons.org/licenses/by/4.0/>. Article available at <http://microbialcell.com/researcharticles/2022a-berroyer-microbial-cell/>.

#### 4.1 Introduction to Top1 Mutants

Top1 is an enzyme that relieves helical stress in DNA accumulated during transcription. Top1 binds duplex DNA, nicking one-strand of DNA with a catalytic tyrosine residue, and re-ligates the nick in DNA after controlled strand swiveling (Stewart et al., 1998). The complete deletion of the Top1-encoding gene in yeast or the siRNA-mediated knockdown of Top1 in mouse lymphoma B-cells results in highly elevated instability at genomic loci containing actively transcribed G4 DNA-forming sequences (Kim and Jinks-Robertson, 2011; Yadav et al., 2014; Husain et al., 2016). Further examination in yeast demonstrated that removal of transcription-associated negative excess helical tension by Top1 suppresses G4-induced recombination at these loci, indicating that Top1 plays an important role in protecting the genome by preventing G4-formation (Yadav et al., 2016). Top1 is a clinically relevant enzyme; it is the sole target of widely used anti-cancer chemotherapeutic camptothecin (CPT) and its derivatives including irinotecan and topotecan (Pommier et al., 2010). CPT targets Top1 by stabilizing the Top1-cleavage complex (Top1cc) consisting of Top1 covalently attached to the 3' end of a DNA nick and then by preventing the re-ligation step. To repair Top1ccs trapped by either DNA lesions or Top1 inhibitors including CPT, single-strand break repair TDP1 in complex with ligase III and XRCC1 can remove 3'-tyrosine residues remaining after degradation of Top1 (Rashid et al., 2021). Therefore, CPT cytotoxicity is mainly dependent on nuclear influx of CPT and DNA cleavage by Top1. While there are multiple ways in which cancer cells can become resistant to CPT, a prevalent mechanism of resistance is the mutation of Top1 such that the enzyme can no longer bind or cleave DNA. In fact, mutations that reduce Top1 DNA binding and/or cleavage are documented in cancer

patients and cells treated with CPT or CPT-derivatives (Beretta et al., 2013). Whether these Top1 mutants that arise in cancer cells in response to CPT-treatment affect G4-induced genomic instability has not been examined.

Interestingly, Top1 itself can bind tightly to G4 structures and can even promote formation of intermolecular G4s *in vitro* (Arimondo, 2000; Marchand et al., 2002). The yeast Top1 catalytic mutant Y727F, which can bind but not cleave duplex DNA, also binds to G4 DNA-forming oligos *in vitro* (Berroyer and Kim, 2020). Expression of Top1Y727F in yeast results in extremely elevated instability at a model G4-motif that is significantly higher than *top1* null mutation, possibly resulting from the high-affinity binding of Top1Y727F to co-transcriptionally formed G4s *in vivo* (Yadav et al., 2016). Here, Nayun and I investigated how CPT-resistance conferring Top1 mutants, like *yTop1Y727F* and those found in cancers, affect G4-induced genomic instability in *S. cerevisiae*.

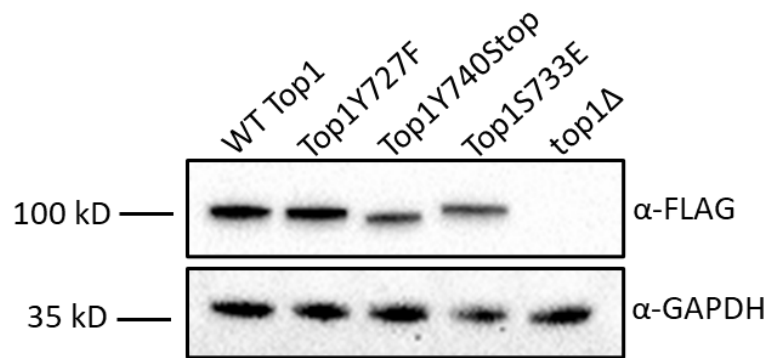
In addition to *yTop1Y727F*, I constructed two other yeast strains with *TOP1* mutant alleles (Table 4). *yTop1Y740Stop* is analogous to the human Top1W763Stop mutant identified in a lung cancer patient treated with the CPT-derivative irinotecan and is predicted to have reduced catalytic activity, while *yTop1S733E* is a duplex DNA binding mutant analogous to human Top1T729E mutant that confers CPT-resistance when expressed in yeast (Tsurutani et al., 2002; Losasso et al., 2008). Construction of both the *yTOP1Y740STOP* and *yTOP1S733E* strains allowed for me to distinguish between Top1 DNA cleavage-defects and Top1 DNA binding-defects as causes of G4-associated genomic instability. Expression of all C-terminally 3XFLAG-tagged Top1 mutants enabled their detection by western blotting (Figure 16A and

B). yTop1Y740Stop yielded a significantly lower steady-state protein level than WT yTop1, yTop1Y727F, or yTop1S733E.

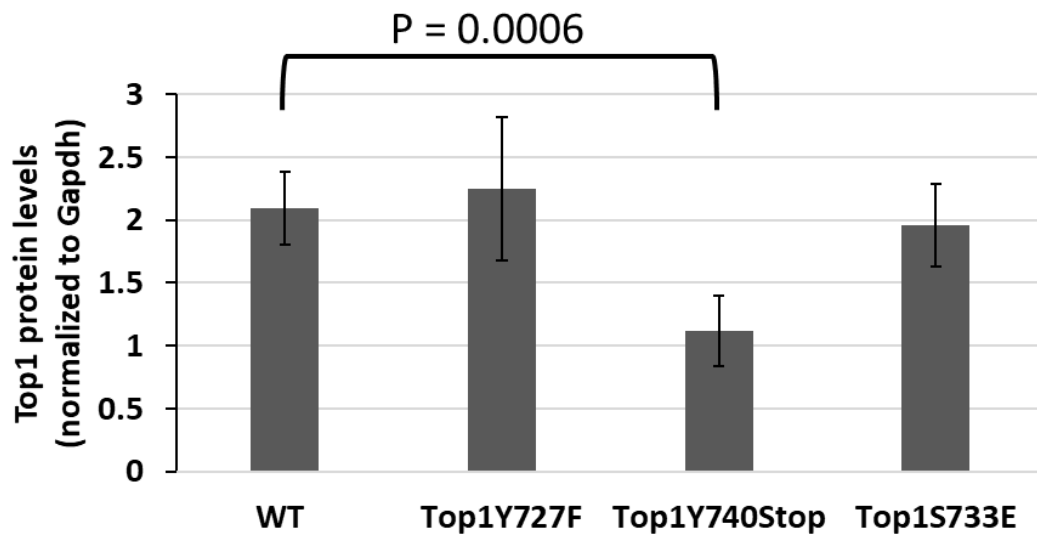
<b>Human Mutant</b>	<b>Yeast Mutant</b>	<b>DNA Binding</b>	<b>DNA Cleavage</b>	<b>Source</b>
Y723F	Y727F	Yes	No	Top1 crystal structure study (Redbino et al., <i>Science</i> , 1998)
W736STOP	Y740STOP	Yes	No	Non-small cell lung cancer patient treated with irinotecan (Tsurutani et al., <i>Lung Cancer</i> , 2002)
T729E	S733E	No	No	Study characterizing CPT-resistance conferring resistance (Losasso et al., <i>NAR</i> , 2008)

**Table 4. Top1 mutants studied**

**A.**



**B.**

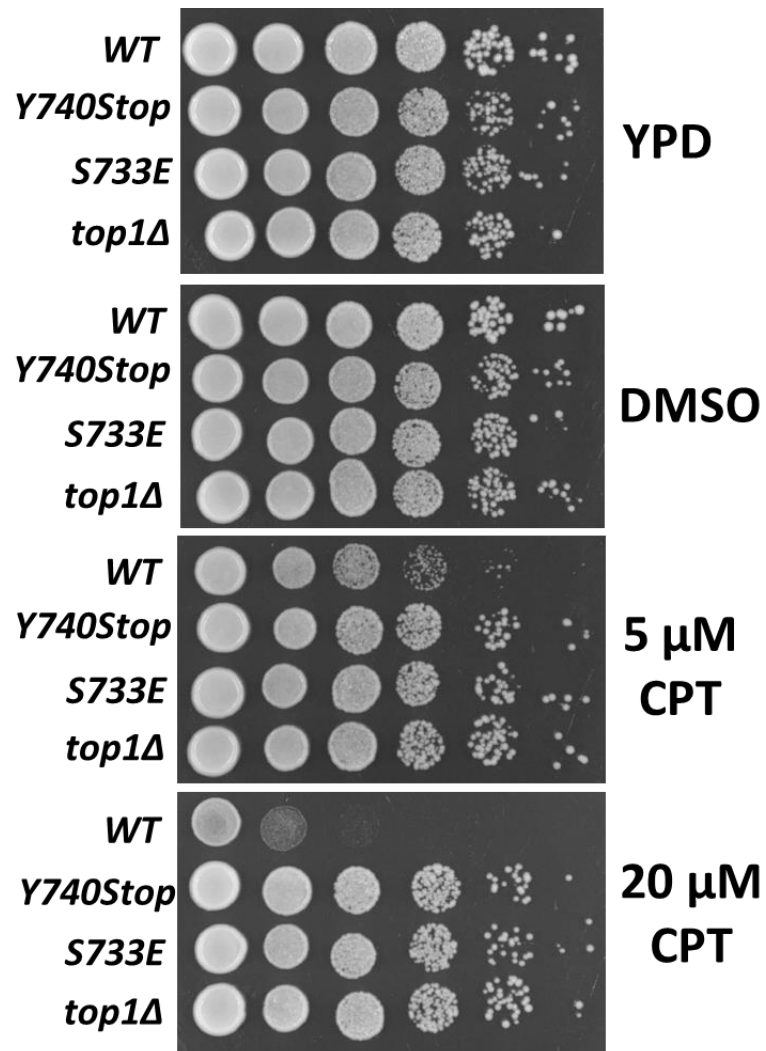


**Figure 16. Yeast homologs of human Top1 mutants found in cancer cells are expressed. A.** Western blot of cell extracts prepared from yeast strains expressing FLAG-tagged Top1 proteins. Blots were probed with  $\alpha$ -FLAG-HRP (Sigma) and  $\alpha$ -GAPDH-HRP (Invitrogen) antibodies. **B.** Quantification of Top1 mutant protein levels. Means and standard deviations of Top1 pixel intensities normalized to Gapdh pixel intensities from 5 western blots performed as in A. Significance of statistical differences was determined by Student T-tests (GraphPad Prism).

In order to confirm that the newly constructed yeast Top1 mutants Y740Stop and S733E are functionally defective, as expected from the study of analogous human mutations, I performed a spot assay using plates that contained CPT. Of note, I used *MUS81*-deletion yeast strains to perform the CPT spot assay so the repair of Top1ccs would be partly disabled. While the yeast strain expressing WT Top1 showed sensitivity to 5  $\mu$ M and 20  $\mu$ M CPT, *yTOP1Y740Stop* and *yTOP1S733E* strains were resistant to CPT like the *top1 $\Delta$*  yeast strain (Figure 17). To further confirm the yeast Top1 mutants are catalytically dead, Nayun and I used a yeast genetic assay where four AG repeats were inserted into the reversion window of the *lys2 $\Delta$ A746,NR* frameshift allele (Cho et al., 2013; Cho et al., 2015). As reported earlier, the reversion mutation at *lys2 $\Delta$ A746,NR,(AG)<sub>4</sub>* allele, which requires a net of two base pair deletion, is dependent upon the presence of functional Top1, particularly when RNase H2 is absent. The RNase H2 complex normally keeps genomic DNA free of ribonucleotides by initiating the excision repair of ribonucleotides incorporated during replication. In *rnh201 $\Delta$*  backgrounds, ribonucleotides remaining embedded in DNA are subsequently cleaved by Top1 and, in the case of repetitive sequences, frequently lead to slippage or frameshift mutations. For the *lys2 $\Delta$ A746,NR,(AG)<sub>4</sub>* allele, processing of the Top1-nicked DNA ends leads to two base pair deletions within the (AG)<sub>4</sub> repeats (Cho et al., 2013; Cho et al., 2015). As reported earlier, absence of functional Top1, as in *top1 $\Delta$*  backgrounds, results in >200-fold decrease in the mutation rate at the *lys2 $\Delta$ A746,NR,(AG)<sub>4</sub>* allele (Cho et al., 2013) (Table 5). Since mutagenic processing of ribonucleoside monophosphates (rNMPs) is dependent on Top1's ability to cleave DNA, we expressed Top1 mutants in the *lys2 $\Delta$ A746,NR,(AG)<sub>4</sub>rnh201 $\Delta$*  background to measure Top1-dependent rNMP cleavage as a proxy for Top1 catalytic activity. *LYS2* reversion

mutation rates of the *yTOP1Y727F*, *yTOP1Y740Stop*, and *yTOP1S733E lys2ΔA746,NR,(AG)<sub>4</sub> rnh201Δ* were all statistically indistinguishable from the *top1Δ lys2ΔA746,NR,(AG)<sub>4</sub> rnh201Δ* rate, confirming all three of the yeast Top1 mutants are catalytically inactive (Table 5).





**Figure 17.** Yeast Top1 mutants Y740Stop and S733E are resistant to CPT. Spot assays of *MUS81*-deletion yeast cells on control plates (YPD and DMSO) or plates containing the indicated concentrations of CPT. Cells were diluted 1:10 serially before spotting on plates. Images shown were taken 2 days after growth at 30 °C.

Strain	Lys+ Mutation Rate x 10 <sup>-8</sup> (CI)
<i>WT rnh201Δ</i>	1980 (1860-2250)
<i>top1Δ rnh201Δ</i>	9.21 (7.52-13.7)
<i>TOP1Y727F rnh201Δ</i>	13.1 (10.2-17.5)
<i>TOP1Y740STOP rnh201Δ</i>	10.9 (8.30-13.9)
<i>TOP1S733E rnh201Δ</i>	10.7 (9.23-15.8)

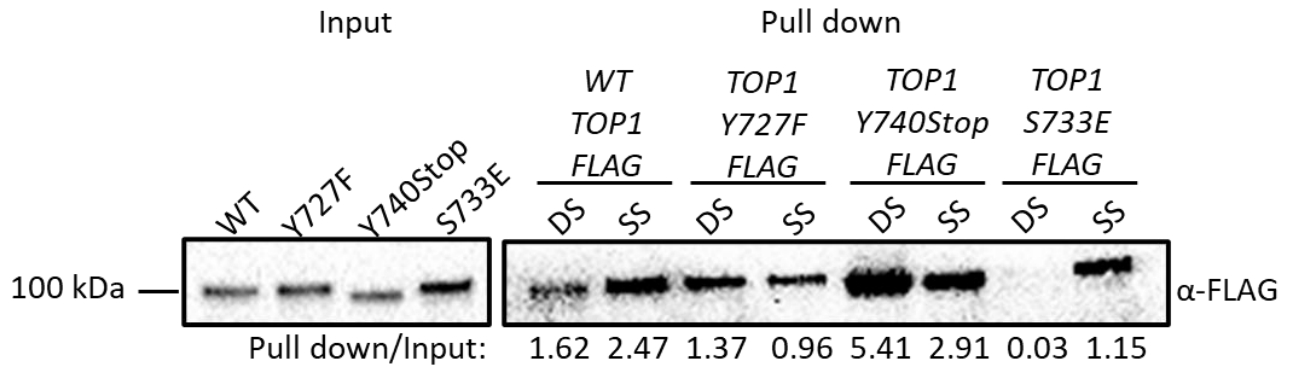
**Table 5. Catalytic activity of Top1 mutants determined by measuring the Lys+ mutation rate of yeast strains harboring a *pTET-lys2-AG4* reporter.** Rates are considered statistically significantly different if their 95% confidence intervals do not overlap (Spell & Jinks-Robertson, 2004). CI = 95% confidence interval. Yeast strains for this experiment were constructed by Alexandra Berroyer and mutation rates were measured by Dr. Nayun Kim.

I next measured the ability of the Top1 mutants to bind to a duplex DNA oligo using an *in vitro* biotinylated oligonucleotide (oligo) pulldown assay. The biotinylated SS oligo was annealed to the non-biotinylated RC-SS oligo to create a duplex substrate that could be used to pull down the Top1 mutants from yeast whole cell lysates if binding occurs (Figure 18A). As expected, the  $\gamma$ Top1Y727F and  $\gamma$ Top1Y740Stop cleavage-defective mutants bound duplex DNA while the  $\gamma$ Top1S733E DNA binding-defective mutant did not (Figure 18B). These DNA binding results coupled with the *lys2 $\Delta$ A746,NR,(AG)<sub>4</sub>* allele reversion mutation analysis results listed above validate the phenotypes of all three yeast Top1 mutants listed in Table 4.

**A.**

Oligo	Sequence
SS	5' GCACGCGTATCTTTTTGGCGCAGGTG
RC-SS	5' CACCTGCGCCAAAAAGATACGCGTGC

**B.**



**Figure 18. Top1 catalytic mutants bind duplex DNA substrates *in vivo*, while Top1 DNA binding mutants do not.** A. Sequences of oligonucleotides (oligos) used in *in vitro* binding oligo binding assay shown in B. B. Western blot of *in vitro* binding assay utilizing biotinylated oligos conjugated to streptavidin magnetic beads and yeast whole cell lysates prepared from cells expressing FLAG-tagged Top1 proteins. Blot was probed with an  $\alpha$ -FLAG-HRP (Sigma) antibody. SS = biotinylated single-stranded oligo control. DS = SS biotinylated oligo annealed to non-biotinylated RC oligo to create a duplex substrate. Quantification of binding is below blot and was calculated by dividing Input pixel intensities from Pull down pixel intensities.

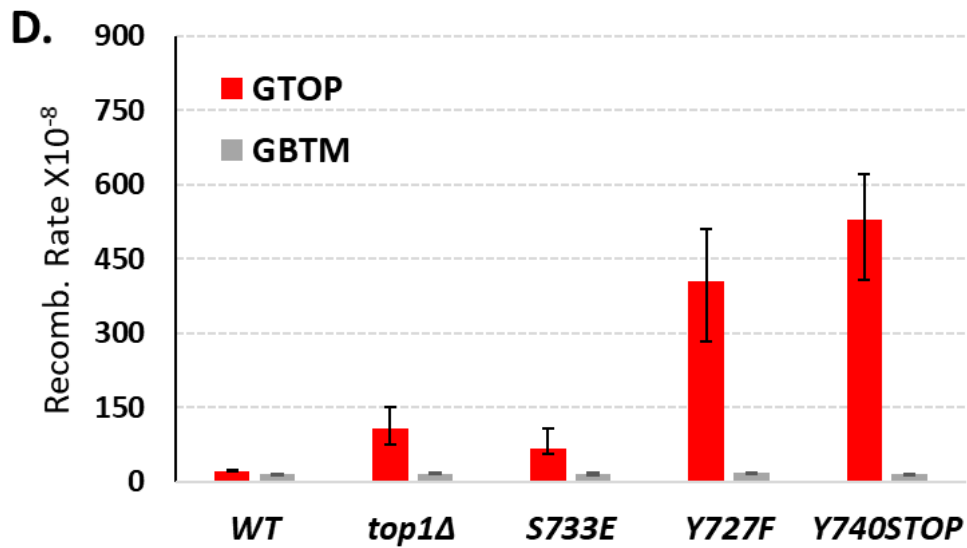
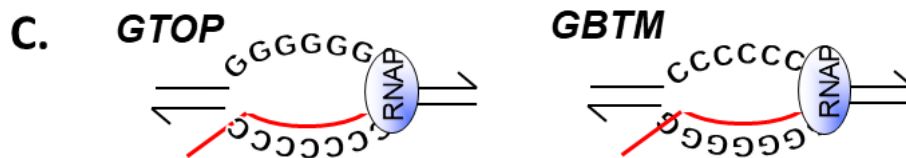
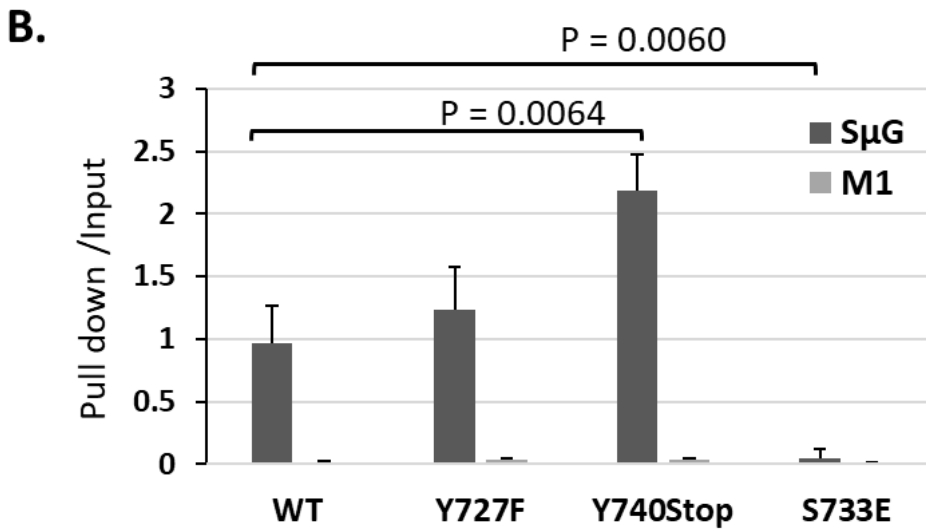
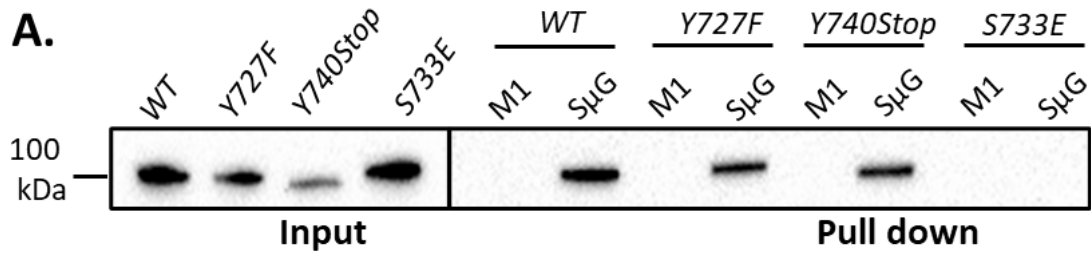
## 4.2 Results

### 4.2.1 Cleavage-Defective Top1 Mutants Bind G4s *in vitro*.

WT human Top1 binds to G4-forming oligos with high-affinity *in vitro* (Arimondo, 2000; Marchand et al., 2002). For the yeast protein, I previously showed that WT yTop1 and yTop1Y727F can bind G4 DNA *in vitro* (Berroyer and Kim, 2020). Enrichment of yTop1 and yTop1Y727F at a G4-capable telomeric region in yeast was also shown by ChIP experiments. To further examine G4-binding of WT yTop1 and yTop1Y727F and to determine if yTop1Y740Stop and yTop1S733E interact with G4s, I performed *in vitro* DNA binding assays using an oligo that is G4-capable (S $\mu$ G) (Table 6). As a non-G4 control, I used M1, where two guanine-runs present in S $\mu$ G were interrupted. WT yTop1, yTop1Y727F, and yTop1Y740Stop all formed stable complexes with the S $\mu$ G oligo (Figure 19A). However, yTop1S733E lacked appreciable binding to S $\mu$ G oligo, indicating that Top1 duplex DNA binding mutants do not form stable complexes with G4 structures like WT yTop1 or the Top1 cleavage-defective mutants yTop1Y727F and yTop1Y740Stop. Interestingly, while quantification revealed that yTop1Y740Stop bound the S $\mu$ G oligo significantly better than WT yTop1 or yTop1Y727F (Figure 19B), none of the Top1 proteins tested bound the G4-incapable M1 oligo, indicating that WT yTop1 and Top1 cleavage-defective mutants possess G4 structure-specificity in this assay (Figure 19A and B).

<b>Oligo</b>	<b>Sequence</b>	<b>Length</b>	<b>G-score*</b>
<b>S<math>\mu</math>G</b>	5'GAGCT <b>GGGG</b> TGAGCT <b>GGG</b> CTGA GCT <b>GGGG</b> TGAGCT <b>GGG</b> CTGAGCT	45 nt	70
<b>M1</b>	5'GAGCTGaGGT <b>G</b> AGCTGGGCTGA GCTGaGGT <b>G</b> AGCTGGGCTGAGCT	45 nt	N/A

**Table 6. Oligonucleotides used in binding assays.** S $\mu$ G oligo can adopt a G4 conformation while the M1 oligos cannot. Guanine-runs in S $\mu$ G oligo are in bold. G>A mutations in M1 oligo introduced to disrupt guanine-runs are indicated as lower case letters. \*G-score was calculated using QGRS Mapper (<https://bioinformatics.ramapo.edu/QGRS/index.php>; parameters: Max Length – 44, Min G-Group Size – 3, and Loop Size – 0 to 10).



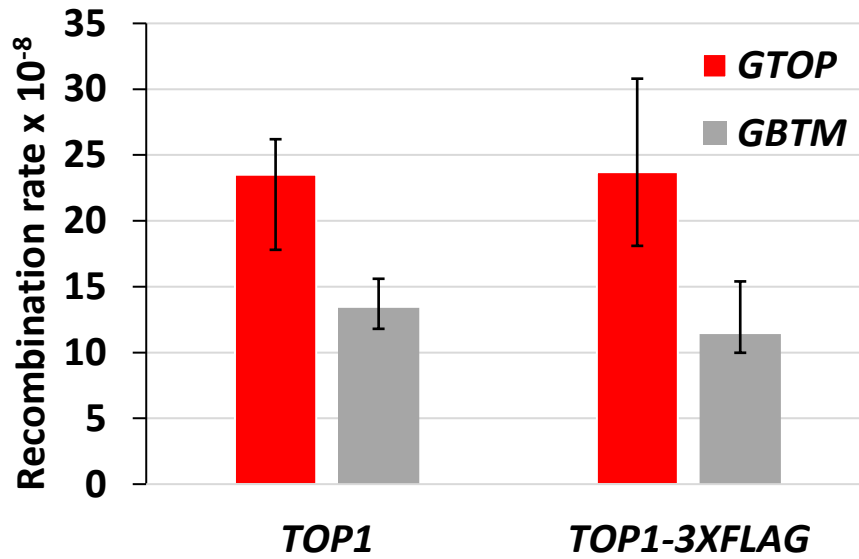
**Figure 19. Top1 mutants bind G4 oligos *in vitro* and increase G4-induced recombination in yeast.** *A.* Western blot from biotinylated oligo pulldown assay performed with streptavidin magnetic beads and yeast whole cell lysates expressing FLAG-tagged Top1 proteins. Blot was probed with an  $\alpha$ -FLAG-HRP (Sigma) antibody. Quantification of binding is listed below blot and was calculated by dividing input pixel intensities from pull down pixel intensities. *B.* Quantification (means and standard deviations of pull-down pixel intensities normalized to input pixel intensities) of western blots from 3 *in vitro* biotin oligo pulldown assays with S $\mu$ G and M1 oligos and Top1 proteins performed as in *A.* Vertical error bars only show positive standard deviations. P-value derived from students T-test (GraphPad Prism). *C.* Transcription orientations and the guanine-runs in the S $\mu$ -containing recombination reporter constructs. Guanine-runs are located on the non-transcribed or the transcribed strand in the *pTET-lys2-GTOP* or *pTET-lys2-GBTM* construct, respectively. The red line within the transcription bubble indicates the transcript. RNA polymerase complex is indicated as blue oval in front of the transcription bubble. *D.* Recombination rates of Top1 mutant-expressing yeast strains. Rates are considered statistically significantly different if their 95% confidence intervals (shown as error bars) do not overlap (Spell and Jinks-Robertson, 2004).



#### 4.2.2 Top1 Cleavage-Defective Mutants Increase G4-Induced Recombination More Than a Top1 Duplex DNA Binding Mutant.

I examined the effect of transcription on G4-induced recombination using a reporter construct containing a model G4-motif from the mouse immunoglobulin switch region Mu (*SμG4*). In this reporter, a segment of *SμG4* was integrated into the yeast genome within the *LYS2* gene under the control of tetracycline/doxycycline-repressible promoter (*pTET-lys2*) (Kim and Jinks-Robertson, 2011). The *SμG4*-motif was inserted into the *pTET-lys2* allele in two different transcriptional orientations, each disrupting the *LYS2* ORF. In *GTOP* orientation (*pTET-lys2-GTOP*), the guanine runs of *SμG4* are on the non-transcribed strand (NTS) which is transiently single-stranded during transcription, facilitating G4-formation (Figure 19C). In *GBTM* orientation (*pTET-lys2-GBTM*), the guanines of *SμG4* are on the transcribed strand (TS), which is base paired with the nascent mRNA during transcription and thus not likely to adopt G4-conformations. In this reporter assay, DNA strand breaks at *SμG4* *GTOP* or *GBTM* are repaired via recombination utilizing a truncated genomic copy of *LYS2*. The recombination rate at *SμG4* can thus be inferred from the emergence of Lys<sup>+</sup> recombinants. Any factor that affects G4-formation or -stability involves recombination starting at *GTOP*, but not *GBTM* (Kim and Jinks-Robertson, 2011). The Kim lab has previously shown that more recombination occurs at *pTET-lys2-GTOP* than at *-GBTM* under active transcription, and that this difference in recombination is increased significantly by the absence of Top1, supporting the hypothesis that Top1 functions to prevent co-transcriptional G4-formation by averting excessive torsional stress on DNA (Yadav et al., 2016).

Using the *pTET-lys2-GTOP* and *-GBTM* reporter strains, I first confirmed that the C-terminal 3X-FLAG tag does not affect the functioning of WT yTop1 in a recombination reporter assay (Figure 20). For yeast strains expressing yTop1Y727F or yTop1Y740Stop mutants, the recombination rates for the *pTET-lys2-GTOP* reporter construct were ~ 3.8- or ~ 4.9-fold higher than the rate for *top1Δ* strain, respectively (Figure 19D). The *pTET-lys2-GTOP* recombination rate in yTop1S733E-expressing strain was significantly lower than for the yTop1Y727F- or yTop1Y740Stop-expressing mutants and was statistically similar to the *top1Δ* strain. These results indicate that the exacerbated genomic instability at G4s in yeast cells expressing the Top1 cleavage-defective mutants yTop1Y727F or yTop1Y740Stop is due to a mechanism distinctly different from the *top1Δ* or the DNA binding-defective mutant yTop1S733E-expressing cells. Importantly, the effect of Top1 mutation on recombination is G4-specific since the recombination rates at the *pTET-lys2 GBTM* reporter construct did not significantly change by the expression of any of the Top1 mutants (Figure 19D). When transcription through *SμG4* was suppressed by adding doxycycline to culture media, the *pTET-lys2-GTOP* recombination rates of the *top1Δ*, *yTOP1Y727F*, *yTOP1Y740Stop*, and *yTOP1S733E* strains were slightly elevated relative to the *WT GTOP* rate but to much less extent than without doxycycline (Figure 21). Thus, the effect of *TOP1*-deletion and mutation on *GTOP* recombination is largely transcription-dependent.



**Figure 20. C-terminal 3X-FLAG tag does not impact the function of WT  $\gamma$ Top1.** Recombination rates of WT  $\gamma$ Top1 tagged and untagged recombination reporter strains. *GTOP* indicates recombination at G4-motif while *GBTM* indicates recombination at a loci not prone for G4-formation. Rates are considered statistically significantly different if their 95% confidence intervals (shown as error bars) do not overlap (Spell & Jinks-Robertson, 2004).

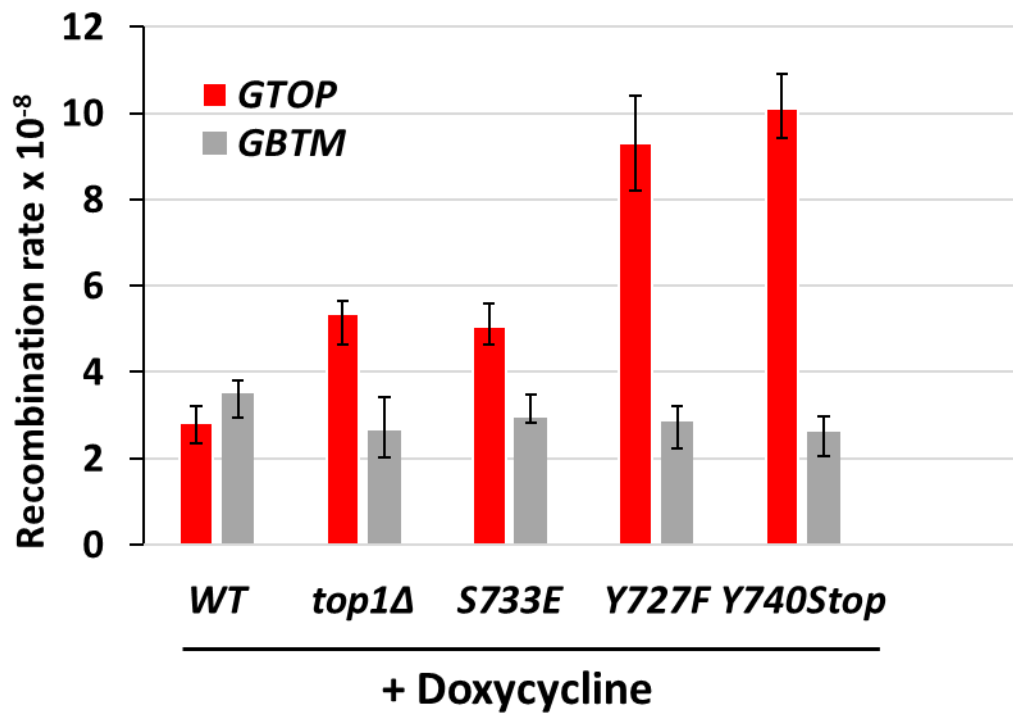
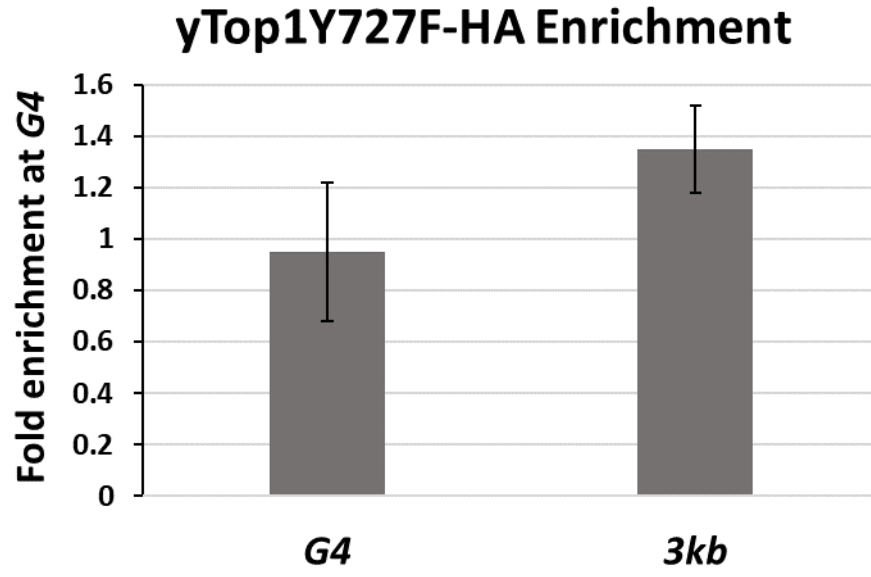


Figure 21. The effect of transcriptional repression on the *GTOp* and *GBTM* recombination rates of Top1 mutant strains. Doxycycline was added to the media of fluctuation cultures to repress transcription from the *pTET* promoter upstream of the *SμG4* *GTOp* and *GBTM* constructs. Rates are considered statistically significantly different if their 95% confidence intervals (shown as error bars) do not overlap (Spell & Jinks-Robertson, 2004).

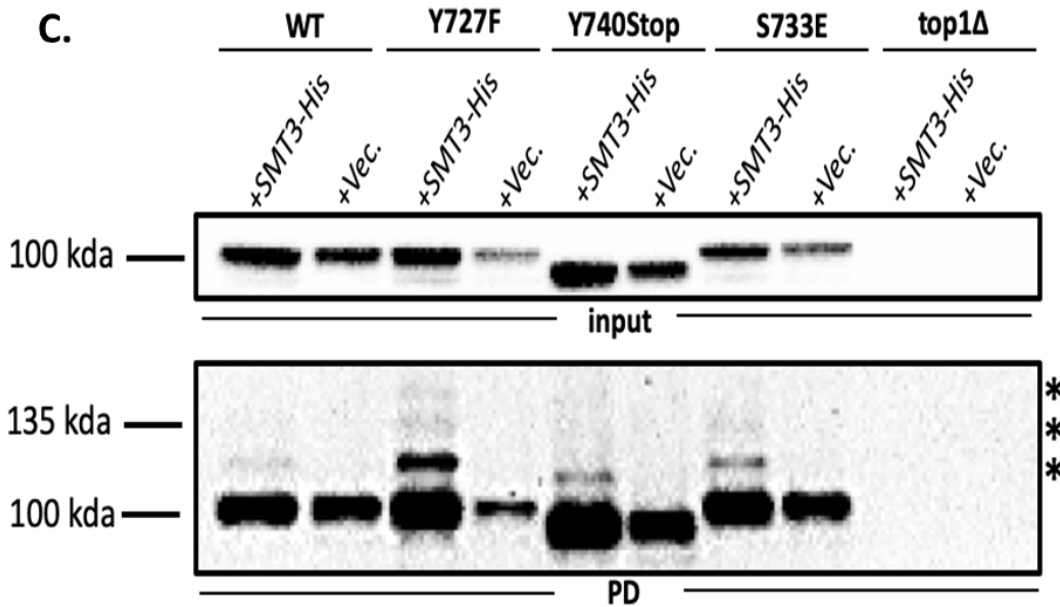
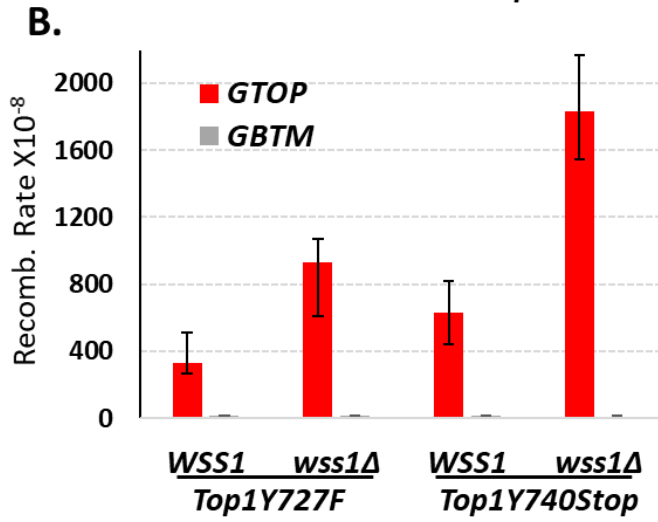
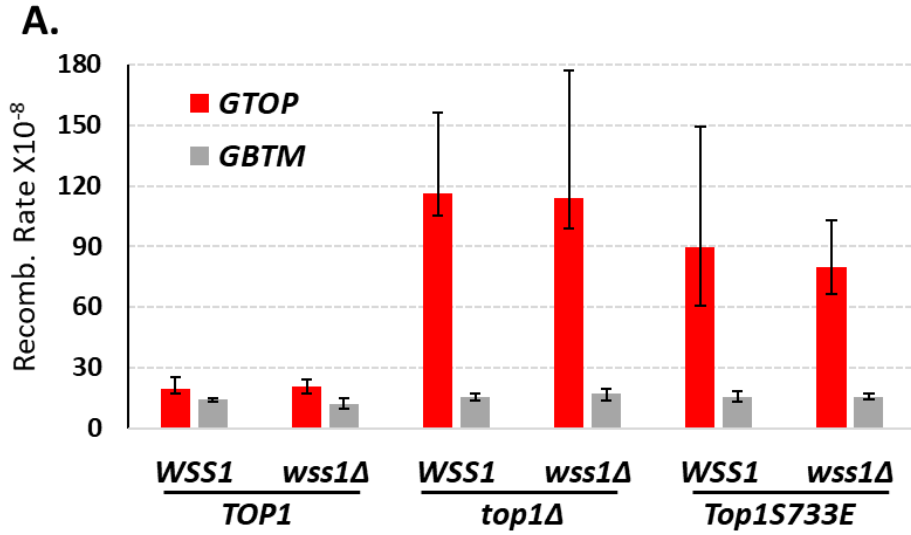
The *in vitro* binding datum in Figure 19A coupled with the G4-recombination datum in Figure 19D prompted me to investigate the *in vivo* binding of Top1 cleavage-defective mutants at  $S\mu G4$ - $GTOP$ . Thus, CHIP-qPCR was performed in *top1Δ* cells transformed with a plasmid expressing a  $\gamma TOP1Y727F$ -HA allele. To my surprise, no further enrichment of  $\gamma Top1Y727F$ -HA was observed at the “G4”  $S\mu G4$ - $GTOP$  locus relative the negative control, non-G4 capable “3kb” locus located ~3 kb from  $S\mu G4$  (Figure 22). However, both the “G4” and “3kb” loci are located in the same transcriptional unit in the *pTET-LYS2-SμG4* recombination reporter. Thus, both *G4* and *3kb* are constitutively transcribed in the absence of doxycycline. Of note, Top1 is known to be localized to genomic regions undergoing transcription, and even interacts with RNA pol2 (Phatnani et al., 2004; Baranello et al., 2016). Therefore, this observation that  $\gamma Top1Y727F$ -HA  $S\mu G4$ - $GTOP$ -binding does not occur more than  $\gamma Top1Y727F$ -HA binding to duplex DNA at a highly transcribed region could be due to  $\gamma Top1Y727F$ -HA still being localized to sites of active transcription despite being catalytically inactive. Further, while  $\gamma Top1Y727F$  G4-binding *in vivo* may not occur more often than binding of  $\gamma Top1Y727F$  at another non-G4 locus, the binding of  $\gamma Top1Y727F$  to G4s could be highly mutagenic, leading to the elevated recombination observed at co-transcriptionally formed G4s (Figure 19D). Additionally, the results depicted in Figure 22 could be due to experimental limitations of my CHIP-qPCR method.



**Figure 22.** Fold enrichment of yTop1Y727F-HA at the *S $\mu$ G4* *G4* and *3kb* loci *in vivo*. Averages and standard deviations of yTop1Y727F-HA fold enrichment in a *top1 $\Delta$*  *GTOP* background (n=6).

### 4.2.3 The DNA-Dependent Protease Wss1 Alleviates G4-Associated Genome Instability Exacerbated by Cleavage-Defective Top1 Mutants.

*WSS1* encodes a DNA-dependent protease that degrades proteins bound to DNA (Stingele et al., 2017; Stingele et al., 2014). Top1ccs trapped on DNA are Wss1's best characterized substrate. Deletion of *WSS1* did not affect the recombination in WT or *top1Δ* backgrounds (Figure 23A) but significantly elevated the rates of recombination at the *pTET-lys2-GTOP* construct only in *γTOP1Y727F-* or *γTOP1Y740Stop*-expressing strains (Figure 23B). These results support that the cleavage-defective Top1 mutants trapped on *SμG4 in vivo* were substrates for Wss1. *WSS1*-deletion did not affect the rate of recombination at the *pTET-lys2-GTOP* construct in *γTOP1S733E*-expressing strains, consistent with the *in vitro* oligo binding assays showing that *γTop1S733E* does not bind G4s formed on *SμG4 GTOP* (Figure 23A). Additionally, *WSS1*-deletion did not affect recombination rate at the *pTET-lys2-GBTM* construct in any of the strains (Figure 23A and 23B), indicating that Wss1 has a specificity for proteins bound to G4s in our fluctuation assays and that the Top1 catalytic mutants are not persistently bound to *SμG4* when G4-formation is not supported. In *wss1Δ* strains, Top1 WT and mutant proteins are SUMOylated with Top1Y727F being most extensively modified by SUMO ligation (Figure 23C).

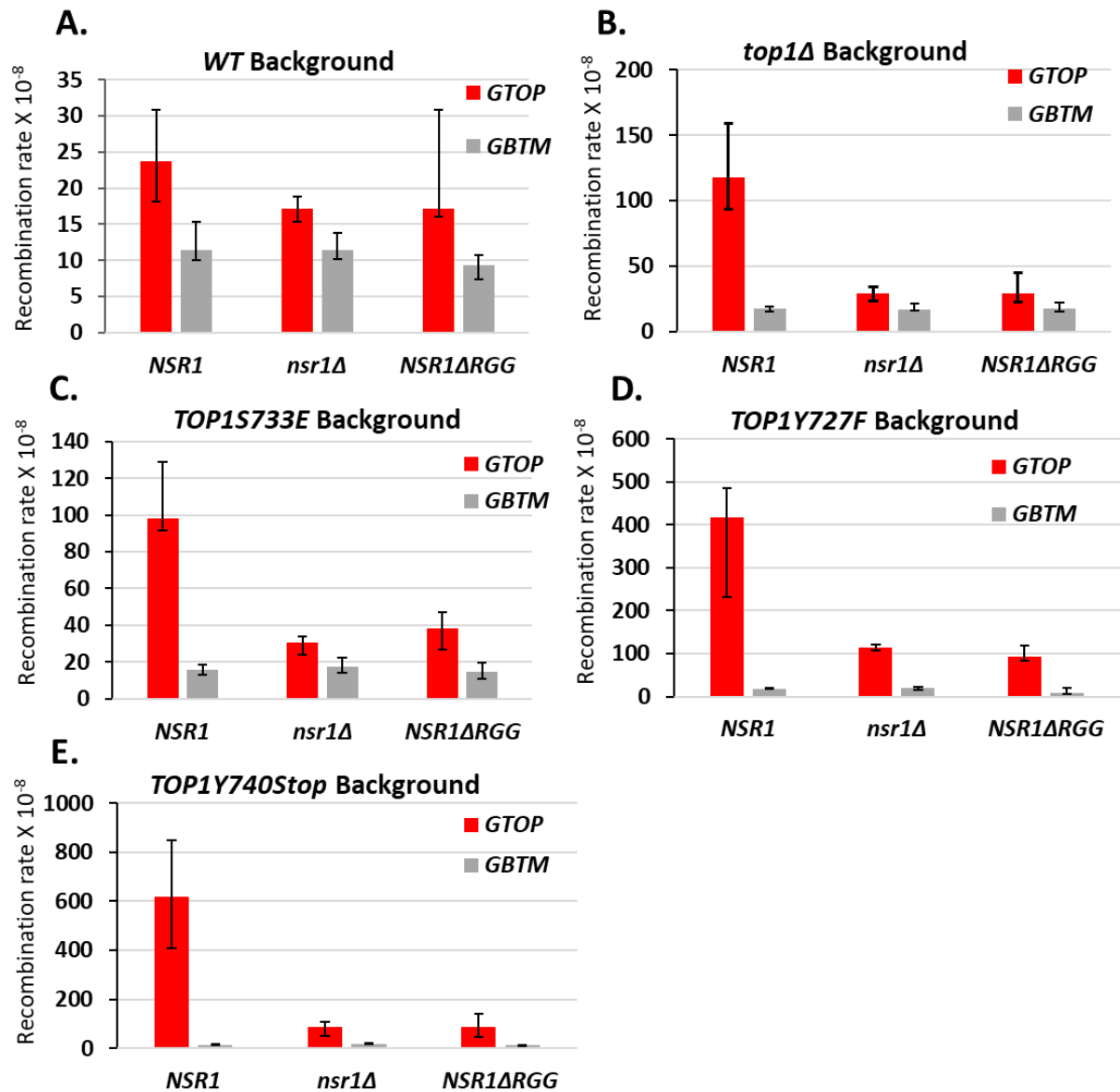




**Figure 23. Deletion of *WSS1* increases G4-induced recombination in Top1 catalytic mutant strains.** A.-B. Recombination rates of *WSS1* and *wss1Δ* yeast strains expressing indicated Top1 alleles. Rates are considered statistically significantly different if their 95% confidence intervals (shown as error bars) do not overlap (Spell and Jinks-Robertson, 2004). C. Top1-SUMO pull down experiment. All experiments were carried out in *wss1Δ* strains expressing C-terminally 3X-FLAG-tagged Top1 proteins and N-terminally 7XHis-tagged Smt3. SUMO-modified proteins were pulled down using Ni<sup>+</sup> beads and Top1 proteins detected by western blotting with α-FLAG-HRP antibody. Top panel – inputs; bottom panel – pull down samples (PD). \* denoted bands are mono- or poly-SUMOylated Top1 proteins.

#### 4.2.4 N-terminal domain and RGG repeats of Nsr1 are required for synergistic elevation of G4-induced genomic instability in cells with a Top1 cleavage-defective mutant.

The yeast protein Nsr1, a homolog of human nucleolin (NCL), is a G4-binding protein required for G4-induced recombination in *top1Δ* cells (Singh et al., 2020). Nsr1 binds to co-transcriptionally formed G4s in the absence of  $\gamma$ Top1 leading to replication stalling along the *SμG4*-containing *pTET-lys2-GTOP* locus. I deleted *NSR1* in the Top1 mutant backgrounds and measured recombination at the *SμG4*-containing recombination reporters. *NSR1*-deletion significantly reduced recombination rates at the *pTET-lys2-GTOP* reporter in all backgrounds (Figure 24A-E). Deletion of *NSR1* in *top1Δ* and *yTOP1S733E* strains reduced recombination at the *pTET-lys2 GTOP* reporter to *WT* background levels (Figure 24B-C). In cells expressing *yTOP1Y727F* or *yTOP1Y740Stop*, recombination rates at the *pTET-lys2 GTOP* reporter were reduced by *NSR1*-deletion but were still significantly higher than *WT* and similar to those measured in *top1Δ* (Figure 24A-B & Figure 24D-E). *NSR1*-deletion did not affect the recombination rate at the *pTET-lys2-GTOP* reporter in a *WT* background (Figure 24A).

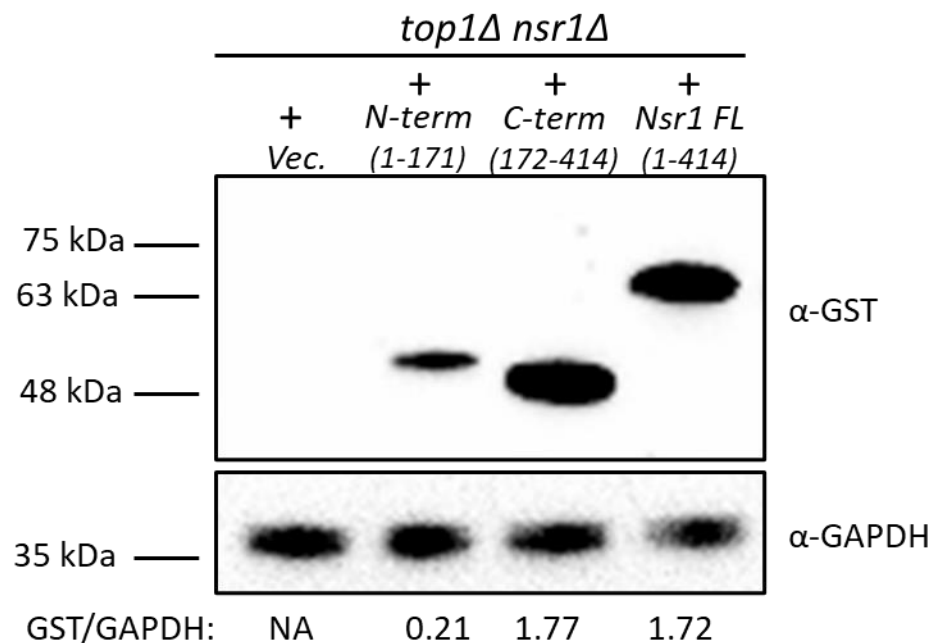


**Figure 24. Deletion of *NSR1* and deletion of the RGG domain of Nsr1 reduces G4-induced recombination in Top1 mutant strains.** A.-E. Recombination rates of *NSR1*, *nsr1Δ*, and *nsr1ΔRGG* strains expressing indicated Top1 alleles. Rates are considered statistically significantly different if their 95% confidence intervals (shown as error bars) do not overlap (Spell and Jinks-Robertson, 2004).

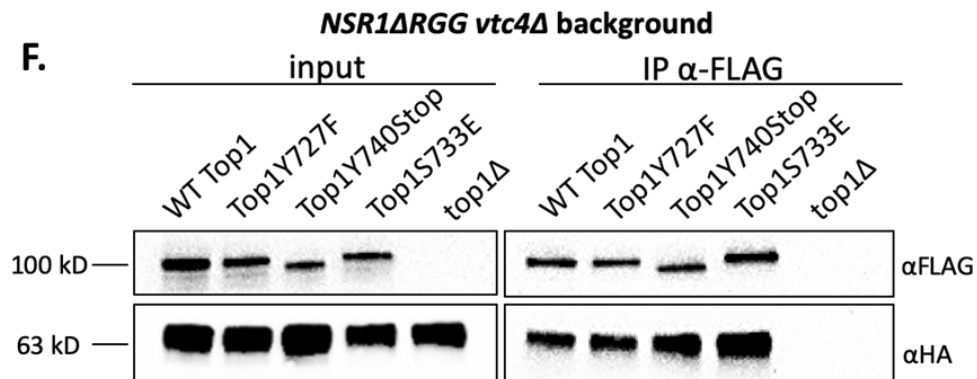
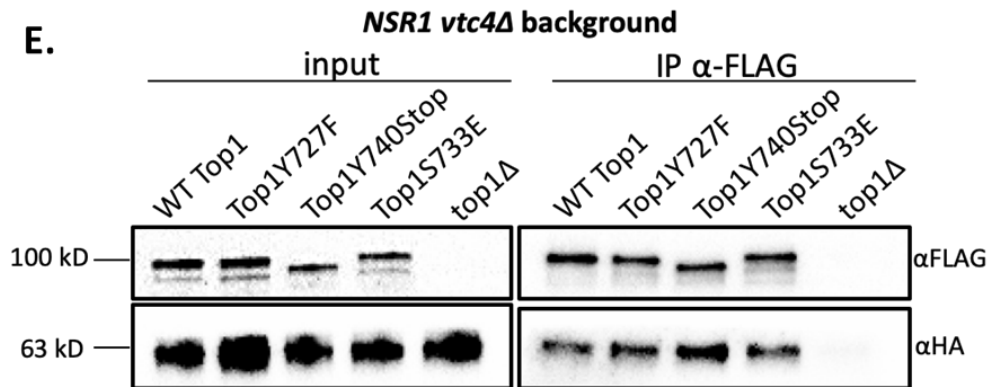
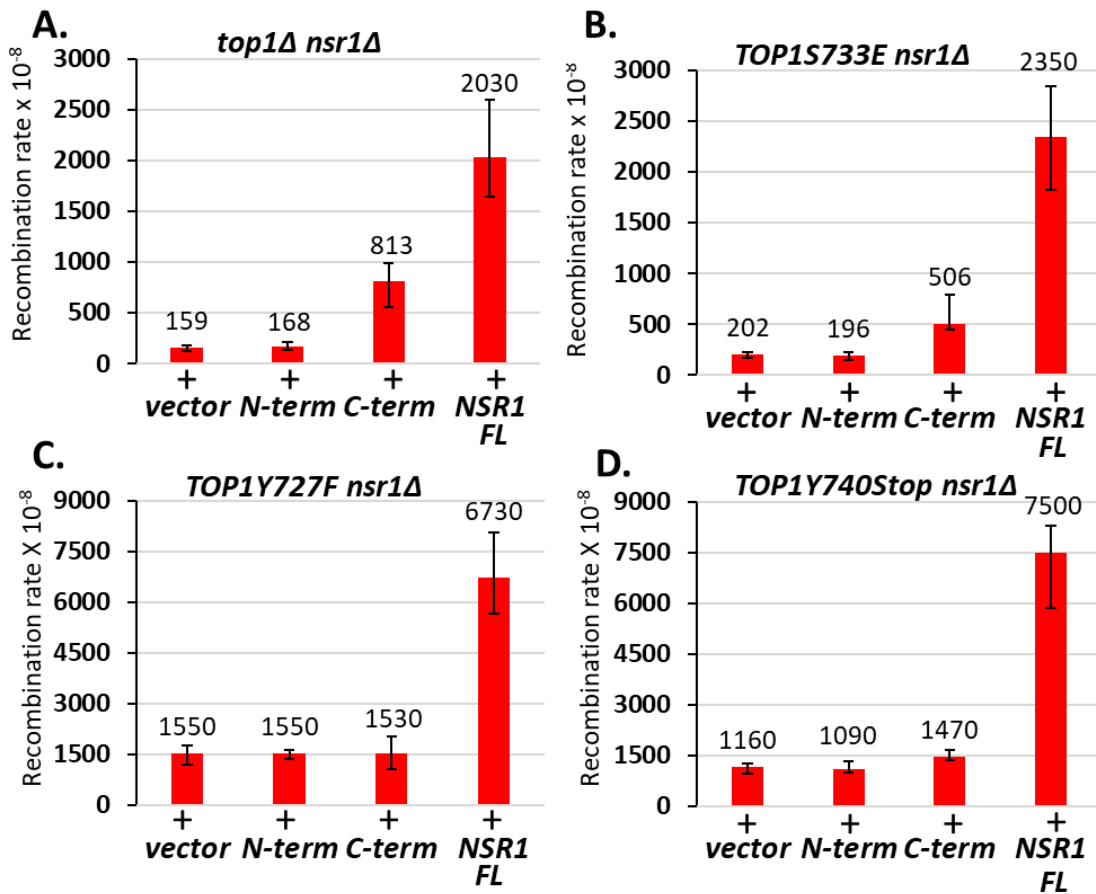
The C-terminally located RGG repeats of Nsr1 are required for high-affinity G4-binding (Singh et al., 2020). When I expressed a truncated form of Nsr1 lacking the RGG domain (Nsr1 $\Delta$ RGG = Nsr1 amino acid residues 1-350), recombination at the *pTET-lys2-GTOP* reporter was significantly reduced in all Top1 mutant strains (Figure 24C-E). As observed in the above experiments with *NSR1*-deletion, recombination at the *pTET-lys2 GTOP* reporter was reduced to that in *WT* level in the *yTOP1S733E NSR1 $\Delta$ RGG* strain but remained above *WT* levels in *yTOP1Y727F NSR1 $\Delta$ RGG* and *yTOP1Y740Stop NSR1 $\Delta$ RGG* strains (Figure 24A & Figure 24C-E). For the *pTET-lys2-GBTM* reporter construct, recombination rates were not affected by either *NSR1*-deletion or deletion of Nsr1's RGG domain in any background, except in the *yTOP1Y740Stop nsr1 $\Delta$*  strain where the recombination rate was significantly higher than the rate in the *yTOP1Y740Stop* background (Figure 24A-E).

I additionally performed functional complementation experiments in *nsr1 $\Delta$*  strains. The N-terminus of Nsr1 (N-term Nsr1; amino acids 1-171), the C-terminus of Nsr1 (C-term Nsr1; amino acids 171-414), or full-length Nsr1 (Nsr1 FL; amino acids 1-414) was expressed in *nsr1 $\Delta$*  cells expressing one of the Top1 mutants using a high copy  $2\mu$  plasmid and verified by western blot analyses (Figure 25). Expression of C-term Nsr1 in a *top1 $\Delta$  nsr1 $\Delta$*  background increased the recombination rate at the *pTET-lys2-GTOP* compared to either the vector control or the N-term Nsr1 (Figure 26A) (Singh et al., 2020). Similar effects were observed in the *yTOP1S733E nsr1 $\Delta$*  strain, where expression of Nsr1 C-term, but not Nsr1 N-term, increased recombination rates at the *pTET-lys2-GTOP* (Figure 26B). Notably, expression of full-length Nsr1 resulted in greater elevation in recombination rates at *pTET-lys2-GTOP* than C-term Nsr1 in both *top1 $\Delta$  nsr1 $\Delta$*  and *yTOP1S733E nsr1 $\Delta$*  backgrounds. In the *yTOP1Y727F*

*nsr1Δ* strain, the expression of neither C-term Nsr1 nor N-Term Nsr1 significantly changed recombination rates at the *pTET-lys2-GTOP* (Figure 26C). Only expression of full-length Nsr1 (Nsr1 FL) elevated the recombination rate at the *pTET-lys2-GTOP* above the control in the *yTOP1Y727F nsr1Δ* strain. Similar results were seen with the *yTOP1Y740Stop nsr1Δ* strain, where expression of Nsr1 FL resulted in a ~ 6.5-fold elevation in recombination rates at the *pTET-lys2-GTOP* relative to the vector control (Figure 26D). In the *yTOP1Y740Stop nsr1Δ GTOP* strain, expression of C-term Nsr1 led to a significant but relatively moderate increase in the recombination rate at the *pTET-lys2-GTOP* relative to the vector control - ~ 1.3-fold increase compared to the 5- or 2.5-fold increase seen with the expression of C-Term Nsr1 in *top1Δ nsr1Δ* or *yTOP1S733E nsr1Δ* backgrounds, respectively (Figure 26A, B, and D). In summary, the Top1-interacting N-terminus of Nsr1 is required to exacerbate G4-associated genome instability induced by *yTop1Y727F* or *yTop1Y740STOP*.



**Figure 25. Nsr1 constructs are expressed from pADH1-Nterm Nsr1, pADH1-Cterm Nsr1, and pADH1-Nsr1 in yeast.** Western blot of lysates prepared from *top1Δ nsr1Δ* cells transformed with control, pADH1-Nterm Nsr1, pADH1-Cterm Nsr1, and pADH1-Nsr1 plasmids. Nsr1-GST construct blots were probed with a primary α-GST antibody (Invitrogen) and a secondary α-mouse IgG-HRP antibody (R&D Systems-biotechne). Loading control blots were probed with an α-GAPDH antibody (Invitrogen). N-term = pADH1-Nterm Nsr1, C-term = pADH1-Cterm Nsr1, Nsr1 FL = pADH1-Nsr1, and Vec. = vector control. Quantifications of Nsr1 construct protein levels are listed below blot.



**Figure 26. Expression of full-length *NSR1* is required to greatly exacerbate G4-induced recombination in *TOP1Y727F nsr1Δ* and *TOP1Y740Stop nsr1Δ* strains.** *A-D*: Recombination rates of *top1Δ nsr1Δ* (*A*), *TOP1S733E nsr1Δ* (*B*), *TOP1Y727F nsr1Δ* (*C*), or *TOP1Y740STOP nsr1Δ* (*D*) yeast strains expressing the indicated Nsr1 constructs. Rates are listed above their respective bars and are considered statistically significantly different if their 95% confidence intervals (shown as error bars) do not overlap. N-term = pADH1-Nterm Nsr1, C-term = pADH1-Cterm Nsr1, NSR1 FL = pADH1-Nsr1, and vector = pRS426. *E* and *F*: Co-immunoprecipitation (co-IP) experiments conducted with *vtc4Δ* yeast strains expressing 3XFLAG-tagged Top1 proteins and either 6X-HA-tagged full-length Nsr1 (*E*) or NsrΔRGG (*F*). Pull down was carried out with αFLAG antibody-coated agarose beads. Blots were probed with either αFLAG or αHA antibodies. Quantification of binding was calculated by dividing FLAG IP pixel intensities from HA-IP pixel intensities and presented in graphs in Figure 31.

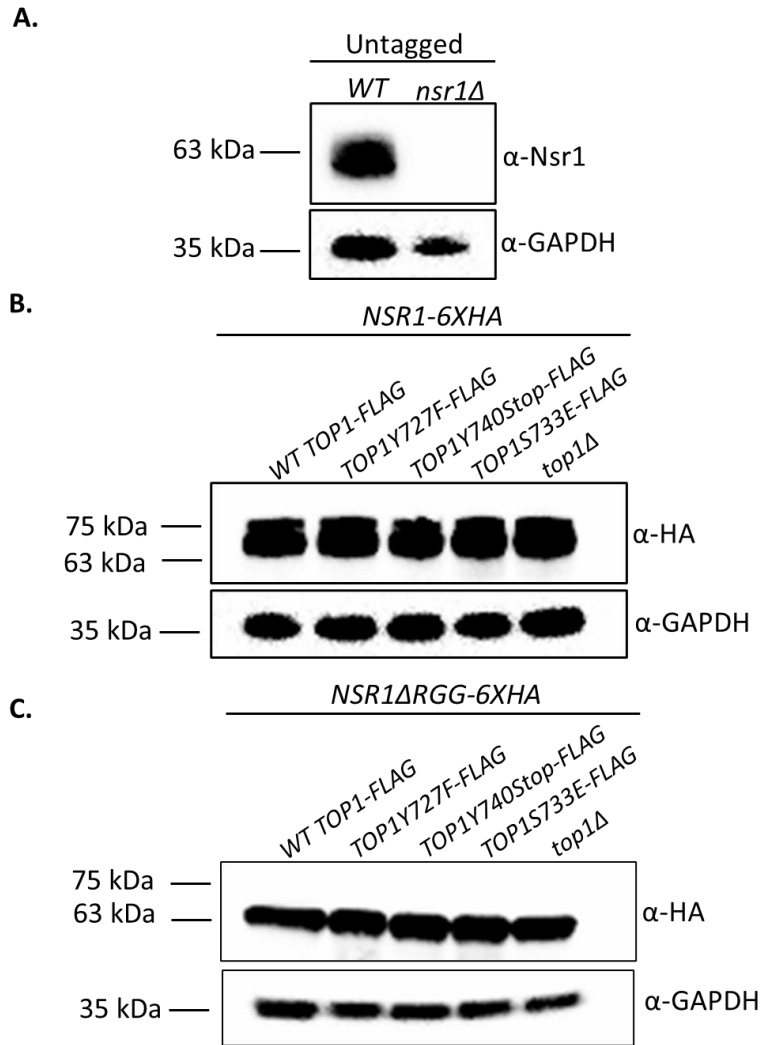


#### 4.2.5 Top1 Mutants Interact with Nsr1.

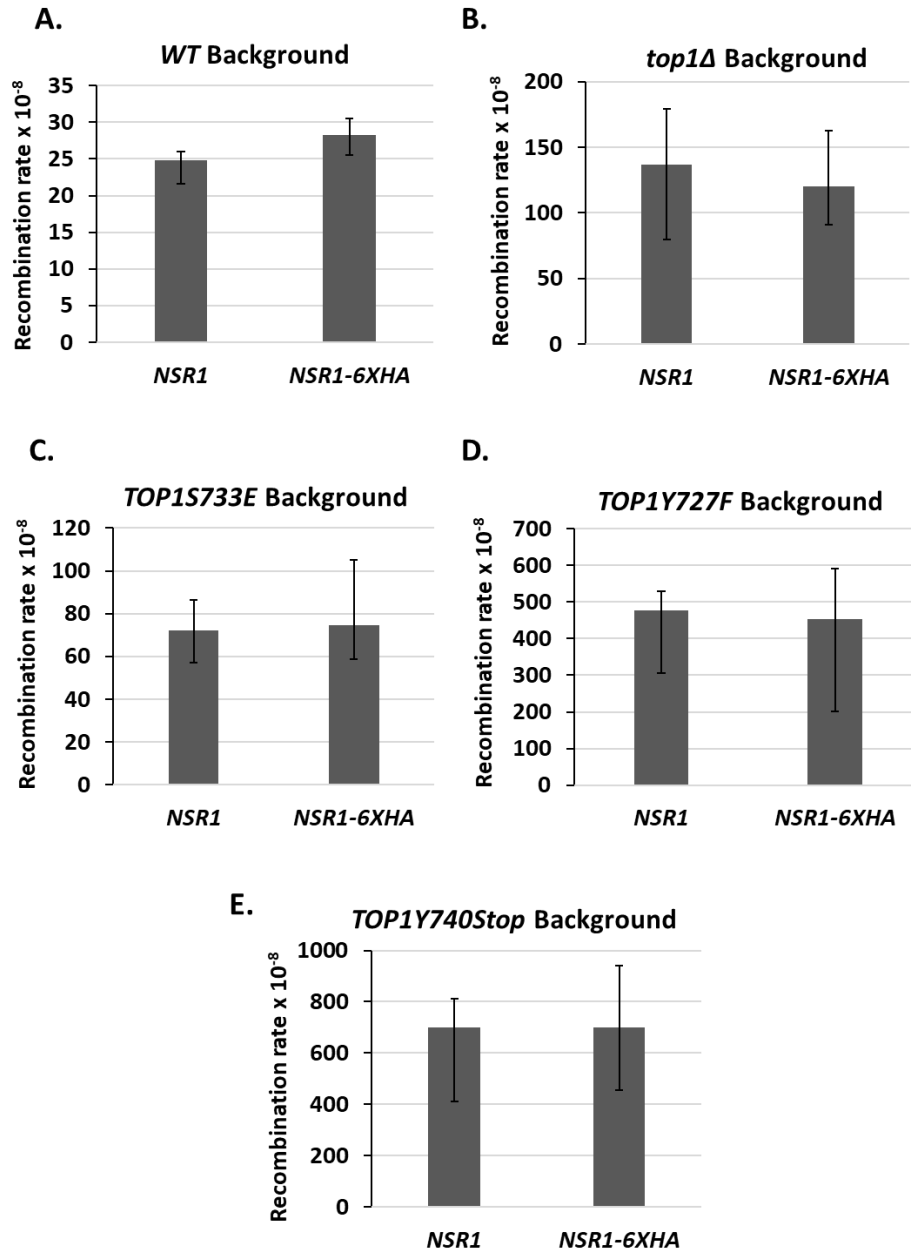
Because the complete deletion of *NSR1* or the expression of the Nsr1 $\Delta$ RGG-truncated protein significantly reduces G4-associated recombination in cells expressing Top1 cleavage-defective mutants, I tested whether Top1 cleavage-defective mutants can bind Nsr1 to cooperatively interact on G4 DNAs. Indeed, a G4 DNA/Nsr1/Top1-mutant complex could be a potent block to DNA replication and thus explain the exacerbated G4-induced instability observed in the presence of Top1 catalytic mutants and Nsr1. While interactions between WT Top1 and Nsr1 have been shown (Edwards et al., 2000; Azevedo et al., 2015), interactions between a Top1 mutant and Nsr1 have not been previously reported.

For co-immunoprecipitation (co-IP) experiments, C-terminally 3XFLAG-tagged Top1 mutants and C-terminally 6XHA-tagged Nsr1 were expressed in yeast cells. First, we showed that Nsr1-6XHA is stably expressed in all Top1 mutant backgrounds (Figure 27B) and that the C-terminal 6XHA-tag on Nsr1 does not affect G4-induced recombination at the *pTET-lys-GTOP* reporter (Figure 28). All co-IP experiments were conducted in a *VTC4*-deletion background because Top1/Nsr1 interactions are not easily detectable in the presence of Vtc4 protein, which is required for the synthesis of post-translational poly-phosphorylation (Figure 29) (Hothorn et al., 2009; Azevedo et al., 2015). *VTC4*-deletion did not affect the rate of recombination at *pTET-lys2-GTOP* reporter in any strain expressing either full-length Nsr1 or the Nsr1 $\Delta$ RGG truncation (Figure 30). In *vtc4* $\Delta$  backgrounds, Top1/Nsr1 interactions were detected in lysates prepared from all Top1 protein strains tested (Figure 26E) and quantification of multiple co-IP experiments revealed that all the Top1 proteins tested undergo a similar level of interaction with Nsr1 (Figure 31A). Interaction between Nsr1 $\Delta$ RGG

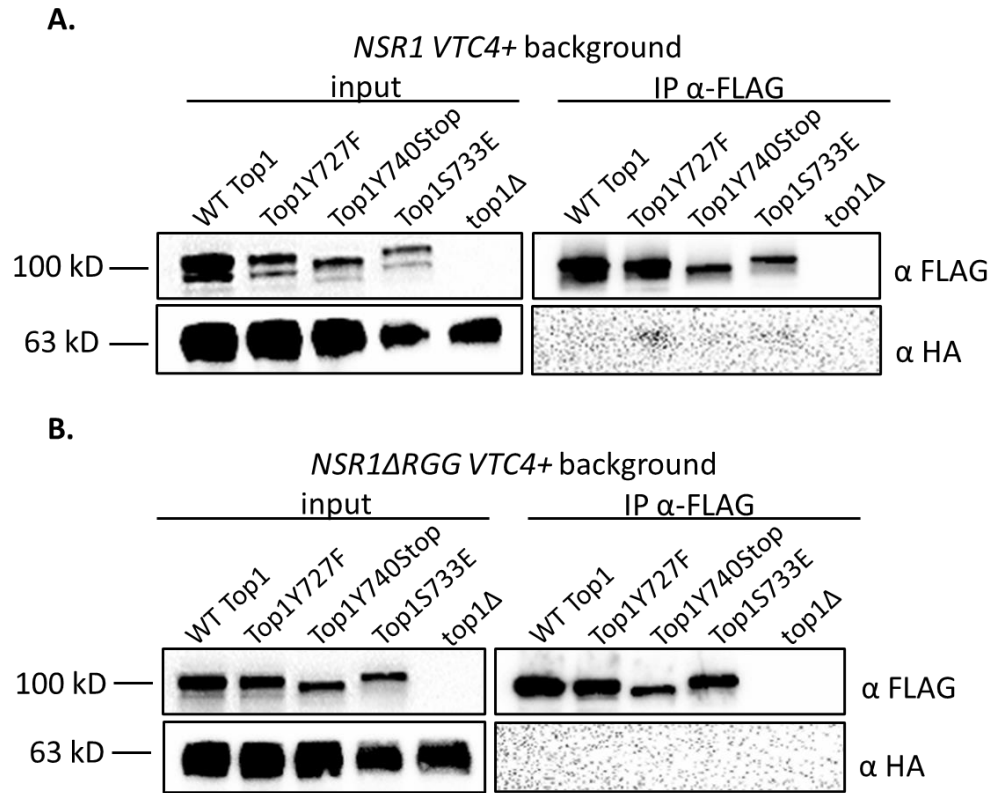
and Top1 mutants were tested through co-IP experiments in *vtc4Δ* strains (Figure 26F); all three Top1 mutants, i.e. Top1Y727F, Top1Y740STOP, and Top1S733E, interacted with Nsr1ΔRGG to a similar degree (Figure 31B). Our composite co-IP data demonstrate that neither Top1 mutation nor deletion of Nsr1's RGG domain impacts the Top1-Nsr1 interaction.



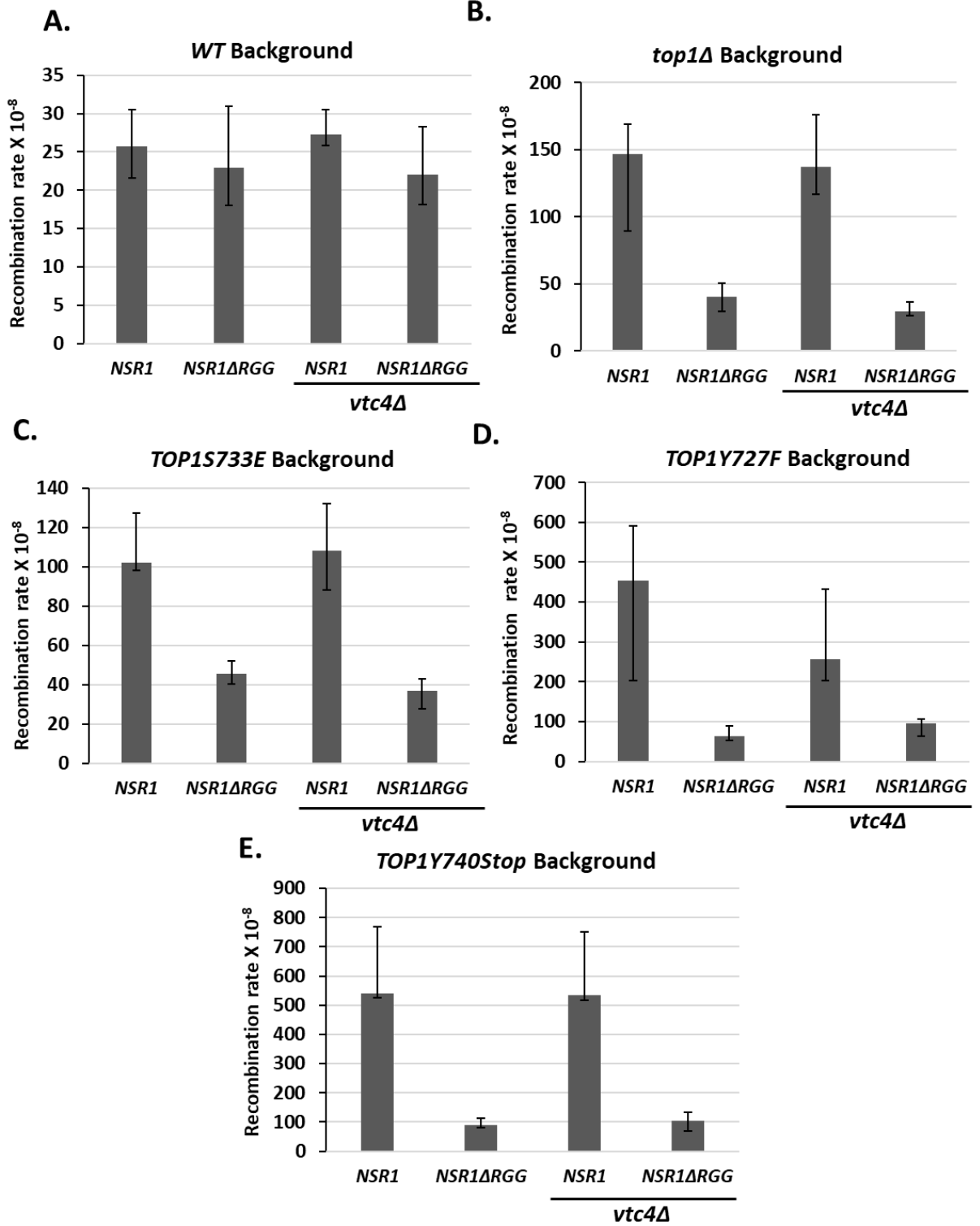
**Figure 27. *NSR1-6XHA* and *NSR1ΔRGG-6XHA* proteins are expressed in Top1 mutant backgrounds.** A. Western blot of lysates prepared from yeast cells expressing untagged, endogenous Nsr1 or *nsr1Δ* yeast cells. Untagged Nsr1 blot was probed with a primary α-Nsr1 antibody (Invitrogen) and a secondary α-mouse IgG-HRP antibody (R&D Systems-biotechne). B-C. Western blot of lysates prepared from Top1 mutant cells expressing either HA-tagged full length Nsr1 or HA-tagged Nsr1ΔRGG. HA-tagged Nsr1 and Nsr1ΔRGG blots were probed with an α-HA-HRP antibody (Sigma). Loading control blots in A.-C. were probed with an α-GAPDH antibody (Invitrogen).



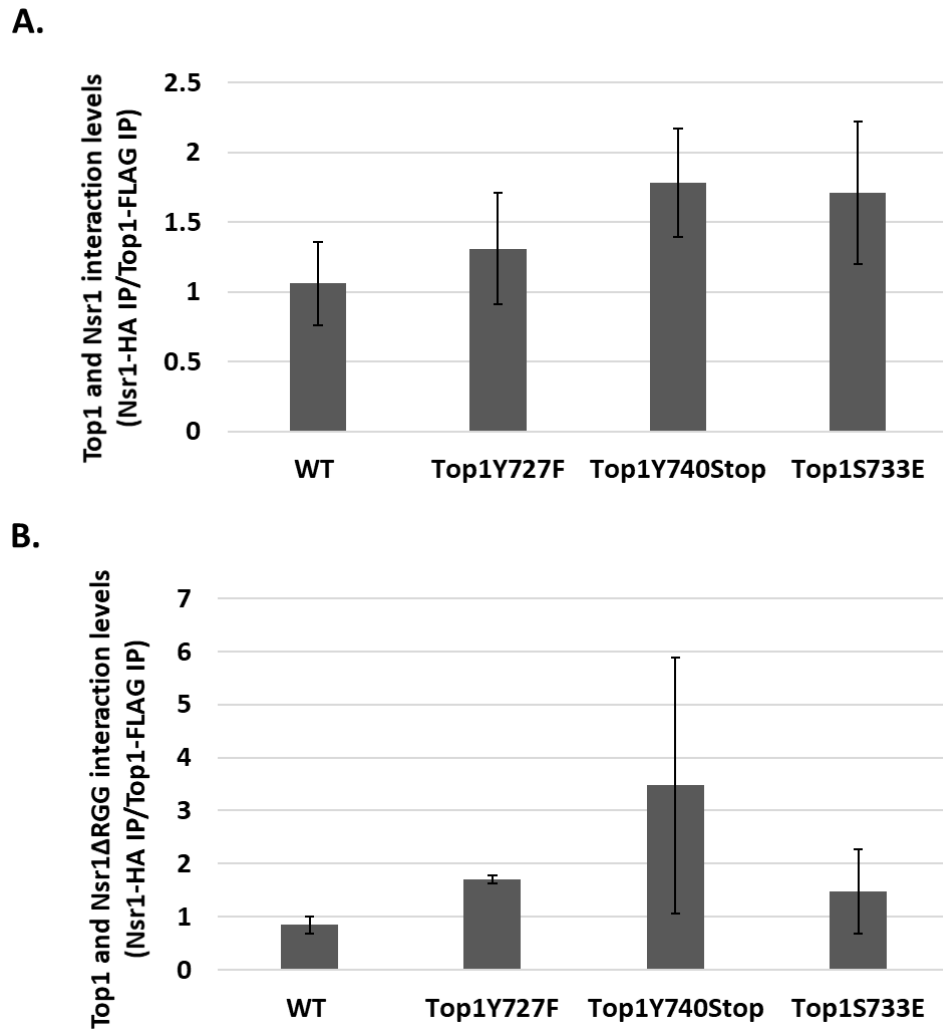
**Figure 28. *NSR1-6XHA* *GTOP* strains and untagged *NSR1* *GTOP* strains have comparable G4-induced recombination rates** A.-E. Recombination rates of Top1 mutant *NSR1-6XHA* *GTOP* strains shown with the rates of respective parental strains. Rates are considered statistically significantly different if their 95% confidence intervals (shown as error bars) do not overlap (Spell & Jinks-Robertson, 2004).



**Figure 29. Top1 mutant and Nsr1 co-immunoprecipitation (co-IP) experiments in the presence of *VTC4*.** A. Co-IP experiments conducted with *VTC4+* yeast strains expressing Top1 proteins that are C-terminally tagged with 3X-FLAG and full-length Nsr1 that is C-terminally tagged with 6X-HA.  $\alpha$ -FLAG coated agarose beads (Sigma) were incubated with yeast whole cell lysates. After pull down and washing, proteins were eluted from the beads by incubation with 3X-FLAG peptide (Sigma) and samples were run on an SDS PAGE followed by western blotting. Blots were probed with either  $\alpha$ -FLAG-HRP (Sigma) or  $\alpha$ -HA-HRP (Sigma) antibodies. B. Co-IP experiments conducted with *VTC4+* yeast strains expressing Top1 proteins that are C-terminally tagged with 3X-FLAG and Nsr1 $\Delta$ RGG that is C-terminally tagged with 6X-HA. Co-IP and western blotting was performed as in A.



**Figure 30. Deletion of *VTC4* does not affect G4-induced recombination in Top1 mutant strains.** A.-E. Recombination rates of Top1 mutant *VTC4+* and *vtc4Δ* *GTOP* yeast strains. Rates are considered statistically significantly different if their 95% confidence intervals (shown as error bars) do not overlap (Spell & Jinks-Robertson, 2004).



**Figure 31. Quantification of Top1 mutant interactions with Nsr1.** *A.* Quantification (means and standard deviations of Nsr1-6XHA IP pixel intensities normalized to Top1-3XFLAG pixel intensities) of western blots from 3 co-immunoprecipitation (co-IP) experiments performed as in Figure 26E with yeast cells expressing FLAG-tagged Top1 proteins and HA-tagged full length Nsr1. *B.* Quantification (means and standard deviations of Nsr1ΔRGG-6XHA IP pixel intensities normalized to Top1-3XFLAG pixel intensities) of western blots from 3 co-IP experiments performed as in Figure 26F with yeast cells expressing FLAG-tagged Top1 proteins and HA-tagged Nsr1ΔRGG.



### 4.3 Discussion

Co-transcriptional helical stress, promoting formation of DNA secondary structures including G4 DNA, is expected to accumulate when Top1's normal function of relieving DNA supercoils by binding and cleaving DNA is completely disrupted (Redinbo et al., 1998). Multiple studies including our lab's previous works support the notion that Top1 functions to prevent co-transcriptional G4-formation by removing negative helical stress (Yadav et al., 2016; Husain et al., 2016). It remained to determine how the expression of catalytic or DNA binding Top1 mutants found in CPT-resistant cancer cells would affect DNA aberrations at G4s. The C-terminal domain of Top1 partly forms a tight loop around duplex DNA and contains catalytically important residues, including the phosphotyrosyl bond-forming tyrosine (Y727 and Y723 in yeast and human, respectively) (Redinbo et al., 1998). The high conservation between yeast and human C-terminal domains of Top1 allowed us to measure the effect of C-terminal Top1 mutants found in CPT-resistant human cancer cells on G4-induced instability by expressing the analogous mutants in yeast.

In an earlier study, the Kim lab found that expression of the catalytically null mutant Top1Y727F results in severely elevated genome instability at G4 DNA-forming recombination reporter in yeast cells. Unexpectedly, the rate of G4-associated recombination in Top1Y727F-expressing cells was significantly higher than in *top1Δ* cells (Yadav et al., 2016). Here, I expressed another cleavage defective mutant, yTop1Y740Stop, and observed similarly acute elevation of G4-associated instability, whereas the DNA binding-defective mutant yTop1S733E had a more moderate effect on G4-induced instability (Figure 19D). Specifically, yTop1Y727F and yTop1Y740STOP expression resulted in *pTET-lys2-GTOP* recombination rates

that were around 5.9-fold and 7.8-fold higher than when was  $\gamma$ Top1S733E expressed. For the *pTET-lys2-GBTM* reporter where G4 DNA formation was unfavorable, Top1 mutants did not impact the rate of recombination, irrespective of the specific mutation, suggesting that the effect of Top1 mutation on recombination is G4-specific.

Multiple studies have documented interactions of Top1 with G4s (Arimondo, 2000; Marchand et al., 2002; Lotito et al., 2008; Berroyer and Kim, 2020). Therefore, Nayun and I postulated that the severe elevation in recombination rates at the *pTET-lys2-GTOP* reporter observed upon expression of Top1 cleavage-defective mutants Top1Y727F and Top1Y740STOP was the result of the binding and stabilization of co-transcriptionally formed G4s. It is possible that, while WT Top1 undergoes transient interactions with G4s *in vivo*, the inability to cleave DNA subsequent to binding could leave Top1Y727F and Top1Y740STOP trapped on G4s that form during transcription, significantly disrupting replication. Cleavage-defective mutants  $\gamma$ Top1Y727F and  $\gamma$ Top1Y740STOP, but not  $\gamma$ Top1S733E, bind a G4-forming oligo but not the control M1 oligo *in vitro* (Figure 19A and 19B). This result is in an agreement with previously published data showing purified calf thymus Top1 has a specificity for G4-capable oligos over non-G4 capable DNA substrates (Shuai et al., 2010). My *in vitro* binding datum is in line with the *GTOP* recombination datum, where Top1 catalytic mutants induce significantly greater recombination at *S $\mu$ G4* than the Top1 duplex DNA binding mutant (Figure 19D). Furthermore, when combined with the low steady-state protein level of  $\gamma$ Top1Y740Stop (Figure 16A-B), the high G4-induced recombination in yeast cells expressing this Top1 mutant (Figure 19D) levels supports a possible dominant negative phenotype of hTop1W736Stop.

*WSS1* encodes a SUMO-dependent metalloprotease (Stingele et al., 2014; Zhang et al., 2020). Wss1 and its mammalian homolog SPRTN degrade proteins forming covalent, irreversible complexes with DNA or DNA–protein cross-links (DPCs) and thus suppress genome instability incurred by DPC-induced replication impediments (Stingele et al., 2016). Besides DPCs occurring as trapped enzymatic intermediates, they can be generated by crosslinking of DNA to proteins by reactive agents such as formaldehyde and metal ions. Most Wss1 targets are SUMOylated and both Wss1 and SPRTN recognize and target Top1ccs that are post-translationally modified (Stingele et al., 2014; Stingele et al., 2017; Juhasz et al., 2012; Centore et al., 2012). Top1ccs have been shown to be SUMOylated (Mao et al., 2000), and human (Top1Y723F) and yeast (Top1Y727F) Top1 catalytic mutants were found to be more heavily SUMOylated than WT Top1 proteins (Heideker et al., 2011; Horie et al., 2002, Chen et al., 2007). Even through non-covalent interaction, certain proteins such as yeast Fob1 or *E. coli* Tus can form effective DPC-like stable, high-affinity complexes with DNA that block replication fork movement (Hizume and Araki, 2019). Recently, histones in such non-covalent DPCs were shown to be substrates for Wss1 (Maddi et al., 2020). Nayun and I postulate that Top1 mutants Y727F and Y740STOP, if forming replication barriers through high-affinity interaction with G4 DNA, could be SUMOylated and targeted by Wss1-dependent proteolysis.

Upon deletion of *WSS1*, the recombination rates at *pTET-lys2-GTOP* in *yTop1Y727F*- and *yTop1Y740STOP*-expressing strains were each elevated by ~ 3-fold (Figure 23B). Since the *pTET-lys2-GTOP* recombination rates in *WT* and *top1Δ* were not significantly changed by *WSS1*-deletion (Figure 23A), the effect of *WSS1*-deletion on G4-induced recombination is specific to the Top1 cleavage-defective mutants. Consistent with the observation where

yTop1S733E does not appear to be tightly bound to G4 DNA *in vitro* and where the recombination at *pTET-lys2-GTOP* did not further elevate upon expression of yTop1S733E in *top1Δ* background, WSS1-deletion did not affect the recombination rate in the yTop1S733E-expressing strain (Figure 23A). These results provide indirect evidence that Top1 mutants Y727F and Y740STOP form stable complexes with G4 DNA *in vivo* and further indicate that Wss1 can partly suppress the genome instability instigated by Top1 cleavage-defective mutants in complex with G4 DNA.

Nsr1 is the yeast homolog of human nucleolin, a clinically relevant protein that exhibits altered expression and localization in cancer cells (Koutsidoumpa and Papadimitriou, 2014; Berger et al., 2015; Huang et al., 2019). While the primary functions of nucleolin and Nsr1 are in pre-ribosomal RNA processing (Tajrishi et al., 2011), both nucleolin and Nsr1 are G4-binding proteins (Hanakahi et al., 1999; Fry, 2007; Singh et al., 2020). The biological consequence of the nucleolin-G4 DNA interaction was demonstrated by transcriptional change in several oncogenes including *MYC* upon binding of nucleolin to G4s present in the promoter (Gonzalez et al., 2009; Gonzalez and Hurley, 2010). Our group recently published the first results showing that, in *top1Δ* cells, Nsr1-G4-binding is responsible for the elevated recombination at the *SμG4*-containing *pTET-lys2-GTOP* reporter as well as for the significant lag in DNA replication timing observed at the *SμG4*-containing genomic locus (Singh et al., 2020). In the current study, I show that *NSR1*-deletion significantly reduces the G4-associated genomic instability observed in all Top1 mutant yeast strains tested (Figure 24C-E). In yTop1S733E-expressing cells, the decrease in recombination rates due to deletion of *NSR1* resembles that seen in *top1Δ* background, indicating that the DNA-binding defect in this Top1

mutant essentially mimics the lack of functional Top1 with no additional detrimental effect. However, in *yTop1Y727F-* and *yTop1Y740Stop*-expressing cells, rates of recombination at the *pTET-lys2-GTOP* reporter were reduced upon deletion of *NSR1* but were still significantly higher than in *WT* background. This indicates that Top1 cleavage-defective mutants can instigate G4-associated instability even in the absence of Nsr1.

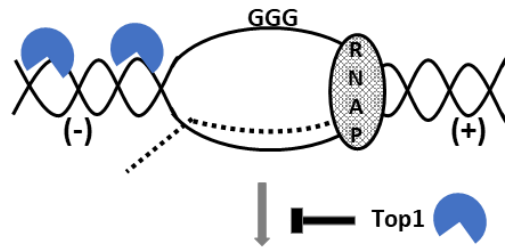
The C-terminal RGG domain of human nucleolin is important for the protein's high-affinity interaction with G4 structures (Hanakahi et al., 1999; Ghosh and Singh, 2018). Moreover, phenylalanine residues in the RGG domain of nucleolin participate in G4-binding and G4-folding, as shown through electrophoretic mobility-shift assays and circular dichroism spectroscopy experiments (Masuzawa and Oyoshi, 2020). Our lab's prior work uncovered that the C-terminally located RGG-domain of Nsr1 is required for G4-binding and the induction of co-transcriptional G4-induced instability in the absence of Top1 (Singh et al., 2020). In the current study, I show that deletion of the RGG domain of Nsr1 also significantly reduces recombination rates at the *pTET-lys2-GTOP* reporter construct in Top1 mutant-expressing strains (Figure 24C-E). As seen with *NSR1*-deletion, the level of decrease in recombination rates due to *Nsr1Δ RGG* in *yTop1S733E*-expressing cells resembles that seen in *top1Δ* background. However, in *yTop1Y727F-* and *yTop1Y740Stop*-expressing cells, the rates of recombination at the *pTET-lys2-GTOP* reporter were significantly higher than in the *WT* background even after deleting the RGG domain. Altogether, the *NSR1*-deletion and *Nsr1ΔRGG* fluctuation data allude to two possible models explaining how expression of both Top1 catalytic mutants and Nsr1 leads to severely heightened recombination at *pTET-lys2-GTOP*. One model is that the high level of G4-associated recombination in yeast cells

expressing Top1 cleavage-defective mutants is the result of an additive effect where G4s are more frequently bound by either Nsr1 or Top1 mutants. Alternatively, Top1 mutants and Nsr1 interacting and cooperatively binding to G4s could result in a synergistic effect.

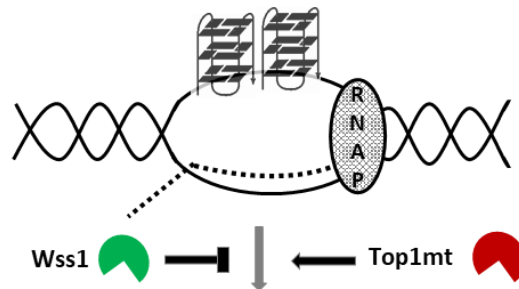
While the exact biological relevance of the Top1/Nsr1 interaction is not completely understood, it is thought to be related to the localization of Top1 to the nucleolus (Edwards et al., 2000). Since rDNA located within the nucleolus is highly transcribed and Top1 relieves transcriptional helical stress, it is conceivable that Nsr1 recruits Top1 to the rDNA locus to maintain optimal helical torsion. I confirmed through co-IP experiments that WT  $\gamma$ Top1 and Nsr1 interact (Figure 26E). In addition, I found that neither Top1 mutation nor deletion of Nsr1's RGG domain affects Top1/Nsr1 interaction (Figure 26F). These co-IP data suggest the effect of Top1 catalytic mutants and Nsr1 on G4-induced recombination could be synergistic since Top1 mutants and Nsr1 interact and our lab previously showed that Nsr1 is significantly enriched at *S $\mu$ G4* *GTOP* relative to non-G4 loci in ChIP experiments (Singh et al., 2020). I also showed that only the expression of full-length Nsr1, but not the Nsr1 C-term or N-term (required for interaction with Top1(Edwards et al., 2000)), increased the recombination rate at the *pTET-lys2-GTOP* reporter relative to the vector control in *TOP1Y727F nsr1 $\Delta$*  and *TOP1Y740STOP nsr1 $\Delta$*  cells (Figure 26C and 26D). This indicates that the synergistic effect on G4-induced recombination requires both interaction between Nsr1 and Top1 and the interaction between Nsr1 and G4 DNA. Altogether, my data suggest a model in which Top1 catalytic mutants and Nsr1 bind to and interact on co-transcriptionally formed G4s to form a highly mutagenic complex that prevents G4-resolution and potentially impedes replication fork movement through G4-motifs (Figure 32). Future experiments in the Kim lab will focus

on uncovering if replication through G4-motifs is disrupted in Top1 catalytic mutant cells expressing full-length Nsr1 to better elucidate this mutagenic mechanism.

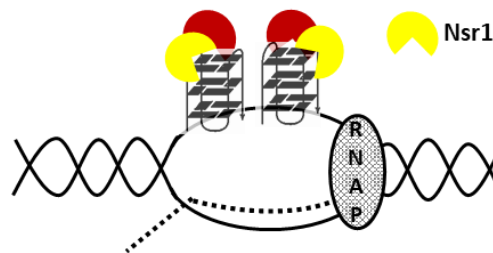
**G-run sequences at a transcribed region**



**Negative tension promotes G4 DNA formation**



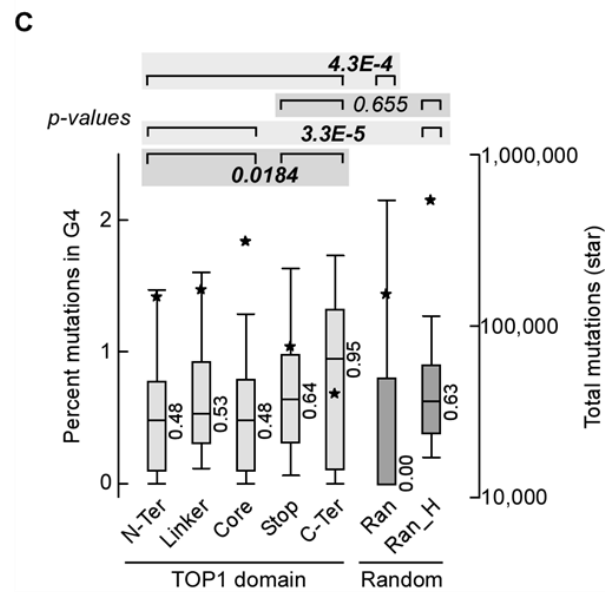
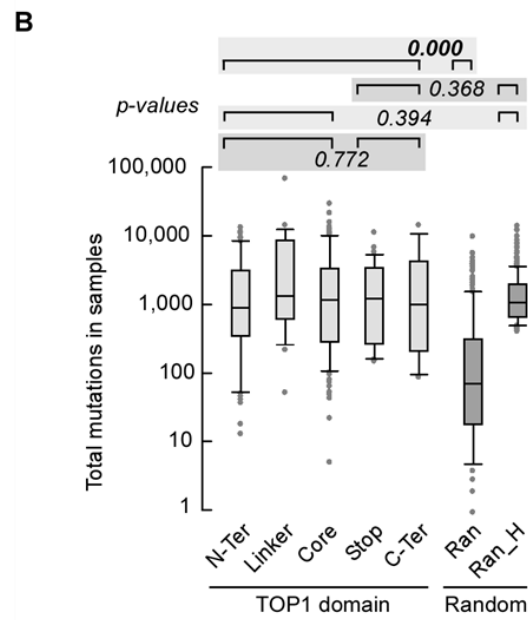
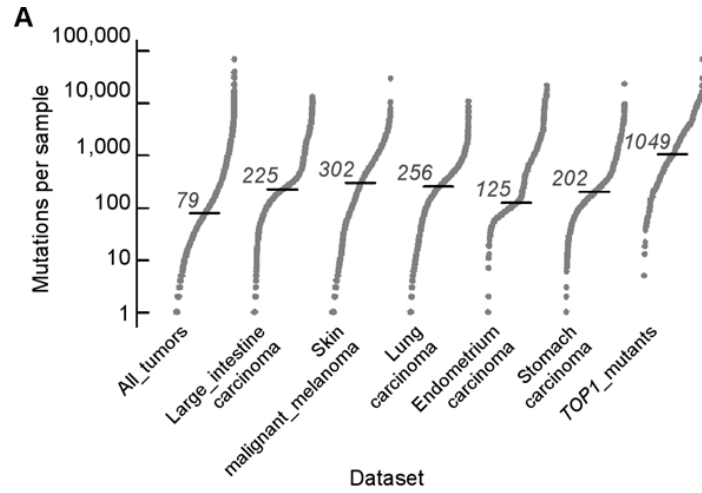
**Top1mt, Nsr1, and G4 DNA form stable complexes that block replication and cause genomic instability**



**Figure 32. Model of co-transcriptional G4-formation and the effect of Top1 activity and mutation on G4-induced genomic instability.** RNAP = RNA polymerase complex. Dotted line = the nascent transcript. (-) = negative tension behind the transcription complex. (+) = positive tension ahead of the transcription complex. Top1mt = Top1 mutant. N = Nsr1 N-term. C = Nsr1 C-term. Top1mts capable of G4-binding (i.e. Top1Y727F or Top1Y740STOP) but not Top1mts incapable of G4-binding (i.e. Top1S733E) form Top1mt/Nsr1/G4 DNA complexes that block and cause genomic instability.



Validating my findings regarding Top1 mutants and G4s in yeast, our collaborator, Dr. Albino Bacolla from the Tainer lab at the University of Texas MD Anderson Cancer Center, performed a computational analysis of the cancer genome sequencing data available in the Catalogue of Somatic Mutations in Cancer (COSMIC) database (<https://cancer.sanger.ac.uk/cosmic>) to show that Top1 catalytic mutants are linked with increased G4-instability in cancers (Figure 33). Specifically, the number of mutations (single nucleotide polymorphisms (SNPs), short insertions, and short deletions (indels)) in 35,887 cancer genomes representing 37 different tissues (All\_tumors-Figure 33A) were compared to the number of mutations present in a group of 239 cancer genomes all harboring mutations in human *TOP1* (*TOP1\_mutants*-Figure 33A). The *TOP1\_mutants* group had a 12-fold higher median number of mutations per genome relative to the All\_tumors group. The *TOP1\_mutants* group also had a higher median number of mutations compared to the median numbers of mutations in 5 different tumor types displaying the greatest number of *TOP1* mutants (large intestine carcinoma, skin malignant melanoma, lung carcinoma, endometrium carcinoma, and stomach carcinoma).



**Figure 33. Somatic mutations in the *TOP1* gene are associated with high mutation rates in cancer.** This bioinformatic analysis was conducted by Dr. Albino Bacolla of the Tainer lab at the University of Texas MD Anderson Cancer Center. (A) S-plots of number of mutations exome-wide. *All\_tumors*, 35,887 samples from Cosmic v.94 comprising all exome-wide screens (i.e. field “Genome-wide screen” corresponding to “y”) and non-redundant sample codes; *Large intestine carcinoma*, 2,355 samples from *All\_tumors* comprising carcinomas of the large intestine, 44 of which had mutations in *TOP1*; *skin\_malignant\_melanoma*, 1,372 samples from *All\_tumors* with malignant melanoma of the skin, 31 with mutations in *TOP1*; *Lung\_carcinoma*, 2,512 lung carcinoma samples from *All\_tumors*, 27 with *TOP1* mutations; *Endometrium\_carcinoma*, 606 samples from *All\_tumors* with carcinoma in the lining of the womb, 23 with mutations in *TOP1*; *Stomach\_carcinoma*, 1,349 samples from *All\_tumors* with stomach carcinoma, 16 with *TOP1* mutations; *TOP1\_mutants*, all 239 samples from *All\_tumors* with mutations in *TOP1*. Horizontal dash, median. (B) Box plot shows number of mutations in samples carrying mutations in different *TOP1* domains. *N-Ter*, 61 samples with mutations in the amino terminus domain (median = 878); *Linker*, 26 samples with mutations in the Linker (median = 1321.5); *Core*, 102 samples with mutations in the Core domain (median = 1,149); *Stop*, 36 samples with nonsense mutations (median = 1,219); *C-Ter*, 14 samples with mutations in the carboxyl terminus domain (median = 982.5); *Ran*, 300 samples chosen at random among *All\_tumors* (median = 73.5); *Ran\_H*, random\_high: a pool of 1,500 random samples with at least 400 mutations each were chosen from *All\_tumors* and 300 entries were then chosen from the pool, after removing samples with identical codes but assigned to different types of tumor in COSMIC (median = 1049). P-values were from

Wilcoxon tests. For the purpose of single Wilcoxon tests, we combined the numbers of mutations and numbers of samples when applicable. (C) Box plots of mutations at G4 tracts. For each sample the value refers to the percent mutations that overlapped with G4-forming repeats. Data sets are as in panel B. *Stars*, total number of mutations; median values are shown. P-values from Wilcoxon tests. For the purpose of single Wilcoxon tests, we combined the percent mutations at G4 and numbers of samples when applicable. Outliers were removed for clarity.

Next, the *TOP1*\_mutant tumors group was divided into 5 groups consisting of genomes with mutations in either TOP1's amino-terminus (N-Ter), linker domain, core domain, carboxy-terminus (C-Ter) as well as TOP1 truncation mutations (Stop). Since the carboxy-terminus of TOP1 contains amino acid residues important for catalytic function, we reasoned that C-Ter and Stop mutations are most likely to affect TOP1's ability to cleave DNA. The comparison between the N-Ter + Linker + Core group with the C-Ter + Stop group resulted in no significant difference in the median number of mutations (Figure 33B). The N-Ter + Linker + Core group as well as the C-Ter + Stop group exhibited significantly higher median numbers of mutations compared to a set of 300 cancer genomes selected randomly from All\_tumors. However, both N-Ter + Linker + Core and C-Ter + Stop groups had median numbers of mutations similar to the group Ran\_H, another set of genomes from All\_tumors harboring a median number of mutations similar to that of *TOP1*\_mutants.

While the data in Figures 33A and B show that TOP1 mutant genomes are hypermutated in general, we wanted to know how TOP1 mutants affect instability occurring at G4 potential non-B DNA-forming sequences (PONDS). Since replication stress through G4-PONDS is expected to elevate SNPs and indels through mutagenic DNA break repair (Malkova and Ira, 2013; Scully et al., 2019; Eckelmann et al., 2020), the percent mutations at G4 PONDS was further examined. When all Top1 mutant cancer genomes were grouped together, the percentage of mutations occurring at G4 PONDS was higher than the G4 PONDS mutations percentage of the Ran group (medians 0.52 and 0.00, respectively) (Figure 33C). When the N-Ter + Linker + Core Top1 mutant group was compared to Ran\_H, a lower percentage of G4 PONDS mutations was observed for N-Ter + Core + Linker. The combined C-Ter + Stop group

had a percentage of mutations at G4 PONDS similar to the Ran\_H group, but significantly higher than the N-Ter + Core + Linker group (p-value = 0.0184). Therefore, we conclude that mutations that impact TOP1 catalytic function affect G4-instability in cancers more than mutations present in other TOP1 domains. This supports that expression of TOP1 cleavage-defective mutants may worsen cancer patient prognosis as a result of G4-stabilization.

In summary, I have found that expression of Top1 mutants, some of which are found in CPT-resistant cancer cells, sharply increases the genomic instability associated with co-transcriptionally-formed G4s in yeast. A model of genome instability at G4 DNA exacerbated by the cleavage-defective Top1 mutants is shown in Figure 32. While co-transcriptionally-formed negative supercoils accumulate in the absence of functional Top1 due to either complete loss of the Top1-encoding gene or mutations leading to defects in DNA binding or DNA cleavage, the G4-binding and -stabilization by the cleave-defective Top1 mutants further enhances the instability and recombination occurring at G4-forming genomic sites. I also discovered a new role of Wss1 in suppressing G4-associated genomic instability in presence of Top1 cleavage-defective mutants, putatively by removing Top1 mutants trapped on co-transcriptionally formed G4s. Another important finding is that the instability at G4 DNA is exacerbated by the interaction between yeast-nucleolin (Nsr1) and Top1 mutants. The findings reported here are clinically relevant since Top1 mutants arise in cancer cells in response to treatment with CPT or CPT-derivatives (Beretta et al., 2013) and human nucleolin are frequently overexpressed or mis-regulated in cancer cells (Berger et al., 2015; Koutsioumpa and Papadimitriou, 2014; Huang et al., 2019). The clinical relevance is further underscored by Albino's finding that mutations in Top1 correlate with high mutation

frequencies throughout the genome, and that mutations in the catalytic carboxy terminal domain of Top1 correlate with enrichment of mutations at G4 PONDSD. Overall, our results suggest that the expression of Top1 mutants could induce additional genome rearrangements in cancer cells by supporting G4-formation and -stabilization. The resulting genomic rearrangements originating at G4-motifs may lead to secondary cancer development greatly complicating patient treatment. Other studies have documented secondary cancer development in patients following treatment with CPT-derivatives (Merrouche et al., 2006; Li et al., 2019). In the future, it will be valuable to explore how CPT-treatment and subsequent emergence of Top1 mutants can lead to further genome instability and potential secondary cancers.

## **Chapter 5: Discussion and Future Directions**



## 5.1 Summary

Many protein factors that regulate DNA topology, transcription, and chromosome packaging interact with G-quadruplexes and impact the genomic instability associated with them. The dissertation work described here set out to uncover how loss of Top1 activity, either through Top1-deletion or mutation, enhances the genomic instability associated with G-quadruplex non-B DNA structures in *S. cerevisiae*. This work also explored if and how the activities of different yeast proteins (Wss1 and Nsr1) enhance or suppress G4-induced instability in the absence of functional Top1.

Chapter 3 describes my attempts at determining if a greater number of G4s form in the highly-transcribed regions of the yeast genome in the absence of Top1 than in Top1's presence. While my approach of expressing the G4-specific antibody BG4 from a vector in yeast to perform ChIP-seq and immunofluorescence for G4 quantification was unsuccessful, I discuss ways to improve both ChIP-seq and immunofluorescence approaches. I also discuss alternative approaches to test our hypothesis that more G4s form in a *TOP1*-deletion strain than in a *WT* strain. In Chapter 4 of this dissertation, I explore expressing of "loss of function" mutants of Top1 impacts G4-induced genomic instability in yeast and find that Top1 cleavage-defective mutants enhance G4-induced genomic instability significantly more than a Top1 duplex DNA binding mutant. While the Wss1 protease of the DNA-protein crosslink repair pathway plays a role in preventing G4-induced recombination in yeast strains expressing Top1 cleavage-defective mutants, the G4-stabilizing protein Nsr1 works together with Top1 cleavage-defective mutants to enhance G4-instability in a synergistic fashion. The results presented in Chapter 4 are clinically relevant as Top1 cleavage-defective mutants are found

in cancer cells resistant to the anticancer Top1-targeting drug CPT. Collectively, the work described in Chapter 4 indicates that Top1 mutants found in chemotherapy-resistant cancer cells could induce secondary genomic rearrangements involving G4-capable loci. In the next sections of this chapter, I discuss future directions of the work outlined in this dissertation.

## **5.2 Enumeration of G4s in the complete absence of Top1 in Yeast**

In light of multiple pieces of evidence showing human and yeast Top1 suppress co-transcriptional G4-induced genomic instability (Kim and Jinks-Robertson, 2011; Yadav et al., 2014; Yadav et al., 2016; Husain et al., 2016), we hypothesized that a greater number of G4 DNA structures form during transcription in the absence of functional Top1. Understanding how Top1 depletion impacts G4-formation on a genome-wide scale is important since Top1 is the target of widely-used anticancer drugs that can lead to loss of function Top1 mutants (Pommier et al., 2010) and G4s are associated with oncogenic translocations and contribute to the mutational burden of cancer cells (Bacolla et al., 2016; Bacolla et al., 2019).

Here, I attempted to express a FLAG-tagged version of the BG4 antibody to perform ChIP-seq and immunofluorescence experiments to compare the number of G4s that form in *WT* and *top1Δ* cells genome-wide. I expected to observe significantly more BG4 enrichment at highly transcribed regions of the yeast genome in *TOP1*-deletion cells relative to *WT* cells in ChIP-seq. I also expected to observe more BG4 puncta in the nuclei of *TOP1*-deletion cells than in the nuclei of *WT* cells (Figure 12). While BG4-FLAG was expressed in yeast (Figure 8) and was found to be functional in terms of G4-binding *in vitro* (Figure 10), ChIP experiments revealed that BG4-FLAG expressed from pGAL-BG4-FLAG was not enriched at our model *SμG4*-motif in yeast (Figure 11). The immunofluorescence experiments performed in cells

transformed with pGAL-BG4-FLAG were also unsuccessful (Figure 13). One reason the ChIP and immunofluorescence experiments failed could be that BG4-FLAG expressed from pGAL-BG4-FLAG was not localized to the nucleus. Thus, a nuclear localization signal sequence should be added to the expressed BG4 construct in effort to increase the likelihood that BG4 will contact G4s in the yeast genome. Additionally, purified BG4-FLAG could be used for ChIP-seq and immunofluorescence experiments.

In addition to optimizing methods to quantify the number of G4s in yeast in the absence of Top1, the genome-wide formation of other non-B DNAs in Top1-deficient cells should be quantified in the future as well. Multiple lines of evidence indicate that Top1 suppresses the genomic instability associated with other non-B DNA structures by removing co-transcriptional negative helical tension. For example, R-loops are formed during transcription when the nascent mRNA loops back to hybridize with the transcribed DNA strand and are a source of genomic instability (Hamperl and Cimprich, 2014). TopA in *E. coli* and Top1 in yeast are linked to the suppression of R-loop mediated cytotoxic effects. TopA is the enzyme responsible for removing DNA negative supercoils in *E. coli* (Drolet, 2006). In the absence of TopA function, the accumulation of negative supercoils can perturb *E. coli* growth and RNA production (Drolet et al., 1994; Baaklini et al., 2004; Baaklini et al., 2008). Because TopA null *E. coli* RNA synthesis and growth defects are rescued in part by overexpression of the R-loop resolving enzyme RNase H, it was concluded that excessive DNA negative supercoils promote R-loop formation in *E. coli* (Drolet et al., 1995; Baaklini et al., 2004). In yeast, ChIP experiments from El Hage et al., 2010 showed that R-loop formation at ribosomal DNA in cells conditionally depleted of Top1 is increased in the absence of R-loop processing

enzymes RNase H1 and RNase H2. Electron microscopy experiments from this same work revealed that pile ups of RNA polymerase I on ribosomal DNA occur more frequently in cells lacking Top1 than in WT cells, suggesting that increased R-loop formation in the absence of Top1 blocks RNA polymerase I progression. It has also been shown that Top1 suppresses genomic instability associated with repetitive DNA loci capable of forming DNA hairpin structures. DNA sequences containing multiple CAG/CTG repeats can form stable DNA hairpins composed of intra-strand base pairs (Petruska et al., 1996). In human cells, knock down of *TOP1* with siRNAs resulted in significantly increased contraction of a trinucleotide repeat track containing 95 CAG repeats (Hubert et al., 2011). Of note, the instability of the CAG repeat track in *TOP1* knock down cells was dependent on transcription. Since Top1 suppresses genomic instability associated with CAG repeats, it is possible that a greater amount of hairpin structures form at inverted repeats on a genome-wide scale in the absence of Top1. Cruciforms (two DNA hairpins located opposite to one another in dsDNA), triplexes (three-stranded DNA), and Z-DNA (left-handed DNA helix) are other non-B DNA structures that form in negatively supercoiled DNA (Zhao et al., 2010); therefore, *TOP1*-depletion is expected to increase their formation as well.

Further, the impact of other Topoisomerases, such as Topoisomerase 2 and Topoisomerase 3, on the prevention of formation of G4s and other non-B DNAs should be examined. Yeast topoisomerase 2 (Top2) is a type 2A topoisomerase that cleaves both strands of duplex DNA and relieves helical torsion via the passage of an intact DNA duplex through the cut DNA duplex (Pommier et al., 2016). In addition to relieving both positive and negative DNA supercoils, yeast Top2 is essential due to its ability to decatenate sister chromatids that

become interlocked during replication. A temperature sensitive Top2 mutant (*top2-ts*) that exhibits no Top2 activity at 35 °C and very weak Top2 activity at 26 °C has allowed for researchers to investigate how Top2-depletion affects helical stress (Goto and Wang, 1985; Christman et al., 1988; Trigueros and Roca, 2001). In yeast strains either lacking Top1 or expressing *top2-ts*, transcription-induced DNA supercoiling of the rDNA repeats was visualized using high-resolution microscopy and showed that Top1 primarily relieves DNA negative supercoils accumulating behind RNA polymerase I while Top2 primarily relieves DNA positive supercoils that accumulating in front of RNA polymerase I (French et al., 2011). Although the study mentioned above indicates distinct roles of Top1 and Top2 in resolving negative and positive supercoils, respectively, Top2 does have the ability to relieve DNA negative supercoils and is thought to play redundant roles in cells as Top1 when Top1 is absent (Trigueros and Roca, 2002; Pommier et al., 2016). The redundant role of Top2 in the relief of transcription-associated negative supercoils is underscored by results showing that the slow growth phenotype of yeast cells depleted of both Top1 and Top2 function at the same time is rescued by the controlled ectopic expression of either Top1 or Top2 alone (Trigueros and Roca, 2002). Interestingly, overexpression of *E. coli* TopA, an enzyme that exclusively removes negative supercoils from DNA, also rescues the slow growth phenotype of the *top1Δ top2-ts* yeast cells at 26 °C, tying this observed growth defect to negative helical torsion (Trigueros and Roca, 2002). This piece of data taken together with the observation that *TOP1*-deletion yeast cells do not display a slow growth phenotype suggest that Top2 does act to relieve negative helical tension, at least partially, in the absence of Top1. Therefore, the potential extra negative supercoil accumulation occurring during Top2

depletion in the absence of Top1 may increase R-loop and/or G4-formation in yeast and should be explored in the future.

Yeast Top3 is a type 1A topoisomerase that uses  $Mg^{2+}$  as a metal cofactor to perform its catalytic cycle (Pommier et al., 2016). Top3 can resolve DNA negative supercoils, hemicatenanes, double Holliday junctions, and displacement (D-loops). Vegetatively growing yeast cells can tolerate *TOP3*-deletion, while cells reproducing sexually cannot complete meiosis in the absence of Top3 (Wallis et al., 1989; Gangloff et al., 1999). Loss of Top3 $\beta$  in human cells led to an increase in R-loops, indicating that human Top3 plays a role in suppressing R-loop formation and/or stability (Zhang et al., 2019). Interestingly, yeast Top3 was shown to have a strong binding preference for single-stranded DNA in an experiment utilizing purified protein and a negatively supercoiled heteroduplex DNA molecule containing a 29 base pair melted region (Kim and Wang, 1992). Further supporting yeast Top3 as a single-stranded DNA specific topoisomerase, yeast Top3 was shown to dissolve D-loops that were formed by Rad51 and Rad54 activities (Fasching et al., 2015). D-loops, while made of only DNA, are similar to R-loops in that they comprise a displaced single DNA strand (Kasamatsu et al., 1971). Since human Top3 $\beta$  can resolve R-loops and yeast Top3 can resolve D-loops and has reported specificity for single-stranded DNA substrates, it is possible that deletion of yeast *TOP3* will result in increased formation of R-loops. Although elevated recombination at G4 in the absence of Top1 is not dependent on R-loop-stability as overexpression of RNase H1 does not reduce G4-recombination at the *S $\mu$ G4-GTOP* reporter in yeast (Yadav et al., 2014), increased R-loop-formation due to Top3 absence could still increase the chances of

G4-formation. Therefore, it will be interesting to uncover if Top3 plays a role in suppressing G4-formation/instability.

Lastly, in addition to preventing the formation of G4s and other non-B DNAs, Top1 may play an important role in the resolution of formed G4s. Human Top1 interacts with the SV40 T antigen, which unwinds duplex DNA as well as G-quadruplex DNA (Stahl et al., 1986; Baran et al., 1997; Haluska and Rubin, 1998; Tuesuwan et al., 2008). Interactions between human Top1 and the Werner G4-helicase have also been documented (Lebel et al., 1999; Mendoza et al., 2016). Thus, it is possible that human Top1 recruits Werner and/or SV40 T antigen helicases to G4s through its capability to bind G4s (Arimondo, 2000; Marchand et al., 2002). The yeast homolog of the Werner helicase is Sgs1, which also interacts with Top1 (Watt et al., 1996; Mankouri and Morgan, 2001). Therefore, the relevance of the Top1/Sgs1 interaction in suppressing G4-induced genomic instability should be further explored as well as other potential Top1/G4-resolvase interactions.

### **5.3 Role of Cleavage-Defective Top1 Mutants in G4-Induced Genomic Instability**

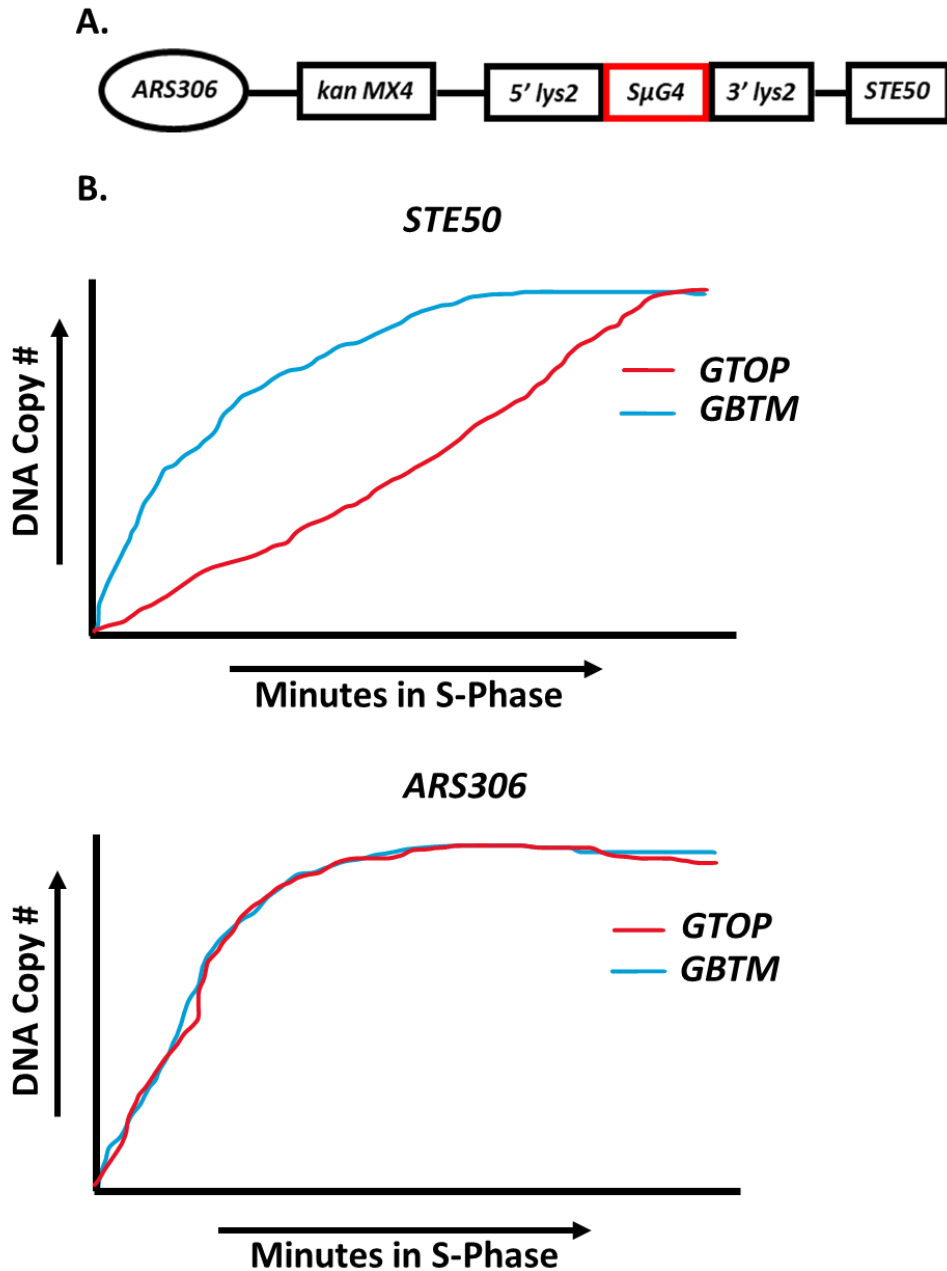
As described above, *TOP1*-deletion significantly enhances recombination at the highly-transcribed *S $\mu$ G4-GTOP* locus (Kim and Jinks-Robertson, 2011; Yadav et al., 2016). This increase in G4-instability is hypothesized to be the result of increased G4-formation due to rises in co-transcriptionally-derived DNA negative supercoils caused from lack of Top1 function. In past work, our group uncovered that expression of the cleavage-defective yeast Top1 mutant, *yTop1Y727F*, increases co-transcriptional G4-recombination even more than *TOP1*-deletion (Yadav et al., 2016). Since the same levels of DNA negative supercoils are expected to form in *top1 $\Delta$*  and *yTOP1Y727F* cells, the finding that expression of a Top1

catalytic mutant induces greater levels of G4-instability in yeast was surprising. Further investigation revealed that yTop1Y727F binds G4s *in vitro* (Berroyer and Kim, 2020) (Figure 3). Thus, I hypothesized that G4-binding of yTop1Y727F stabilizes G4s and leads to the observed hyper-recombination at the *SμG4-GTOP* construct. Instability and recombination at G4 is acutely elevated by the expression of a cancer-associated catalytically inactive Top1 mutant (yTop1Y740Stop) capable of binding both duplex DNA (Figure 18) and G4 DNA (Figure 19A-B). This suggests that cleavage-defective Top1 mutants induce genomic instability at G4s by a shared mechanism.

Since the tight-binding of proteins to DNA can block DNA replication and cause DNA breaks (Gadaleta et al., 2017), I further hypothesize that the exacerbated G4-associated recombination observed in cleavage-defective Top1 mutant yeast cells is caused from replication stress induced from Top1 mutants binding at G4s. This hypothesis should be tested in the future via a technique developed in the Kim lab to assess kinetics of DNA copy number change during S-phase. This new method was used to show expression of full-length Nsr1, but not Nsr1ΔRGG, results in a significant delay in replication timing through a G4-motif (*SμG4-GTOP*) relative to a non-G4-motif (*SμG4-GBTM*) in the absence of Top1 (Singh et al., 2020). This indicates Nsr1-G4 binding induces replication stress through the *SμG4-GTOP* locus. In this experiment, yeast cells were first arrested at G1 with  $\alpha$ -factor followed by release into S-phase. Genomic DNA was isolated from the cells at 10-minute intervals from 1 to 120 minutes following release from G1 arrest. Isolated DNA from every time point was then subjected to droplet digital PCR (ddPCR), a technique used to quantify the absolute number of nucleic acid molecules present in a sample (Batrakou et al., 2018), and assessment



of DNA copy number throughout S-phase was used to determine replication kinetics (i.e. time it takes cells to duplicate DNA at a targeted genomic locus). To measure the absolute copy number of genomic loci surrounding *SμG4* throughout S-phase in cleavage-defective Top1 mutant backgrounds, primers targeting multiple loci located at and downstream of *ARS306* on yeast chromosome III can be used in ddPCR experiments. Locations primers can target upstream of *SμG4* (*ARS306* and *kanMX4*) or downstream of *SμG4* (3' *lys2* and *STE50*) on chromosome III are shown in Figure 34A. If binding of cleavage-defective Top1 mutants to G4s induces replication stress, I expect a significant lag in replication will be observed in *GTOP* cells relative to *GBTM* cells at loci downstream of *SμG4* (3' *lys2* and *STE50*) (Figure 34B). Conversely, no differences in replication kinetics are expected between *GTOP* and *GBTM* in Top1 mutant strains at loci upstream of *SμG4* (*ARS306* and *kanMX4*) (Figure 34B).



**Figure 34. Expected results of proposed ddPCR replication kinetics experiment.** A. Schematic of chromosome III showing locations of *ARS306* and loci upstream and downstream of *SμG4* to be targeted in ddPCR. B. Expected outcome of ddPCR experiment of 2 loci presented as DNA copy number of cell populations throughout S-phase in Top1 cleavage-defective mutant *GTOP* and *GBTM* strains.

As described in Chapter 4, I also found that Wss1 of the DNA protein crosslink removal pathway plays a role in suppressing G4-induced genomic instability in yeast strains expressing cleavage-defective Top1 mutants. Yeast Wss1 and its human homolog SPTRN are proteases that degrade replication blocking protein DNA complexes (Stingele et al., 2014; Stingele et al., 2016). Wss1 harbors a SUMO-interacting motif and degrades SUMOylated proteins (Mullen et al., 2010). Since I found that Wss1 partly suppresses recombination at *SμG4-GTOP* in Top1 catalytic mutant backgrounds (Figure 23B) and that Top1 catalytic mutants are SUMOylated (Figure 23C), Wss1 likely targets the SUMOylated form of cleavage-defective Top1 mutants. Therefore, future experiments should be conducted with lysates from *WSS1+* and *WSS1-* yeast cells to determine if the presence of Wss1 reduces the cellular levels of SUMOylated Top1 catalytic mutants (yTop1Y727F and yTop1Y740Stop). Further, Top1 has been shown to be SUMOylated at three specific amino acid residues located in the N-terminus of the enzyme: Lys65, Lys91, and Lys92 (Chen et al., 2007). Mutation of all three lysines to arginines results in a ~95% reduction in Top1 SUMOylation. In order to probe the relevance of Top1 mutant SUMOylation in the suppression of G4-induced genomic instability, all three Top1 SUMOylation sites should be mutated to arginines and its effect on recombination at *SμG4-GTOP* and *SμG4-GBTM* should be determined. Yeast SUMO ligases Siz1 and Siz2 (Jalal et al., 2017) should also be deleted in Top1 mutant yeast strains to uncover if their ability to SUMOylate Top1 is relevant in the suppression of G4-induced genomic instability.

The Kim lab has previously shown that Nsr1 binds G4s and instigates G4-induced genomic instability in the absence of yeast Top1 (Singh et al., 2020). I found that Nsr1 also

contributes to G4-recombination in the presence of Top1 mutants (Figure 24C-D). Since deletion of the C-terminal RGG domain of Nsr1, which is responsible for Nsr1-G4 binding, significantly reduces G4-recombination in Top1 mutant expressing strains (Figure 24C-D), I conclude that G4-binding is the mechanism of Nsr1-induced G4-instability in Top1 mutant cells. While *NSR1*-deletion or truncation of Nsr1 reduces *GTOP* recombination to *WT* levels in *top1Δ* and *yTOP1S733E* strains, deletion or truncation of Nsr1 only reduces *yTOP1Y727F* and *yTOP1Y740Stop* *GTOP* recombination to *TOP1*-deletion levels (Figure 24B-D). This indicates that while Top1 cleavage-defective mutants and Nsr1 are a highly mutagenic combination in terms of the instigation of G4-induced genomic instability, expression of Top1 cleavage-defective mutants alone can cause G4-associated recombination. I also found that the N-terminus of Nsr1 is required for induction of Nsr1-mediated G4-induced genomic instability in the *yTOP1Y727F* and *yTOP1Y740Stop* cleavage-defective mutant strains (Figure 26C-D). This suggests that the physical interaction of Nsr1 and Top1 mutants as well as binding of G4 DNA by both proteins is necessary to exert a synergistic effect on elevating G4-associated recombination (Figure 32). To further elucidate the mechanism of G4-induced genomic instability in cells expressing both Top1 cleavage-defective mutants and Nsr1, future experiments should be conducted to determine if Top1 mutants and Nsr1 bind the same G4-molecule at the same time. Such future experiments could include a sequential pull down where photocleavable biotinylated G4-oligos conjugated to streptavidin magnetic beads are first mixed with yeast cell lysates containing FLAG-tagged Top1 mutants to separate G4-binding proteins from non-G4 binding proteins. Following photocleavage to separate G4-oligos from the streptavidin magnetic beads, a second pull down can be done with  $\alpha$ -FLAG

coated agarose beads to pull down G4 oligos bound by FLAG-tagged Top1 mutants. After proteins are eluted from oligos and  $\alpha$ -FLAG beads, western blotting utilizing  $\alpha$ -FLAG and  $\alpha$ -Nsr1 antibodies can be performed to determine if Top1 mutants and Nsr1 bind G4 oligos simultaneously. Additionally, Nsr1-ChIP should be conducted to determine whether enrichment of Nsr1 at G4 loci is enhanced by its interaction with Top1 catalytic mutants.

One of the ways cells combat G4-induced genomic instability is through the resolution of G4s by a subset of helicases. ChIP-seq experiments performed in *S. cerevisiae* revealed that the G4-unwinder Pif1 is enriched at G4-motifs and that deletion of Pif1 results in replication stress and increased DNA breaks at G4-motifs (Paeschke et al., 2011). Additionally, while it is hard to imagine G4 helicases are able to unwind G4s in the presence of cleavage-defective Top1 mutants, past work from our group suggests Pif1 may resolve a portion of G4s in Top1 catalytic mutant cells. Lopez et al., 2017 demonstrated that the yeast transcription factor Sub1, homolog to human PC4, binds G4s and prevents G4-induced recombination in *TOP1*-deletion cells. Genetic experiments further showed that Sub1 likely suppresses G4-instability by recruiting G4-helicase Pif1 to co-transcriptionally formed G4s so they can be unwound. Interestingly, deletion of *SUB1* significantly increases G4-induced recombination in yeast in the *yTOP1Y727F* background (Lopez et al., 2017). This means it is possible that Sub1 recruits Pif1 to G4s in the presence of Top1 cleavage-defective mutants. While it remains to be tested if Pif1 can metabolize G4s that are bound by Top1 catalytic mutants, Pif1 may be able to access G4s after Wss1 has proteolytically removed Top1 cleavage-defective mutants from the structure. Further work needs to be done to elucidate how Sub1 reduces G4-instability in the *yTOP1Y727F* background and if Pif1 is involved. To start, *PIF1* can be deleted in *yTOP1Y727F*

and *yTOP1Y727F sub1Δ* backgrounds to uncover if Sub1 and Pif1 work together to suppress G4-induced recombination. Future work should also determine if Sub1 impacts G4-associated recombination in the cancer mutant homolog *yTOP1Y740Stop* background.

#### **5.4 Yeast Chromatin Remodelers and G4-Induced Genomic Instability**

In addition to being tied to transcriptional regulation, G4s are also linked to epigenetics and chromatin structure. First off, G4-formation can impact the epigenetic modification of DNA bases. Cytosine residues can be methylated by DNA methyltransferase enzymes (Li et al., 1992; Okano et al., 1999; Liao et al., 2015). One DNA methyltransferase, DNMT1, binds G4s both *in vitro* and *in vivo*, and DNMT1 G4-binding inhibits DNMT1's catalytic activity (Mao et al., 2018). Interestingly, G4s were found to form near sparsely methylated CpG islands in the human genome, therefore, it was concluded that G4-formation can prevent CpG island methylation by preventing DNMT1 activity. Since cytosine methylation can impact transcription by affecting DNA structure and ability of transcription factors to bind to their target sites (Yin et al., 2017), any change in cytosine methylation resulting from G4-formation could impact transcription. Therefore, loss of Top1 function could impact transcription by promoting G4-formation and consequently decreasing cytosine methylation. G4-formation and -stability can also impact the presence of modified histone proteins in the genome. During DNA replication, histone marks of parental nucleosomes are maintained by chaperone proteins that insert the modified parental nucleosomes into the newly synthesized DNA (Alabert et al., 2017; Hammond et al., 2017). Stabilized G4s can block the progression of DNA polymerases, leading to post-replicative gaps where the replication of problematic G4-capable motifs becomes uncoupled from bulk DNA replication when replication restarts

downstream of the G4-block (Sarkies et al., 2010; Sarkies et al., 2012). Such uncoupling of replication is accompanied by the deposition of new nucleosomes lacking parental histone modifications into G4-capable genomic regions, resulting in a loss of epigenetic information that impacts transcription. Depletion of Rev1, a translesion polymerase associated with the destabilization of G4s (Eddy et al., 2014), and of FANCI, a helicase that is capable of unwinding G4s (Wu et al., 2008), leads to loss of parental epigenetic marks at G4-motifs in cells; thus, loss of Top1 may impact the recycling of modified parental histones in G4-motifs during replication as well since *TOP1*-deletion increases G4-instability. Other data showing that 98% of G4s in human keratinocytes are located in nucleosome free genomic regions (Hänsel-Hertsch et al., 2016) suggest that G4s may prevent nucleosome deposition. The suppression of nucleosome assembly on DNA by G4s located near TSSs could explain the observed association of genes harboring G4s with increased transcriptional output (Du et al., 2008). Additionally, since G4s can exclude nucleosomes, increased formation and stabilization of G4s in the absence of functional Top1 could greatly impact the chromatin landscape of cells.

While G4s may serve to exclude nucleosomes from DNA and affect epigenetic modification of DNA bases and histones, thus potentially impacting the overall architecture of chromatin, multiple studies have documented connections of chromatin remodeling proteins with G4s and other non-B DNAs. For example, ATRX was shown to suppress the formation of R-loops at telomeric DNA sequences and bind to G-quadruplex (G4) DNA structures *in vitro* and *in vivo* (Law et al., 2010; Nguyen et al., 2017). ATRX is a chromatin remodeler that belongs to the SWI/SNF family that functions in complex with the histone chaperone DAXX to suppress transcription by insertion of the H3.3 histone variant into DNA

(Argentaro et al., 2007; Wong et al., 2010; Lewis et al., 2010). Interestingly, ATRX's genomic localization is enriched at centromeric, pericentromeric, ribosomal, and telomeric DNA sequences which all contain tandem repeats and have the ability to adopt non-B DNA structures (McDowell et al., 1999; Gibbons et al., 2000; Law et al., 2010; De La Fuente et al., 2011; Elsasser et al., 2015). In addition to being enriched at telomeres, ATRX has also been shown to suppress the alternative lengthening of telomeres (ALT) pathway (Heaphy et al., 2011), a mutagenic mechanism involving the homology directed repair of broken DNA that cancer cells use to preserve the length of their telomeres (Cesare and Reddel, 2010). Further, since telomeres contain G4-capable DNA sequences, it is reasonable to postulate that ATRX may have some connection to G4-mediated genomic instability. Nguyen et al., 2017 investigated the recruitment of ATRX to G4-capable telomeric DNA repeats that were inserted into an intron of a genomic *GFP* gene controlled by a doxycycline-inducible promoter in mouse cells. Usage of this reporter system containing G4-capable telomeric repeats lead to the discovery that ATRX only bound co-transcriptionally-formed G4s when guanine-rich DNA is present on the non-transcribed strand (NTS), which is consistent with other studies that showed the interaction of bona-fide G4-binding proteins with G4-capable DNA is dependent on transcriptional orientation (Lopez et al., 2017; Singh et al., 2020). Additionally, Nguyen et al., 2017 observed that the levels of ATRX *in vivo* G4-binding were dependent on the number of telomeric repeats present in the GFP gene, where longer tracts of telomeric repeats with a high propensity to form G4s resulted in greater ATRX enrichment than shorter telomeric repeat tracts. Based on this finding, Nguyen et al., 2017 proposed that ATRX is recruited to telomeres by G4s, and this recruitment may play a role in suppressing ALT. A

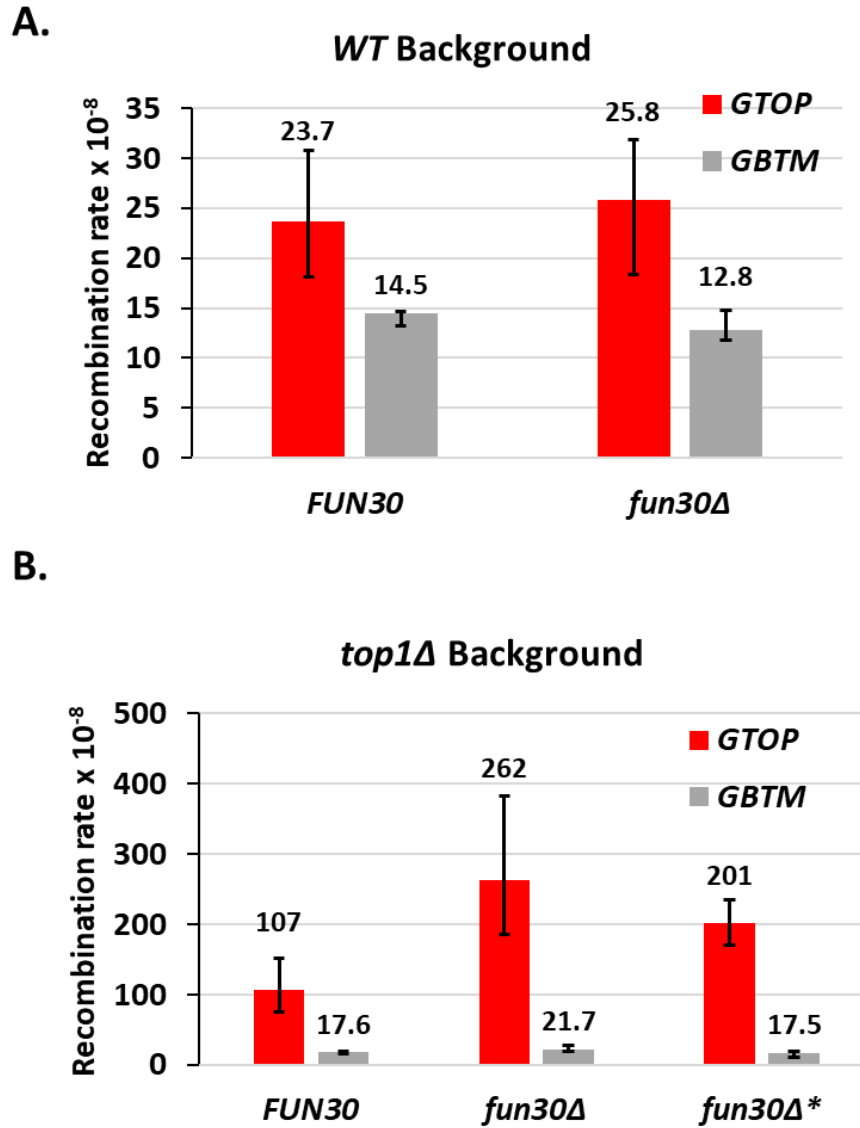


separate study by Teng et al., 2021 revealed that ATRX and DAXX insert H3.3 into DNA harboring G4-motifs to promote the formation heterochromatin which ultimately suppresses the genomic instability associated with G4s by hindering G4-formation (Teng et al., 2021b). The two aforementioned studies taken together with results demonstrating that ATRX exogenous expression decreases the number of G4s in cells (Wang et al., 2019) supports a model where ATRX may have a role in G4-resolution and in the suppression of G4-formation to protect the genome from replication stress.

Another study implemented the chromatin remodeler SMARCA4 in suppressing G4-induced genomic instability. SMARCA4 is an ATP-dependent chromatin remodeler that plays roles in the regulation of transcription and DNA repair (Chetty and Serra, 2020). Interestingly, SMARCA4 has been documented to play both tumor suppressive and tumor supportive roles in cancers (Ramos et al., 2014; Kim et al., 2021). Interestingly, a study showed that SMARCA4 siRNA knock down results in increased *Igh/c-Myc* translocations in mouse B-cell lymphoma cells (Husain et al., 2016). Because the *Igh/c-Myc* translocation is a G4-associated translocation where G4s form at the *Igh* and *c-Myc* break sites (Duquette et al., 2004; Sun and Hurley, 2009), SMARCA4 is linked to G4 DNA. Experiments assessing the *in vivo* enrichment of Top1 at the *Igh* locus showed that Top1 recruitment to *Igh* was significantly reduced when SMARCA4 is depleted (Husain et al., 2016). It was also shown that knock down of either SMARCA4 or Top1 increased negative supercoil levels of the highly transcribed *Igh* locus in mouse B-cells. Therefore, it was concluded that SMARCA4 recruits Top1 to G4-capable loci so that Top1's activity of resolving DNA negative supercoils suppresses G4-formation and associated DNA breaks.

It remains unknown if yeast chromatin remodelers are involved in protecting the genome from G4-mediated DNA damage. Thus, I set out to determine if deletion of yeast chromatin remodelers influences recombination at the *GTOP* and *GBTM* reporters. Since yeast chromatin remodeler Fun30 was identified as a G4-interacting protein in whole cell extracts in an *in vitro* G4-oligo pull down assay followed by mass spectroscopy (Kim lab unpublished datum, not shown), I started by exploring if *FUN30*-deletion affects G4-induced genomic instability. Fun30 is a SWI/SNF (SWitch/ Sucrose Non-Fermentable), ATP-dependent chromatin remodeler with roles in histone-dimer exchange, gene silencing, and end resection of double-strand DNA breaks (Flaus et al., 2006; Neves-Costa et al., 2009; Awad et al., 2010; Chen et al., 2012). I found that, in the *WT* yeast background, deletion of *FUN30* did not affect *GTOP* or *GBTM* recombination (Figure 35A). However, when I deleted *FUN30* in the *TOP1*-deletion background, I observed a significant ~2-fold increase in recombination at *S $\mu$ G4-GTOP* (Figure 35B). Since *GBTM* recombination was not affected by *FUN30*-deletion in the absence of Top1, Fun30 plays a role in suppressing recombination that is G4-specific. To verify that the absence of Fun30 increases G4-induced recombination in the absence of Top1, I next deleted *TOP1* in the *fun30 $\Delta$*  background (Figure 35B, *fun30 $\Delta$ \** strain). Again, I saw that the absence of both Top1 and Fun30 significantly increases recombination at the *S $\mu$ G4-GTOP* reporter only. Therefore, I conclude that Fun30 has an unknown role in suppressing G4-induced genomic instability in the absence of Top1 when G4-formation is favored due to increased accumulation of co-transcriptionally formed negative DNA supercoils. In the future, experiments should be conducted to uncover exactly how Fun30 suppresses G4-induced genomic instability. For example, ChIP could be performed to uncover if Fun30 interacts with

G4-capable loci *in vivo*. And *in vitro* experiments could be completed to decipher if Fun30 binds G4-oligos directly. Other areas of future exploration include assessing if Fun30's ATPase and/or helicase activities are required for suppression of G4-induced recombination.



**Figure 35. Deletion of *FUN30* increases G4-induced recombination in *TOP1*-deletion strains.**

A-B. Recombination rates of *fun30Δ* *GTOP* and *GBTM* strains in WT (A) and *top1Δ* (B) backgrounds. Rates for each strain are listed above their respective bars in graph. Recombination rates are considered statistically significantly different if their 95% confidence intervals (shown as error bars) do not overlap (Spell and Jinks-Robertson, 2004). In B, *fun30Δ\** denotes strains where *TOP1* was deleted in a *FUN30*-deletion background.

Since Fun30 deletion alone does not affect G4-induced recombination and only impacts instability when Top1 is also deleted (Figure 35), it is not likely that Fun30 has a role in recruiting Top1 to G4-capable loci undergoing transcription as has been documented with the SMARCA4 chromatin remodeler (Husain et al., 2016). Instead, Fun30 may have another role in suppressing G4-induced recombination related to its documented role in promoting silencing of reporter genes inserted into ribosomal DNA repeats and telomeres in yeast (Neves-Costa et al., 2009). Of note, both ribosomal DNA and telomeres are G4-capable (Capra et al., 2010), indicating it is possible that Fun30's role in gene silencing could involve G4-recognition and/or binding. In support of Fun30 having a possible role in suppressing transcription of G4-capable loci or loci proximal to G4s, mammalian ATRX has been shown to promote the formation of heterochromatin as a mechanism to block transcription and prevent genomic instability at G4-capable genomic loci in murine embryonic stem cells (Teng et al., 2021b). While the exact mechanism of Fun30 gene silencing at the ribosomal DNA loci remains to be elucidated, it is possible that Fun30 functions to repress G4-induced genomic instability in a similar fashion to ATRX by inserting histone proteins into DNA to promote chromatin condensation and repress transcription. In addition, whether SMARCAD1, a human homolog of Fun30 (Awad et al., 2010), represses G4-instability in mammalian cells remains unknown and should be tested in the future.

In addition to Fun30, the role of other yeast chromatin remodelers (i.e. the Ino80 complex that plays a role in repair of DNA double-strand breaks (Morrison et al., 2004; van Attikum et al., 2007), in genome instability at G4s should be investigated. And as mentioned above, murine SMARCA4 was shown to recruit Top1 to G4s and suppress G4-induced

instability (Husain et al., 2016). Therefore, it is possible a yeast chromatin remodeler also recruits Top1 to highly-transcribed G4-capable genomic loci as well. Studying whether the recruitment of Top1 cleavage-defective mutants by SMARCA4 or other chromatin remodelers could contribute to oncogenic translocations at highly-transcribed G4-capable loci is another attractive area of exploration in the future.

## 5.5 Overall Conclusions

The work attempted in chapter 3 is important because uncovering if a higher number of G4s form in the highly-transcribed regions of *TOP1*-deletion yeast cells relative to *WT* yeast cells will contribute to a better understanding of Top1's role in suppressing G4-induced genomic instability. Investigating the genome-wide formation of G4s in the absence of Top1 is also medically relevant since Top1 is a major target of highly used anti-cancer drugs, such as CPT-derivatives (Pommier et al., 2010).

Top1 mutants arise in cancer cells in response to chemotherapeutic treatment (Beretta et al., 2013). Since these mutants are catalytically faulty, their expression is coupled with an increase of DNA negative supercoils levels which impacts non-B DNA-formation and associated genomic instability. I found that Top1 mutants, including some found in cancers, increased G4-induced genomic instability in yeast.

The scientific findings uncovered in Chapter 4 of this dissertation are important since nucleolin is commonly overexpressed in cancers (Carvalho et al., 2021), G4-DNA contributes the mutational burden of cancer cells (Bacolla et al., 2019), and Top1 is a target of widely used chemotherapeutics that cancer cells can acquire resistance to through alterations of Top1 functioning (Pommier et al., 2010; Beretta et al., 2013). A protein complex involving

Top1 mutants and nucleolin bound to G4s could cause increased G4-mutagenesis in cancer cells as a result of DNA replication blockage and potentially lead to the formation of secondary cancers in patients treated with Top1-targeting anticancer drugs. In support of this, while Top1 mutants are correlated with an overall increase in mutations throughout cancer genomes, Top1 catalytic mutants are associated with increased levels of genomic aberrations occurring at G4-capable loci in cancers (Figure 33). This indicates that Top1 mutants that are catalytically defective may negatively impact cancer patient prognosis due to increased G4 DNA stabilization. Moving forward, experiments should be conducted to investigate if cancer cells treated with CPT have increased levels of G4-induced genomic instability relative to untreated cells.

## Bibliography

- AGUILERA, A. & GAILLARD, H. 2014. Transcription and recombination: when RNA meets DNA. *Cold Spring Harb Perspect Biol*, 6.
- ALABERT, C., JASENCAKOVA, Z. & GROTH, A. 2017. Chromatin Replication and Histone Dynamics. *Adv Exp Med Biol*, 1042, 311-333.
- AMRANE, S., ADRIAN, M., HEDDI, B., SERERO, A., NICOLAS, A., MERGNY, J. L. & PHAN, A. T. 2012. Formation of pearl-necklace monomeric G-quadruplexes in the human CEB25 minisatellite. *J Am Chem Soc*, 134, 5807-16.
- ARAKAWA, Y., SUZUKI, H., SAITO, S. & YAMADA, H. 2006. Novel missense mutation of the DNA topoisomerase I gene in SN-38-resistant DLD-1 cells. *Mol Cancer Ther*, 5, 502-8.
- ARGENTARO, A., YANG, J. C., CHAPMAN, L., KOWALCZYK, M. S., GIBBONS, R. J., HIGGS, D. R., NEUHAUS, D. & RHODES, D. 2007. Structural consequences of disease-causing mutations in the ATRX-DNMT3-DNMT3L (ADD) domain of the chromatin-associated protein ATRX. *Proc Natl Acad Sci U S A*, 104, 11939-44.
- ARIMONDO, P., B., RIOU, J., MERGNY, J., TAZI, J., SUN, J., GARESTIER, T., AND HÉLÈNE, C. 2000. Interaction of human topoisomerase I with G-quartet structures. *Nucleic Acids Research*, 28, 4832-4838.
- ARIS, J. P. & BLOBEL, G. 1988. Identification and characterization of a yeast nucleolar protein that is similar to a rat liver nucleolar protein. *J Cell Biol*, 107, 17-31.
- AWAD, S., RYAN, D., PROCHASSON, P., OWEN-HUGHES, T. & HASSAN, A. H. 2010. The Snf2 homolog Fun30 acts as a homodimeric ATP-dependent chromatin-remodeling enzyme. *J Biol Chem*, 285, 9477-9484.



- AZEVEDO, C., LIVERMORE, T. & SAIARDI, A. 2015. Protein polyphosphorylation of lysine residues by inorganic polyphosphate. *Mol Cell*, 58, 71-82.
- 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. *Mol Microbiol*, 54, 198-211.
- BAAKLINI, I., USONGO, V., NOLENT, F., SANSCARTIER, P., HRAIKY, C., DRILICA, K. & DROLET, M. 2008. Hypernegative supercoiling inhibits growth by causing RNA degradation. *J Bacteriol*, 190, 7346-56.
- BACOLLA, A., TAINER, J. A., VASQUEZ, K. M. & COOPER, D. N. 2016. Translocation and deletion breakpoints in cancer genomes are associated with potential non-B DNA-forming sequences. *Nucleic Acids Res*, 44, 5673-88.
- BACOLLA, A., YE, Z., AHMED, Z. & TAINER, J. A. 2019. Cancer mutational burden is shaped by G4 DNA, replication stress and mitochondrial dysfunction. *Prog Biophys Mol Biol*, 147, 47-61.
- BALAGURUMOORTHY, P., BRAHMACHARI, S. K., MOHANTY, D., BANSAL, M. & SASISEKHARAN, V. 1992. Hairpin and parallel quartet structures for telomeric sequences. *Nucleic Acids Res*, 20, 4061-7.
- BARAN, N., PUCSHANSKY, L., MARCO, Y., BENJAMIN, S. & MANOR, H. 1997. The SV40 large T-antigen helicase can unwind four stranded DNA structures linked by G-quartets. *Nucleic Acids Res*, 25, 297-303.
- BARANELLO, L., WOJTOWICZ, D., CUI, K., DEVAIAH, B. N., CHUNG, H. J., CHAN-SALIS, K. Y., GUHA, R., WILSON, K., ZHANG, X., ZHANG, H., PIOTROWSKI, J., THOMAS, C. J.,

- SINGER, D. S., PUGH, B. F., POMMIER, Y., PRZYTYCKA, T. M., KOUZINE, F., LEWIS, B. A., ZHAO, K. & LEVENS, D. 2016. RNA Polymerase II Regulates Topoisomerase 1 Activity to Favor Efficient Transcription. *Cell*, 165, 357-71.
- BEEN, M. D., BURGESS, R. R. & CHAMPOUX, J. J. 1984. Nucleotide sequence preference at rat liver and wheat germ type 1 DNA topoisomerase breakage sites in duplex SV40 DNA. *Nucleic Acids Res*, 12, 3097-114.
- BELETSKII, A. & BHAGWAT, A. S. 1996. Transcription-induced mutations: increase in C to T mutations in the nontranscribed strand during transcription in *Escherichia coli*. *Proc Natl Acad Sci U S A*, 93, 13919-24.
- BELOTSERKOVSKII, B. P., LIU, R., TORNALETTI, S., KRASILNIKOVA, M. M., MIRKIN, S. M. & HANAWALT, P. C. 2010. Mechanisms and implications of transcription blockage by guanine-rich DNA sequences. *Proc Natl Acad Sci U S A*, 107, 12816-21.
- BELOTSERKOVSKII, B. P., NEIL, A. J., SALEH, S. S., SHIN, J. H., MIRKIN, S. M. & HANAWALT, P. C. 2013. Transcription blockage by homopurine DNA sequences: role of sequence composition and single-strand breaks. *Nucleic Acids Res*, 41, 1817-28.
- BELOTSERKOVSKII, B. P., SOO SHIN, J. H. & HANAWALT, P. C. 2017. Strong transcription blockage mediated by R-loop formation within a G-rich homopurine-homopyrimidine sequence localized in the vicinity of the promoter. *Nucleic Acids Res*, 45, 6589-6599.
- BERETTA, G. L., GATTI, L., PEREGO, P. & ZAFFARONI, N. 2013. Camptothecin resistance in cancer: insights into the molecular mechanisms of a DNA-damaging drug. *Curr Med Chem*, 20, 1541-65.

- BERGER, C. M., GAUME, X. & BOUVET, P. 2015. The roles of nucleolin subcellular localization in cancer. *Biochimie*, 113, 78-85.
- BERROYER, A., ALVARADO, G. & LARSON, E. D. 2019. Response of *Sulfolobus solfataricus* Dpo4 polymerase in vitro to a DNA G-quadruplex. *Mutagenesis*, 34, 289-297.
- BERROYER, A. & KIM, N. 2020. The Functional Consequences of Eukaryotic Topoisomerase 1 Interaction with G-Quadruplex DNA. *Genes (Basel)*, 11.
- BHARTI, A. K., OLSON, M. O., KUFÉ, D. W. & RUBIN, E. H. 1996. Identification of a nucleolin binding site in human topoisomerase I. *J Biol Chem*, 271, 1993-7.
- BIFFI, G., TANNAHILL, D., MCCAFFERTY, J. & BALASUBRAMANIAN, S. 2013. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat Chem*, 5, 182-6.
- BOCHMAN, M. L., PAESCHKE, K. & ZAKIAN, V. A. 2012. DNA secondary structures: stability and function of G-quadruplex structures. *Nat Rev Genet*, 13, 770-80.
- BREWER, B. J. & FANGMAN, W. L. 1988. A replication fork barrier at the 3' end of yeast ribosomal RNA genes. *Cell*, 55, 637-43.
- BROXSON, C., BECKETT, J. & TORNALETTI, S. 2011. Transcription arrest by a G quadruplex forming-trinucleotide repeat sequence from the human c-myc gene. *Biochemistry*, 50, 4162-72.
- BURGE, S., PARKINSON, G. N., HAZEL, P., TODD, A. K. & NEIDLE, S. 2006. Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res*, 34, 5402-15.
- BURGESS, D. J., DOLES, J., ZENDER, L., XUE, W., MA, B., MCCOMBIE, W. R., HANNON, G. J., LOWE, S. W. & HEMANN, M. T. 2008. Topoisomerase levels determine chemotherapy response in vitro and in vivo. *Proc Natl Acad Sci U S A*, 105, 9053-8.

- 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 Comput Biol*, 6, e1000861.
- CARVALHO, L. S., GONCALVES, N., FONSECA, N. A. & MOREIRA, J. N. 2021. Cancer Stem Cells and Nucleolin as Drivers of Carcinogenesis. *Pharmaceuticals (Basel)*, 14.
- CENTORE, R. C., YAZINSKI, S. A., TSE, A. & ZOU, L. 2012. Spartan/C1orf124, a reader of PCNA ubiquitylation and a regulator of UV-induced DNA damage response. *Mol Cell*, 46, 625-35.
- CESARE, A. J. & REDDEL, R. R. 2010. Alternative lengthening of telomeres: models, mechanisms and implications. *Nat Rev Genet*, 11, 319-30.
- 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. *Nat Biotechnol*, 33, 877-81.
- CHATTERJEE, N. & WALKER, G. C. 2017. Mechanisms of DNA damage, repair, and mutagenesis. *Environ Mol Mutagen*, 58, 235-263.
- CHEN, X., CUI, D., PAPUSHA, A., ZHANG, X., CHU, C. D., TANG, J., CHEN, K., PAN, X. & IRA, G. 2012. The Fun30 nucleosome remodeller promotes resection of DNA double-strand break ends. *Nature*, 489, 576-80.
- CHEN, X. L., SILVER, H. R., XIONG, L., BELICHENKO, I., ADEGITE, C. & JOHNSON, E. S. 2007. Topoisomerase I-dependent viability loss in *saccharomyces cerevisiae* mutants defective in both SUMO conjugation and DNA repair. *Genetics*, 177, 17-30.
- CHETTY, R. & SERRA, S. 2020. SMARCA family of genes. *J Clin Pathol*, 73, 257-260.

- CHILKA, P., DESAI, N. & DATTA, B. 2019. Small Molecule Fluorescent Probes for G-Quadruplex Visualization as Potential Cancer Theranostic Agents. *Molecules*, 24.
- CHO, J. E., KIM, N. & JINKS-ROBERTSON, S. 2015. Topoisomerase 1-dependent deletions initiated by incision at ribonucleotides are biased to the non-transcribed strand of a highly activated reporter. *Nucleic Acids Res*, 43, 9306-13.
- CHO, J. E., KIM, N., LI, Y. C. & JINKS-ROBERTSON, S. 2013. Two distinct mechanisms of Topoisomerase 1-dependent mutagenesis in yeast. *DNA Repair (Amst)*, 12, 205-11.
- CHRISTIANSON, T. W., SIKORSKI, R. S., DANTE, M., SHERO, J. H. & HIETER, P. 1992. Multifunctional yeast high-copy-number shuttle vectors. *Gene*, 110, 119-22.
- CHRISTMAN, M. F., DIETRICH, F. S. & FINK, G. R. 1988. Mitotic recombination in the rDNA of *S. cerevisiae* is suppressed by the combined action of DNA topoisomerases I and II. *Cell*, 55, 413-25.
- CORNELIO, D. A., SEDAM, H. N., FERRAREZI, J. A., SAMPAIO, N. M. & ARGUESO, J. L. 2017. Both R-loop removal and ribonucleotide excision repair activities of RNase H2 contribute substantially to chromosome stability. *DNA Repair (Amst)*, 52, 110-114.
- DABROWIAK, J. C., GOODISMAN, J. & WARD, B. 1997. Quantitative DNA footprinting. *Methods Mol Biol*, 90, 23-42.
- DAHAN, D., TSIRKAS, I., DOVRAT, D., SPARKS, M. A., SINGH, S. P., GALLETTO, R. & AHARONI, A. 2018. Pif1 is essential for efficient replisome progression through lagging strand G-quadruplex DNA secondary structures. *Nucleic Acids Res*, 46, 11847-11857.
- DAVIS, L. & MAIZELS, N. 2011. G4 DNA: at risk in the genome. *EMBO J*, 30, 3878-9.

- DE LA FUENTE, R., BAUMANN, C. & VIVEIROS, M. M. 2011. Role of ATRX in chromatin structure and function: implications for chromosome instability and human disease. *Reproduction*, 142, 221-34.
- DE MAGIS, A., GOTZ, S., HAJIKAZEMI, M., FEKETE-SZUCS, E., CATERINO, M., JURANEK, S. & PAESCHKE, K. 2020. Zuo1 supports G4 structure formation and directs repair toward nucleotide excision repair. *Nat Commun*, 11, 3907.
- DEJESUS-HERNANDEZ, M., MACKENZIE, I. R., BOEVE, B. F., BOXER, A. L., BAKER, M., RUTHERFORD, N. J., NICHOLSON, A. M., FINCH, N. A., FLYNN, H., ADAMSON, J., KOURI, N., WOJTAS, A., SENGDY, P., HSIUNG, G. Y., KARYDAS, A., SEELEY, W. W., JOSEPHS, K. A., COPPOLA, G., GESCHWIND, D. H., WSZOLEK, Z. K., FELDMAN, H., KNOPMAN, D. S., PETERSEN, R. C., MILLER, B. L., DICKSON, D. W., BOYLAN, K. B., GRAFF-RADFORD, N. R. & RADEMAKERS, R. 2011. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron*, 72, 245-56.
- DEMPSEY, L. A., SUN, H., HANAKAHI, L. A. & MAIZELS, N. 1999. G4 DNA binding by LR1 and its subunits, nucleolin and hnRNP D, A role for G-G pairing in immunoglobulin switch recombination. *J Biol Chem*, 274, 1066-71.
- DO, N. Q., LIM, K. W., TEO, M. H., HEDDI, B. & PHAN, A. T. 2011. Stacking of G-quadruplexes: NMR structure of a G-rich oligonucleotide with potential anti-HIV and anticancer activity. *Nucleic Acids Res*, 39, 9448-57.
- DONOHUE, J. 1956. Hydrogen-bonded helical configurations of polynucleotides. *Proc Natl Acad Sci U S A*, 42, 60-65.

- DROLET, M. 2006. Growth inhibition mediated by excess negative supercoiling: the interplay between transcription elongation, R-loop formation and DNA topology. *Mol Microbiol*, 59, 723-30.
- DROLET, M., BI, X. & LIU, L. F. 1994. Hypernegative supercoiling of the DNA template during transcription elongation in vitro. *J Biol Chem*, 269, 2068-74.
- DROLET, M., PHOENIX, P., MENZEL, R., MASSE, 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. *Proc Natl Acad Sci U S A*, 92, 3526-30.
- DROUIN, R., THERRIEN, J. P., ANGERS, M. & OUELLET, S. 2001. In vivo DNA analysis. *Methods Mol Biol*, 148, 175-219.
- DU, Z., ZHAO, Y. & LI, N. 2008. Genome-wide analysis reveals regulatory role of G4 DNA in gene transcription. *Genome Res*, 18, 233-41.
- DUCETT, J. K., PETERSON, F. C., HOOVER, L. A., PRUNUSKE, A. J., VOLKMAN, B. F. & CRAIG, E. A. 2013. Unfolding of the C-terminal domain of the J-protein Zuo1 releases autoinhibition and activates Pdr1-dependent transcription. *J Mol Biol*, 425, 19-31.
- 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 Dev*, 18, 1618-29.
- DUQUETTE, M. L., HUBER, M. D. & MAIZELS, N. 2007. G-rich proto-oncogenes are targeted for genomic instability in B-cell lymphomas. *Cancer Res*, 67, 2586-94.

- DUXIN, J. P., DEWAR, J. M., YARDIMCI, H. & WALTER, J. C. 2014. Repair of a DNA-protein crosslink by replication-coupled proteolysis. *Cell*, 159, 346-57.
- ECKELMANN, B. J., BACOLLA, A., WANG, H., YE, Z., GUERRERO, E. N., JIANG, W., EL-ZEIN, R., HEGDE, M. L., TOMKINSON, A. E., TAINER, J. A., & MITRA, S. 2020. XRCC1 promotes replication restart, nascent fork degradation and mutagenic DNA repair in BRCA2-deficient cells. *NAR Cancer*, 2, zcaa013.
- EDDY, J. & MAIZELS, N. 2008. Conserved elements with potential to form polymorphic G-quadruplex structures in the first intron of human genes. *Nucleic Acids Res*, 36, 1321-33.
- EDDY, S., KETKAR, A., ZAFAR, M. K., MADDUKURI, L., CHOI, J. Y. & EOFF, R. L. 2014. Human Rev1 polymerase disrupts G-quadruplex DNA. *Nucleic Acids Res*, 42, 3272-85.
- EDDY, S., MADDUKURI, L., KETKAR, A., ZAFAR, M. K., HENNINGER, E. E., PURSELL, Z. F. & EOFF, R. L. 2015. Evidence for the kinetic partitioning of polymerase activity on G-quadruplex DNA. *Biochemistry*, 54, 3218-30.
- EDDY, S., TILLMAN, M., MADDUKURI, L., KETKAR, A., ZAFAR, M. K. & EOFF, R. L. 2016. Human Translesion Polymerase kappa Exhibits Enhanced Activity and Reduced Fidelity Two Nucleotides from G-Quadruplex DNA. *Biochemistry*, 55, 5218-29.
- EDWARDS, T. K., SALEEM, A., SHAMAN, J. A., DENNIS, T., GERIGK, C., OLIVEROS, E., GARTENBERG, M. R. & RUBIN, E. H. 2000. Role for nucleolin/Nsr1 in the cellular localization of topoisomerase I. *J Biol Chem*, 275, 36181-8.



- 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 Dev*, 24, 1546-58.
- EL HAGE, A., WEBB, S., KERR, A. & TOLLERVEY, D. 2014. Genome-wide distribution of RNA-DNA hybrids identifies RNase H targets in tRNA genes, retrotransposons and mitochondria. *PLoS Genet*, 10, e1004716.
- ELSASSER, S. J., NOH, K. M., DIAZ, N., ALLIS, C. D. & BANASZYNSKI, L. A. 2015. Histone H3.3 is required for endogenous retroviral element silencing in embryonic stem cells. *Nature*, 522, 240-244.
- FASCHING, C. L., CEJKA, P., KOWALCZYKOWSKI, S. C. & HEYER, W. D. 2015. Top3-Rmi1 dissolve Rad51-mediated D loops by a topoisomerase-based mechanism. *Mol Cell*, 57, 595-606.
- FLAUS, A., MARTIN, D. M., BARTON, G. J. & OWEN-HUGHES, T. 2006. Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. *Nucleic Acids Res*, 34, 2887-905.
- FRATTA, P., MIZIELINSKA, S., NICOLL, A. J., ZLOH, M., FISHER, E. M., PARKINSON, G. & ISAACS, A. M. 2012. C9orf72 hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms RNA G-quadruplexes. *Sci Rep*, 2, 1016.
- FRENCH, S. L., SIKES, M. L., HONTZ, R. D., OSHEIM, Y. N., LAMBERT, T. E., EL HAGE, A., SMITH, M. M., TOLLERVEY, D., SMITH, J. S. & BEYER, A. L. 2011. Distinguishing the

- roles of Topoisomerases I and II in relief of transcription-induced torsional stress in yeast rRNA genes. *Mol Cell Biol*, 31, 482-94.
- FRY, M. 2007. Tetraplex DNA and its interacting proteins. *Front Biosci*, 12, 4336-51.
- FUJIKI, H., WATANABE, T. & SUGANUMA, M. 2014. Cell-surface nucleolin acts as a central mediator for carcinogenic, anti-carcinogenic, and disease-related ligands. *J Cancer Res Clin Oncol*, 140, 689-99.
- GADALETA, M. C. & NOGUCHI, E. 2017. Regulation of DNA Replication through Natural Impediments in the Eukaryotic Genome. *Genes (Basel)*, 8.
- GANGLOFF, S., DE MASSY, B., ARTHUR, L., ROTHSTEIN, R. & FABRE, F. 1999. The essential role of yeast topoisomerase III in meiosis depends on recombination. *EMBO J*, 18, 1701-11.
- GAO, J., ZYBAILOV, B. L., BYRD, A. K., GRIFFIN, W. C., CHIB, S., MACKINTOSH, S. G., TACKETT, A. J. & RANEY, K. D. 2015. Yeast transcription co-activator Sub1 and its human homolog PC4 preferentially bind to G-quadruplex DNA. *Chem Commun (Camb)*, 51, 7242-4.
- GELLERT, M., LIPSETT, M. N., & DAVIES, D.R. 1962. Helix formation by guanylic acid. *Proceedings of the National Academy of Sciences of the United States of America*, 48, 2013–2018.
- GERMAN, J. 1993. Bloom syndrome: a mendelian prototype of somatic mutational disease. *Medicine (Baltimore)*, 72, 393-406.

- GHOSH, M. & SINGH, M. 2018. RGG-box in hnRNPA1 specifically recognizes the telomere G-quadruplex DNA and enhances the G-quadruplex unfolding ability of UP1 domain. *Nucleic Acids Res*, 46, 10246-10261.
- GIBBONS, R. J., MCDOWELL, T. L., RAMAN, S., O'ROURKE, D. M., GARRICK, D., AYYUB, H. & HIGGS, D. R. 2000. Mutations in ATRX, encoding a SWI/SNF-like protein, cause diverse changes in the pattern of DNA methylation. *Nat Genet*, 24, 368-71.
- GINISTY, H., SICARD, H., ROGER, B. & BOUVET, P. 1999. Structure and functions of nucleolin. *J Cell Sci*, 112 ( Pt 6), 761-72.
- GOLDSTEIN, M., DERHEIMER, F. A., TAIT-MULDER, J. & KASTAN, M. B. 2013. Nucleolin mediates nucleosome disruption critical for DNA double-strand break repair. *Proc Natl Acad Sci U S A*, 110, 16874-9.
- GOMEZ-GONZALEZ, B. & AGUILERA, A. 2019. Transcription-mediated replication hindrance: a major driver of genome instability. *Genes Dev*, 33, 1008-1026.
- GONZALEZ, V., GUO, K., HURLEY, L. & SUN, D. 2009. Identification and characterization of nucleolin as a c-myc G-quadruplex-binding protein. *J Biol Chem*, 284, 23622-35.
- GONZALEZ, V. & HURLEY, L. H. 2010. The C-terminus of nucleolin promotes the formation of the c-MYC G-quadruplex and inhibits c-MYC promoter activity. *Biochemistry*, 49, 9706-14.
- GOTO, T. & WANG, J. C. 1985. Cloning of yeast TOP1, the gene encoding DNA topoisomerase I, and construction of mutants defective in both DNA topoisomerase I and DNA topoisomerase II. *Proc. Natl. Acad. Sci. U. S. A.* , 82, 7178-82.

- GRIFFIN, W. C., GAO, J., BYRD, A. K., CHIB, S. & RANEY, K. D. 2017. A biochemical and biophysical model of G-quadruplex DNA recognition by positive coactivator of transcription 4. *J Biol Chem*, 292, 9567-9582.
- HAEUSLER, A. R., DONNELLY, C. J., PERIZ, G., SIMKO, E. A., SHAW, P. G., KIM, M. S., MARAGAKIS, N. J., TRONCOSO, J. C., PANDEY, A., SATTLER, R., ROTHSTEIN, J. D. & WANG, J. 2014. C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature*, 507, 195-200.
- HALUSKA, P., JR. & RUBIN, E. H. 1998. A role for the amino terminus of human topoisomerase I. *Adv Enzyme Regul*, 38, 253-62.
- HAMMOND, C. M., STROMME, C. B., HUANG, H., PATEL, D. J. & GROTH, A. 2017. Histone chaperone networks shaping chromatin function. *Nat Rev Mol Cell Biol*, 18, 141-158.
- HAMPERL, S. & CIMPRICH, K. A. 2014. The contribution of co-transcriptional RNA:DNA hybrid structures to DNA damage and genome instability. *DNA Repair (Amst)*, 19, 84-94.
- HAN, H., HURLEY, L. H. & SALAZAR, M. 1999. A DNA polymerase stop assay for G-quadruplex-interactive compounds. *Nucleic Acids Res*, 27, 537-42.
- HANAKAHI, L. A., SUN, H. & MAIZELS, N. 1999. High affinity interactions of nucleolin with G-G-paired rDNA. *J Biol Chem*, 274, 15908-12.
- HANN, C. L., CARLBERG, A. L. & BJORNSTI, M. A. 1998. Intragenic suppressors of mutant DNA topoisomerase I-induced lethality diminish enzyme binding of DNA. *J Biol Chem*, 273, 31519-27.

- HANSEL-HERTSCH, R., BERALDI, D., LENSING, S. V., MARSICO, G., ZYNER, K., PARRY, A., DI ANTONIO, M., PIKE, J., KIMURA, H., NARITA, M., TANNAHILL, D. & BALASUBRAMANIAN, S. 2016. G-quadruplex structures mark human regulatory chromatin. *Nat Genet*, 48, 1267-72.
- HANSEL-HERTSCH, R., DI ANTONIO, M. & BALASUBRAMANIAN, S. 2017. DNA G-quadruplexes in the human genome: detection, functions and therapeutic potential. *Nat Rev Mol Cell Biol*, 18, 279-284.
- HEAPHY, C. M., DE WILDE, R. F., JIAO, Y., KLEIN, A. P., EDIL, B. H., SHI, C., BETTEGOWDA, C., RODRIGUEZ, F. J., EBERHART, C. G., HEBBAR, S., OFFERHAUS, G. J., MCLENDON, R., RASHEED, B. A., HE, Y., YAN, H., BIGNER, D. D., OBA-SHINJO, S. M., MARIE, S. K., RIGGINS, G. J., KINZLER, K. W., VOGELSTEIN, B., HRUBAN, R. H., MAITRA, A., PAPADOPOULOS, N. & MEEKER, A. K. 2011. Altered telomeres in tumors with ATRX and DAXX mutations. *Science*, 333, 425.
- HEDDI, B. & PHAN, A. T. 2011. Structure of human telomeric DNA in crowded solution. *J Am Chem Soc*, 133, 9824-33.
- HEIDEKER, J., PRUDDEN, J., PERRY, J. J., TAINER, J. A. & BODDY, M. N. 2011. SUMO-targeted ubiquitin ligase, Rad60, and Nse2 SUMO ligase suppress spontaneous Top1-mediated DNA damage and genome instability. *PLoS Genet*, 7, e1001320.
- HILL, T. M., HENSON, J. M. & KUEMPEL, P. L. 1987. The terminus region of the Escherichia coli chromosome contains two separate loci that exhibit polar inhibition of replication. *Proc Natl Acad Sci U S A*, 84, 1754-8.

- HILL, T. M., TECKLENBURG, M. L., PELLETIER, A. J. & KUEMPEL, P. L. 1989. *tus*, the trans-acting gene required for termination of DNA replication in *Escherichia coli*, encodes a DNA-binding protein. *Proc Natl Acad Sci U S A*, 86, 1593-7.
- HIZUME, K. & ARAKI, H. 2019. Replication fork pausing at protein barriers on chromosomes. *FEBS Lett*, 593, 1449-1458.
- HORIE, K., TOMIDA, A., SUGIMOTO, Y., YASUGI, T., YOSHIKAWA, H., TAKETANI, Y. & TSURUO, T. 2002. SUMO-1 conjugation to intact DNA topoisomerase I amplifies cleavable complex formation induced by camptothecin. *Oncogene*, 21, 7913-22.
- HOTHORN, M., NEUMANN, H., LENHERR, E. D., WEHNER, M., RYBIN, V., HASSA, P. O., UTTENWEILER, A., REINHARDT, M., SCHMIDT, A., SEILER, J., LADURNER, A. G., HERRMANN, C., SCHEFFZEK, K. & MAYER, A. 2009. Catalytic core of a membrane-associated eukaryotic polyphosphate polymerase. *Science*, 324, 513-6.
- HUANG, F., WU, Y., TAN, H., GUO, T., ZHANG, K., LI, D. & TONG, Z. 2019. Phosphorylation of nucleolin is indispensable to its involvement in the proliferation and migration of non-small cell lung cancer cells. *Oncol Rep*, 41, 590-598.
- HUANG, F. C., CHANG, C. C., WANG, J. M., CHANG, T. C. & LIN, J. J. 2012. Induction of senescence in cancer cells by the G-quadruplex stabilizer, BMVC4, is independent of its telomerase inhibitory activity. *Br J Pharmacol*, 167, 393-406.
- HUANG, S. N., WILLIAMS, J. S., ARANA, M. E., KUNKEL, T. A. & POMMIER, Y. 2017. Topoisomerase I-mediated cleavage at unrepaired ribonucleotides generates DNA double-strand breaks. *EMBO J*, 36, 361-373.

- HUBERT, L., JR., LIN, Y., DION, V. & WILSON, J. H. 2011. Topoisomerase 1 and single-strand break repair modulate transcription-induced CAG repeat contraction in human cells. *Mol Cell Biol*, 31, 3105-12.
- HUPPERT, J. L. & BALASUBRAMANIAN, S. 2005. Prevalence of quadruplexes in the human genome. *Nucleic Acids Res*, 33, 2908-16.
- HUSAIN, A., BEGUM, N. A., TANIGUCHI, T., TANIGUCHI, H., KOBAYASHI, M. & HONJO, T. 2016. Chromatin remodeller SMARCA4 recruits topoisomerase 1 and suppresses transcription-associated genomic instability. *Nat Commun*, 7, 10549.
- HUSTON, J. S., LEVINSON, D., MUDGETT-HUNTER, M., TAI, M. S., NOVOTNY, J., MARGOLIES, M. N., RIDGE, R. J., BRUCCOLERI, R. E., HABER, E., CREA, R. & ET AL. 1988. Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in *Escherichia coli*. *Proc Natl Acad Sci U S A*, 85, 5879-83.
- INDIG, F. E., RYBANSKA, I., KARMAKAR, P., DEVULAPALLI, C., FU, H., CARRIER, F. & BOHR, V. A. 2012. Nucleolin inhibits G4 oligonucleotide unwinding by Werner helicase. *PLoS One*, 7, e35229.
- IROBALIEVA, R. N., FOGG, J. M., CATANESE, D. J., JR., SUTTHIBUTPONG, T., CHEN, M., BARKER, A. K., LUDTKE, S. J., HARRIS, S. A., SCHMID, M. F., CHIU, W. & ZECHIEDRICH, L. 2015. Structural diversity of supercoiled DNA. *Nat Commun*, 6, 8440.
- ITO, H., FUKUDA, Y., MURATA, K. & KIMURA, A. 1983. Transformation of intact yeast cells treated with alkali cations. *J Bacteriol*, 153, 163-8.

- IVESSA, A. S., LENZMEIER, B. A., BESSLER, J. B., GOUDSOUZIAN, L. K., SCHNAKENBERG, S. L. & ZAKIAN, V. A. 2003. The *Saccharomyces cerevisiae* helicase Rrm3p facilitates replication past nonhistone protein-DNA complexes. *Mol Cell*, 12, 1525-36.
- JAIN, A., WANG, G. & VASQUEZ, K. M. 2008. DNA triple helices: biological consequences and therapeutic potential. *Biochimie*, 90, 1117-30.
- JALAL, D., CHALISSERY, J. & HASSAN, A. H. 2017. Genome maintenance in *Saccharomyces cerevisiae*: the role of SUMO and SUMO-targeted ubiquitin ligases. *Nucleic Acids Res*, 45, 2242-2261.
- JANA, J., MOHR, S., VIANNEY, Y. M. & WEISZ, K. 2021. Structural motifs and intramolecular interactions in non-canonical G-quadruplexes. *RSC Chem Biol*, 2, 338-353.
- JUHASZ, S., BALOGH, D., HAJDU, I., BURKOVICS, P., VILLAMIL, M. A., ZHUANG, Z. & HARACSKA, L. 2012. Characterization of human Spartan/C1orf124, an ubiquitin-PCNA interacting regulator of DNA damage tolerance. *Nucleic Acids Res*, 40, 10795-808.
- KALDERON, D., ROBERTS, B. L., RICHARDSON, W. D. & SMITH, A. E. 1984. A short amino acid sequence able to specify nuclear location. *Cell*, 39, 499-509.
- KAMADA, K., HORIUCHI, T., OHSUMI, K., SHIMAMOTO, N. & MORIKAWA, K. 1996. Structure of a replication-terminator protein complexed with DNA. *Nature*, 383, 598-603.
- KASAMATSU, H., ROBBERSON, D. L. & VINOGRAD, J. 1971. A novel closed-circular mitochondrial DNA with properties of a replicating intermediate. *Proc Natl Acad Sci U S A*, 68, 2252-7.



- KATAPADI, V. K., NAMBIAR, M. & RAGHAVAN, S. C. 2012. Potential G-quadruplex formation at breakpoint regions of chromosomal translocations in cancer may explain their fragility. *Genomics*, 100, 72-80.
- KIM, N. 2019. The Interplay between G-quadruplex and Transcription. *Curr Med Chem*, 26, 2898-2917.
- KIM, N., HUANG, S. N., WILLIAMS, J. S., LI, Y. C., CLARK, A. B., CHO, J. E., KUNKEL, T. A., POMMIER, Y. & JINKS-ROBERTSON, S. 2011. Mutagenic processing of ribonucleotides in DNA by yeast topoisomerase I. *Science*, 332, 1561-4.
- KIM, N. & JINKS-ROBERTSON, S. 2011. Guanine repeat-containing sequences confer transcription-dependent instability in an orientation-specific manner in yeast. *DNA Repair (Amst)*, 10, 953-60.
- KIM, N. & JINKS-ROBERTSON, S. 2012. Transcription as a source of genome instability. *Nat Rev Genet*, 13, 204-14.
- KIM, N. & JINKS-ROBERTSON, S. 2017. The Top1 paradox: Friend and foe of the eukaryotic genome. *DNA Repair (Amst)*, 56, 33-41.
- KIM, R. A. & WANG, J. C. 1992. Identification of the yeast TOP3 gene product as a single strand-specific DNA topoisomerase. *J Biol Chem*, 267, 17178-85.
- KIM, S. Y., SHEN, Q., SON, K., KIM, H. S., YANG, H. D., NA, M. J., SHIN, E., YU, S., KANG, K., YOU, J. S., YU, K. R., JEONG, S. M., LEE, E. K., AHN, Y. M., PARK, W. S. & NAM, S. W. 2021. SMARCA4 oncogenic potential via IRAK1 enhancer to activate Gankyrin and AKR1B10 in liver cancer. *Oncogene*, 40, 4652-4662.

- KLEIN, H. L. 2020. Stressed DNA replication generates stressed DNA. *Proc Natl Acad Sci U S A*, 117, 10108-10110.
- KOBAYASHI, M., AIDA, M., NAGAOKA, H., BEGUM, N. A., KITAWAKI, Y., NAKATA, M., STANLIE, A., DOI, T., KATO, L., OKAZAKI, I. M., SHINKURA, R., MURAMATSU, M., KINOSHITA, K. & HONJO, T. 2009. AID-induced decrease in topoisomerase 1 induces DNA structural alteration and DNA cleavage for class switch recombination. *Proc Natl Acad Sci U S A*, 106, 22375-80.
- KOBAYASHI, M., SABOURI, Z., SABOURI, S., KITAWAKI, Y., POMMIER, Y., ABE, T., KIYONARI, H. & HONJO, T. 2011. Decrease in topoisomerase I is responsible for activation-induced cytidine deaminase (AID)-dependent somatic hypermutation. *Proc Natl Acad Sci U S A*, 108, 19305-10.
- KOBAYASHI, T. & HORIUCHI, T. 1996. A yeast gene product, Fob1 protein, required for both replication fork blocking and recombinational hotspot activities. *Genes Cells*, 1, 465-74.
- KONDO, K. & INOUE, M. 1992. Yeast NSR1 protein that has structural similarity to mammalian nucleolin is involved in pre-rRNA processing. *J Biol Chem*, 267, 16252-8.
- KONIG, S. L., EVANS, A. C. & HUPPERT, J. L. 2010. Seven essential questions on G-quadruplexes. *Biomol Concepts*, 1, 197-213.
- KOUTSIOUMPA, M. & PAPADIMITRIOU, E. 2014. Cell surface nucleolin as a target for anti-cancer therapies. *Recent Pat Anticancer Drug Discov*, 9, 137-52.

- KUBOTA, N., KANZAWA, F., NISHIO, K., TAKEDA, Y., OHMORI, T., FUJIWARA, Y., TERASHIMA, Y. & SAIJO, N. 1992. Detection of topoisomerase I gene point mutation in CPT-11 resistant lung cancer cell line. *Biochem Biophys Res Commun*, 188, 571-7.
- KUSHNIROV, V. V. 2000. Rapid and reliable protein extraction from yeast. *Yeast*, 16, 857-60.
- KYPR, J., KEJNOVSKA, I., RENCIUUK, D. & VORLICKOVA, M. 2009. Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Res*, 37, 1713-25.
- LAFUENTE-BARQUERO, J., LUKE-GLASER, S., GRAF, M., SILVA, S., GOMEZ-GONZALEZ, B., LOCKHART, A., LISBY, M., AGUILERA, A. & LUKE, B. 2017. The Smc5/6 complex regulates the yeast Mph1 helicase at RNA-DNA hybrid-mediated DNA damage. *PLoS Genet*, 13, e1007136.
- LAGUERRE, A., STEFAN, L., LARROUY, M., GENEST, D., NOVOTNA, J., PIRROTTA, M. & MONCHAUD, D. 2014. A twice-as-smart synthetic G-quartet: PyroTASQ is both a smart quadruplex ligand and a smart fluorescent probe. *J Am Chem Soc*, 136, 12406-14.
- LAGUERRE, A., WONG, J. M. & MONCHAUD, D. 2016. Direct visualization of both DNA and RNA quadruplexes in human cells via an uncommon spectroscopic method. *Sci Rep*, 6, 32141.
- LAW, M. J., LOWER, K. M., VOON, H. P., HUGHES, J. R., GARRICK, D., VIPRAKASIT, V., MITSON, M., DE GOBBI, M., MARRA, M., MORRIS, A., ABBOTT, A., WILDER, S. P., TAYLOR, S., SANTOS, G. M., CROSS, J., AYYUB, H., JONES, S., RAGOUSSIS, J., RHODES, D., DUNHAM, I., HIGGS, D. R. & GIBBONS, R. J. 2010. ATR-X syndrome protein targets

- tandem repeats and influences allele-specific expression in a size-dependent manner. *Cell*, 143, 367-78.
- LEBEL, M., SPILLARE, E. A., HARRIS, C. C. & LEDER, P. 1999. The Werner syndrome gene product co-purifies with the DNA replication complex and interacts with PCNA and topoisomerase I. *J Biol Chem*, 274, 37795-9.
- LEE, C. Y., MCNERNEY, C., MA, K., ZHAO, W., WANG, A. & MYONG, S. 2020. R-loop induced G-quadruplex in non-template promotes transcription by successive R-loop formation. *Nat Commun*, 11, 3392.
- LERNER, L. K. & SALE, J. E. 2019. Replication of G Quadruplex DNA. *Genes (Basel)*, 10.
- LEWIS, P. W., ELSAESSER, S. J., NOH, K. M., STADLER, S. C. & ALLIS, C. D. 2010. Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres. *Proc Natl Acad Sci U S A*, 107, 14075-80.
- LI, E., BESTOR, T. H. & JAENISCH, R. 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell*, 69, 915-26.
- LI, M., CHEN, W., SUN, X., WANG, Z., ZOU, X., WEI, H., WANG, Z. & CHEN, W. 2019. Metastatic colorectal cancer and severe hypocalcemia following irinotecan administration in a patient with X-linked agammaglobulinemia: a case report. *BMC Med Genet*, 20, 157.
- LIAO, J., KARNIK, R., GU, H., ZILLER, M. J., CLEMENT, K., TSANKOV, A. M., AKOPIAN, V., GIFFORD, C. A., DONAGHEY, J., GALONSKA, C., POP, R., REYON, D., TSAI, S. Q., MALLARD, W., JOUNG, J. K., RINN, J. L., GNIRKE, A. & MEISSNER, A. 2015. Targeted

- disruption of DNMT1, DNMT3A and DNMT3B in human embryonic stem cells. *Nat Genet*, 47, 469-78.
- LIN, C. Y., LOVEN, J., RAHL, P. B., PARANAL, R. M., BURGE, C. B., BRADNER, J. E., LEE, T. I. & YOUNG, R. A. 2012. Transcriptional amplification in tumor cells with elevated c-Myc. *Cell*, 151, 56-67.
- LINDAHL, T. 1993. Instability and decay of the primary structure of DNA. *Nature*, 362, 709-15.
- LOMEN-HOERTH, C., ANDERSON, T. & MILLER, B. 2002. The overlap of amyotrophic lateral sclerosis and frontotemporal dementia. *Neurology*, 59, 1077-9.
- LOPES, J., PIAZZA, A., BERMEJO, R., KRIEGSMAN, B., COLOSIO, A., TEULADE-FICHO, M. P., FOIANI, M. & NICOLAS, A. 2011. G-quadruplex-induced instability during leading-strand replication. *EMBO J*, 30, 4033-46.
- LOPEZ, C. R., SINGH, S., HAMBARDE, S., GRIFFIN, W. C., GAO, J., CHIB, S., YU, Y., IRA, G., RANEY, K. D. & KIM, N. 2017. Yeast Sub1 and human PC4 are G-quadruplex binding proteins that suppress genome instability at co-transcriptionally formed G4 DNA. *Nucleic Acids Res*, 45, 5850-5862.
- LOSASSO, C., CRETAIO, E., FIORANI, P., D'ANNESSA, I., CHILLEMI, G. & BENEDETTI, P. 2008. A single mutation in the 729 residue modulates human DNA topoisomerase IB DNA binding and drug resistance. *Nucleic Acids Res*, 36, 5635-44.
- LOTITO, L., RUSSO, A., CHILLEMI, G., BUENO, S., CAVALIERI, D. & CAPRANICO, G. 2008. Global transcription regulation by DNA topoisomerase I in exponentially growing

- Saccharomyces cerevisiae cells: activation of telomere-proximal genes by TOP1 deletion. *J Mol Biol*, 377, 311-22.
- MA, J. & WANG, M. 2014. Interplay between DNA supercoiling and transcription elongation. *Transcription*, 5, e28636.
- MA, J. & WANG, M. D. 2016. DNA supercoiling during transcription. *Biophys Rev*, 8, 75-87.
- MADDEN, K. R., STEWART, L. & CHAMPOUX, J. J. 1995. Preferential binding of human topoisomerase I to superhelical DNA. *EMBO J*, 14, 5399-409.
- MADDI, K., SAM, D. K., BONN, F., PRGOMET, S., TULOWETZKE, E., AKUTSU, M., LOPEZ-MOSQUEDA, J. & DIKIC, I. 2020. Wss1 Promotes Replication Stress Tolerance by Degrading Histones. *Cell Rep*, 30, 3117-3126 e4.
- MALKOVA, A., & IRA, G. 2013. Break-induced replication: functions and molecular mechanism. *Curr Opin Genet Dev*, 23, 271-279.
- MANKOURI, H. W., & MORGAN, A. 2001. The DNA helicase activity of yeast Sgs1p is essential for normal lifespan but not for resistance to topoisomerase inhibitors. *Mechanisms of Ageing and Development*, 122, 1107-1120.
- MANZO, S. G., HARTONO, S. R., SANZ, L. A., MARINELLO, J., DE BIASI, S., COSSARIZZA, A., CAPRANICO, G. & CHEDIN, F. 2018. DNA Topoisomerase I differentially modulates R-loops across the human genome. *Genome Biol*, 19, 100.
- MAO, S. Q., GHANBARIAN, A. T., SPIEGEL, J., MARTINEZ CUESTA, S., BERALDI, D., DI ANTONIO, M., MARSICO, G., HANSEL-HERTSCH, R., TANNAHILL, D. & BALASUBRAMANIAN, S. 2018. DNA G-quadruplex structures mold the DNA methylome. *Nat Struct Mol Biol*, 25, 951-957.

- MAO, Y., SUN, M., DESAI, S. D. & LIU, L. F. 2000. SUMO-1 conjugation to topoisomerase I: A possible repair response to topoisomerase-mediated DNA damage. *Proc Natl Acad Sci U S A*, 97, 4046-51.
- MARCHAND, C., POURQUIER, P., LACO, G. S., JING, N. & POMMIER, Y. 2002. Interaction of human nuclear topoisomerase I with guanosine quartet-forming and guanosine-rich single-stranded DNA and RNA oligonucleotides. *J Biol Chem*, 277, 8906-11.
- MARSICO, G., CHAMBERS, V. S., SAHAKYAN, A. B., MCCAULEY, P., BOUTELL, J. M., ANTONIO, M. D. & BALASUBRAMANIAN, S. 2019. Whole genome experimental maps of DNA G-quadruplexes in multiple species. *Nucleic Acids Res*, 47, 3862-3874.
- MASKEY, R. S., FLATTEN, K. S., SIEBEN, C. J., PETERSON, K. L., BAKER, D. J., NAM, H. J., KIM, M. S., SMYRK, T. C., KOJIMA, Y., MACHIDA, Y., SANTIAGO, A., VAN DEURSEN, J. M., KAUFMANN, S. H. & MACHIDA, Y. J. 2017. Spartan deficiency causes accumulation of Topoisomerase 1 cleavage complexes and tumorigenesis. *Nucleic Acids Res*, 45, 4564-4576.
- MASUZAWA, T. & OYOSHI, T. 2020. Roles of the RGG Domain and RNA Recognition Motif of Nucleolin in G-Quadruplex Stabilization. *ACS Omega*, 5, 5202-5208.
- MATOS-PERDOMO, E. & MACHIN, F. 2019. Nucleolar and Ribosomal DNA Structure under Stress: Yeast Lessons for Aging and Cancer. *Cells*, 8.
- MATTAJ, I. W. & ENGLMEIER, L. 1998. Nucleocytoplasmic transport: the soluble phase. *Annu Rev Biochem*, 67, 265-306.
- MCDOWELL, T. L., GIBBONS, R. J., SUTHERLAND, H., O'ROURKE, D. M., BICKMORE, W. A., POMBO, A., TURLEY, H., GATTER, K., PICKETTS, D. J., BUCKLE, V. J., CHAPMAN, L.,

- RHODES, D. & HIGGS, D. R. 1999. Localization of a putative transcriptional regulator (ATRX) at pericentromeric heterochromatin and the short arms of acrocentric chromosomes. *Proc Natl Acad Sci U S A*, 96, 13983-8.
- MEGONIGAL, M. D., FERTALA, J. & BJORNSTI, M. A. 1997. Alterations in the catalytic activity of yeast DNA topoisomerase I result in cell cycle arrest and cell death. *J Biol Chem*, 272, 12801-8.
- MENDOZA, O., BOURDONCLE, A., BOULE, J. B., BROSH, R. M., JR. & MERGNY, J. L. 2016. G-quadruplexes and helicases. *Nucleic Acids Res*, 44, 1989-2006.
- MERROUCHE, Y., MUGNERET, F. & CAHN, J. Y. 2006. Secondary acute promyelocytic leukemia following irinotecan and oxaliplatin for advanced colon cancer. *Ann Oncol*, 17, 1025-6.
- METIFIOT, M., AMRANE, S., LITVAK, S. & ANDREOLA, M. L. 2014. G-quadruplexes in viruses: function and potential therapeutic applications. *Nucleic Acids Res*, 42, 12352-66.
- MIRKIN, S. M. 2008. Discovery of alternative DNA structures: a heroic decade (1979-1989). *Front Biosci*, 13, 1064-71.
- MOHANTY, B. K., SAHOO, T. & BASTIA, D. 1996. The relationship between sequence-specific termination of DNA replication and transcription. *EMBO J*, 15, 2530-9.
- MORAWSKA, M. & ULRICH, H. D. 2013. An expanded tool kit for the auxin-inducible degron system in budding yeast. *Yeast*, 30, 341-51.
- MORRISON, A. J., HIGHLAND, J., KROGAN, N. J., ARBEL-EDEN, A., GREENBLATT, J. F., HABER, J. E. & SHEN, X. 2004. INO80 and gamma-H2AX interaction links ATP-dependent chromatin remodeling to DNA damage repair. *Cell*, 119, 767-75.



- MORUNO-MANCHON, J. F., KOELLHOFFER, E. C., GOPAKUMAR, J., HAMBARDE, S., KIM, N., MCCULLOUGH, L. D. & TSVETKOV, A. S. 2017. The G-quadruplex DNA stabilizing drug pyridostatin promotes DNA damage and downregulates transcription of Brca1 in neurons. *Aging (Albany NY)*, 9, 1957-1970.
- MULLEN, J. R., CHEN, C. F. & BRILL, S. J. 2010. Wss1 is a SUMO-dependent isopeptidase that interacts genetically with the Slx5-Slx8 SUMO-targeted ubiquitin ligase. *Mol Cell Biol*, 30, 3737-48.
- NAKATANI, R., NAKAMORI, M., FUJIMURA, H., MOCHIZUKI, H. & TAKAHASHI, M. P. 2015. Large expansion of CTG\*CAG repeats is exacerbated by MutSbeta in human cells. *Sci Rep*, 5, 11020.
- NAMBIAR, M., GOLDSMITH, G., MOORTHY, B. T., LIEBER, M. R., JOSHI, M. V., CHOUDHARY, B., HOSUR, R. V. & RAGHAVAN, S. C. 2011. Formation of a G-quadruplex at the BCL2 major breakpoint region of the t(14;18) translocation in follicular lymphoma. *Nucleic Acids Res*, 39, 936-48.
- NAMBIAR, M., SRIVASTAVA, M., GOPALAKRISHNAN, V., SANKARAN, S. K. & RAGHAVAN, S. C. 2013. G-quadruplex structures formed at the HOX11 breakpoint region contribute to its fragility during t(10;14) translocation in T-cell leukemia. *Mol Cell Biol*, 33, 4266-81.
- NAPIERALA, M., BACOLLA, A. & WELLS, R. D. 2005. Increased negative superhelical density in vivo enhances the genetic instability of triplet repeat sequences. *J Biol Chem*, 280, 37366-76.

- NEVES-COSTA, A., WILL, W. R., VETTER, A. T., MILLER, J. R. & VARGA-WEISZ, P. 2009. The SNF2-family member Fun30 promotes gene silencing in heterochromatic loci. *PLoS One*, 4, e8111.
- NGUYEN, D. T., VOON, H. P. J., XELLA, B., SCOTT, C., CLYNES, D., BABBS, C., AYYUB, H., KERRY, J., SHARPE, J. A., SLOANE-STANLEY, J. A., BUTLER, S., FISHER, C. A., GRAY, N. E., JENUWEIN, T., HIGGS, D. R. & GIBBONS, R. J. 2017. The chromatin remodelling factor ATRX suppresses R-loops in transcribed telomeric repeats. *EMBO Rep*, 18, 914-928.
- NORDHEIM, A., LAFER, E. M., PECK, L. J., WANG, J. C., STOLLAR, B. D. & RICH, A. 1982. Negatively supercoiled plasmids contain left-handed Z-DNA segments as detected by specific antibody binding. *Cell*, 31, 309-18.
- OGLOBLINA, A. M., BANNIKOVA, V. A., KHRISTICH, A. N., ORETSKAYA, T. S., YAKUBOVSKAYA, M. G. & DOLINNAYA, N. G. 2015. Parallel G-Quadruplexes Formed by Guanine-Rich Microsatellite Repeats Inhibit Human Topoisomerase I. *Biochemistry (Mosc)*, 80, 1026-38.
- OGLOBLINA, A. M., KHRISTICH, A. N., KARPECHENKO, N. Y., SEMINA, S. E., BELITSKY, G. A., DOLINNAYA, N. G. & YAKUBOVSKAYA, M. G. 2018. Multi-targeted effects of G4-aptamers and their antiproliferative activity against cancer cells. *Biochimie*, 145, 163-173.
- OHKUNI, K., TAKAHASHI, Y. & BASRAI, M. A. 2015. Protein purification technique that allows detection of sumoylation and ubiquitination of budding yeast kinetochore proteins Ndc10 and Ndc80. *J Vis Exp*, e52482.

- OKANO, M., BELL, D. W., HABER, D. A. & LI, E. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell*, 99, 247-57.
- OUSSATCHEVA, E. A., PAVLICEK, J., SANKEY, O. F., SINDEN, R. R., LYUBCHENKO, Y. L. & POTAMAN, V. N. 2004. Influence of global DNA topology on cruciform formation in supercoiled DNA. *J Mol Biol*, 338, 735-43.
- PAESCHKE, K., BOCHMAN, M. L., GARCIA, P. D., CEJKA, P., FRIEDMAN, K. L., KOWALCZYKOWSKI, S. C. & ZAKIAN, V. A. 2013. Pif1 family helicases suppress genome instability at G-quadruplex motifs. *Nature*, 497, 458-62.
- PAESCHKE, K., CAPRA, J. A. & ZAKIAN, V. A. 2011. DNA replication through G-quadruplex motifs is promoted by the *Saccharomyces cerevisiae* Pif1 DNA helicase. *Cell*, 145, 678-91.
- PARKINSON, G. N., & COLLIE, G. W. 2019. X-Ray Crystallographic Studies of G-Quadruplex Structures. *G-Quadruplex Nucleic Acids*, 131-155.
- PAYNE, B. T., VAN KNIPPENBERG, I. C., BELL, H., FILIPE, S. R., SHERRATT, D. J. & MCGLYNN, P. 2006. Replication fork blockage by transcription factor-DNA complexes in *Escherichia coli*. *Nucleic Acids Res*, 34, 5194-202.
- PEDERSEN, J. M., FREDSOE, J., ROEDGAARD, M., ANDREASEN, L., MUNDBJERG, K., KRUIHOFFER, M., BRINCH, M., SCHIERUP, M. H., BJERGBAEK, L. & ANDERSEN, A. H. 2012. DNA Topoisomerases maintain promoters in a state competent for transcriptional activation in *Saccharomyces cerevisiae*. *PLoS Genet*, 8, e1003128.

- PETRUSKA, J., ARNHEIM, N. & GOODMAN, M. F. 1996. Stability of intrastrand hairpin structures formed by the CAG/CTG class of DNA triplet repeats associated with neurological diseases. *Nucleic Acids Res*, 24, 1992-8.
- PHATNANI, H. P., JONES, J. C. & GREENLEAF, A. L. 2004. Expanding the functional repertoire of CTD kinase I and RNA polymerase II: novel phosphoCTD-associating proteins in the yeast proteome. *Biochemistry*, 43, 15702-19.
- PIAZZA, A., ADRIAN, M., SAMAZAN, F., HEDDI, B., HAMON, F., SERERO, A., LOPES, J., TEULADE-FICHOU, M. P., PHAN, A. T. & NICOLAS, A. 2015. Short loop length and high thermal stability determine genomic instability induced by G-quadruplex-forming minisatellites. *EMBO J*, 34, 1718-34.
- PIAZZA, A., BOULE, J. B., LOPES, J., MINGO, K., LARGY, E., TEULADE-FICHOU, M. P. & NICOLAS, A. 2010. Genetic instability triggered by G-quadruplex interacting Phen-DC compounds in *Saccharomyces cerevisiae*. *Nucleic Acids Res*, 38, 4337-48.
- PIAZZA, A., CUI, X., ADRIAN, M., SAMAZAN, F., HEDDI, B., PHAN, A. T. & NICOLAS, A. G. 2017. Non-Canonical G-quadruplexes cause the hCEB1 minisatellite instability in *Saccharomyces cerevisiae*. *Elife*, 6.
- PLATELLA, C., RICCARDI, C., MONTESARCHIO, D., ROVIELLO, G. N. & MUSUMECI, D. 2017. G-quadruplex-based aptamers against protein targets in therapy and diagnostics. *Biochim Biophys Acta Gen Subj*, 1861, 1429-1447.
- POMMIER, Y., LEO, E., ZHANG, H. & MARCHAND, C. 2010. DNA topoisomerases and their poisoning by anticancer and antibacterial drugs. *Chem Biol*, 17, 421-33.

- POMMIER, Y., SUN, Y., HUANG, S. N. & NITISS, J. L. 2016. Roles of eukaryotic topoisomerases in transcription, replication and genomic stability. *Nat Rev Mol Cell Biol*, 17, 703-721.
- PRINGLE, J. R., ADAMS, A. E., DRUBIN, D. G. & HAARER, B. K. 1991. Immunofluorescence methods for yeast. *Methods Enzymol*, 194, 565-602.
- PRUNUSKE, A. J., WALTNER, J. K., KUHN, P., GU, B. & CRAIG, E. A. 2012. Role for the molecular chaperones Zuo1 and Ssz1 in quorum sensing via activation of the transcription factor Pdr1. *Proc Natl Acad Sci U S A*, 109, 472-7.
- RAMOS, P., KARNEZIS, A. N., HENDRICKS, W. P., WANG, Y., TEMBE, W., ZISMANN, V. L., LEGENDRE, C., LIANG, W. S., RUSSELL, M. L., CRAIG, D. W., FARLEY, J. H., MONK, B. J., ANTHONY, S. P., SEKULIC, A., CUNLIFFE, H. E., HUNTSMAN, D. G. & TRENT, J. M. 2014. Loss of the tumor suppressor SMARCA4 in small cell carcinoma of the ovary, hypercalcemic type (SCCOHT). *Rare Dis*, 2, e967148.
- RASHID, I., HAMMEL, M., SVERZHINSKY, A., TSAI, M. S., PASCAL, J. M., TAINER, J. A. & TOMKINSON, A. E. 2021. Direct interaction of DNA repair protein tyrosyl DNA phosphodiesterase 1 and the DNA ligase III catalytic domain is regulated by phosphorylation of its flexible N-terminus. *J Biol Chem*, 297, 100921.
- RAWAL, P., KUMMARASETTI, V. B., RAVINDRAN, J., KUMAR, N., HALDER, K., SHARMA, R., MUKERJI, M., DAS, S. K. & CHOWDHURY, S. 2006. Genome-wide prediction of G4 DNA as regulatory motifs: role in Escherichia coli global regulation. *Genome Res*, 16, 644-55.

- REDDY, K., ZAMIRI, B., STANLEY, S. Y. R., MACGREGOR, R. B., JR. & PEARSON, C. E. 2013. The disease-associated r(GGGGCC)<sub>n</sub> repeat from the C9orf72 gene forms tract length-dependent uni- and multimolecular RNA G-quadruplex structures. *J Biol Chem*, 288, 9860-9866.
- REDINBO, M. R., STEWART, L., KUHN, P., CHAMPOUX, J. J. & HOL, W. G. 1998. Crystal structures of human topoisomerase I in covalent and noncovalent complexes with DNA. *Science*, 279, 1504-13.
- REINA, C. & CAVALIERI, V. 2020. Epigenetic Modulation of Chromatin States and Gene Expression by G-Quadruplex Structures. *Int J Mol Sci*, 21.
- RENTON, A. E., MAJOUNIE, E., WAITE, A., SIMON-SANCHEZ, J., ROLLINSON, S., GIBBS, J. R., SCHYMICK, J. C., LAAKSOVIRTA, H., VAN SWIETEN, J. C., MYLLYKANGAS, L., KALIMO, H., PAETAU, A., ABRAMZON, Y., REMES, A. M., KAGANOVICH, A., SCHOLZ, S. W., DUCKWORTH, J., DING, J., HARMER, D. W., HERNANDEZ, D. G., JOHNSON, J. O., MOK, K., RYTEN, M., TRABZUNI, D., GUERREIRO, R. J., ORRELL, R. W., NEAL, J., MURRAY, A., PEARSON, J., JANSEN, I. E., SONDERVAN, D., SEELAAR, H., BLAKE, D., YOUNG, K., HALLIWELL, N., CALLISTER, J. B., TOULSON, G., RICHARDSON, A., GERHARD, A., SNOWDEN, J., MANN, D., NEARY, D., NALLS, M. A., PEURALINNA, T., JANSSON, L., ISOVIITA, V. M., KAIVORINNE, A. L., HOLTTA-VUORI, M., IKONEN, E., SULKAVA, R., BENATAR, M., WUU, J., CHIO, A., RESTAGNO, G., BORGHERO, G., SABATELLI, M., CONSORTIUM, I., HECKERMAN, D., ROGAEVA, E., ZINMAN, L., ROTHSTEIN, J. D., SENDTNER, M., DREPPER, C., EICHLER, E. E., ALKAN, C., ABDULLAEV, Z., PACK, S. D., DUTRA, A., PAK, E., HARDY, J., SINGLETON, A.,

- WILLIAMS, N. M., HEUTINK, P., PICKERING-BROWN, S., MORRIS, H. R., TIENARI, P. J. & TRAYNOR, B. J. 2011. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron*, 72, 257-68.
- RIBEYRE, C., LOPES, J., BOULE, J. B., PIAZZA, A., GUEDIN, A., ZAKIAN, V. A., MERGNY, J. L. & NICOLAS, A. 2009. The yeast Pif1 helicase prevents genomic instability caused by G-quadruplex-forming CEB1 sequences in vivo. *PLoS Genet*, 5, e1000475.
- ROCKMILL, B. 2009. Chromosome spreading and immunofluorescence methods in *Saccharomyces cerevisiae*. *Methods Mol Biol*, 558, 3-13.
- RODRIGUEZ, R., MILLER, K. M., FORMENT, J. V., BRADSHAW, C. R., NIKAN, M., BRITTON, S., OELSCHLAEGEL, T., XHEMALCE, B., BALASUBRAMANIAN, S. & JACKSON, S. P. 2012. Small-molecule-induced DNA damage identifies alternative DNA structures in human genes. *Nat Chem Biol*, 8, 301-10.
- SAITO, A., YAMASHITA, T., MARIKO, Y., NOSAKA, Y., TSUCHIYA, K., ANDO, T., SUZUKI, T., TSURUO, T. & NAKANISHI, O. 1999. A synthetic inhibitor of histone deacetylase, MS-27-275, with marked in vivo antitumor activity against human tumors. *Proc Natl Acad Sci U S A*, 96, 4592-7.
- SARKIES, P., MURAT, P., PHILLIPS, L. G., PATEL, K. J., BALASUBRAMANIAN, S. & SALE, J. E. 2012. FANCD1 coordinates two pathways that maintain epigenetic stability at G-quadruplex DNA. *Nucleic Acids Res*, 40, 1485-98.
- SARKIES, P., REAMS, C., SIMPSON, L. J. & SALE, J. E. 2010. Epigenetic instability due to defective replication of structured DNA. *Mol Cell*, 40, 703-13.

- SCHAFFITZEL, C., BERGER, I., POSTBERG, J., HANES, J., LIPPS, H. J. & PLUCKTHUN, A. 2001. In vitro generated antibodies specific for telomeric guanine-quadruplex DNA react with *Stylylonychia lemnae* macronuclei. *Proc Natl Acad Sci U S A*, 98, 8572-7.
- SCHIAVONE, D., GUILBAUD, G., MURAT, P., PAPADOPOULOU, C., SARKIES, P., PRIOLEAU, M. N., BALASUBRAMANIAN, S. & SALE, J. E. 2014. Determinants of G quadruplex-induced epigenetic instability in REV1-deficient cells. *EMBO J*, 33, 2507-20.
- SCHOFIELD, D. J., POPE, A. R., CLEMENTEL, V., BUCKELL, J., CHAPPLE, S., CLARKE, K. F., CONQUER, J. S., CROFTS, A. M., CROWTHER, S. R., DYSON, M. R., FLACK, G., GRIFFIN, G. J., HOOKS, Y., HOWAT, W. J., KOLB-KOKOCINSKI, A., KUNZE, S., MARTIN, C. D., MASLEN, G. L., MITCHELL, J. N., O'SULLIVAN, M., PERERA, R. L., ROAKE, W., SHADBOLT, S. P., VINCENT, K. J., WARFORD, A., WILSON, W. E., XIE, J., YOUNG, J. L. & MCCAFFERTY, J. 2007. Application of phage display to high throughput antibody generation and characterization. *Genome Biol*, 8, R254.
- SCULLY, R., PANDAY, A., ELANGO, R., & WILLIS, N. A. 2019. DNA double-strand break repair-pathway choice in somatic mammalian cells. *Nat Rev Mol Cell Biol*, 20, 698-714.
- SEKIGUCHI, J. & SHUMAN, S. 1994. Requirements for noncovalent binding of vaccinia topoisomerase I to duplex DNA. *Nucleic Acids Res*, 22, 5360-5.
- SEKIGUCHI, J. & SHUMAN, S. 1997. Site-specific ribonuclease activity of eukaryotic DNA topoisomerase I. *Mol Cell*, 1, 89-97.
- SHUAI, L., DENG, M., ZHANG, D., ZHOU, Y. & ZHOU, X. 2010. Quadruplex-duplex motifs as new topoisomerase I inhibitors. *Nucleosides Nucleotides Nucleic Acids*, 29, 841-53.



- SINGH, S., BERROYER, A., KIM, M. & KIM, N. 2020. Yeast Nucleolin Nsr1 Impedes Replication and Elevates Genome Instability at an Actively Transcribed Guanine-Rich G4 DNA-Forming Sequence. *Genetics*, 216, 1023-1037.
- SIU, F. M. & POMMIER, Y. 2013. Sequence selectivity of the cleavage sites induced by topoisomerase I inhibitors: a molecular dynamics study. *Nucleic Acids Res*, 41, 10010-9.
- SMYTH, M. S., & MARTIN, J. H. J. 2000. x Ray crystallography. *Mol Pathol*, 53, 8-14.
- SOLLIER, J., STORK, C. T., GARCIA-RUBIO, M. L., PAULSEN, R. D., AGUILERA, A. & CIMPRICH, K. A. 2014. Transcription-coupled nucleotide excision repair factors promote R-loop-induced genome instability. *Mol Cell*, 56, 777-85.
- SPELL, R. M. & JINKS-ROBERTSON, S. 2004. Determination of mitotic recombination rates by fluctuation analysis in *Saccharomyces cerevisiae*. *Methods Mol Biol*, 262, 3-12.
- SPERLING, A. S., JEONG, K. S., KITADA, T. & GRUNSTEIN, M. 2011. Topoisomerase II binds nucleosome-free DNA and acts redundantly with topoisomerase I to enhance recruitment of RNA Pol II in budding yeast. *Proc Natl Acad Sci U S A*, 108, 12693-8.
- STAHL, H., DROGE, P. & KNIPPERS, R. 1986. DNA helicase activity of SV40 large tumor antigen. *EMBO J*, 5, 1939-44.
- STEWART, L., REDINBO, M. R., QIU, X., HOL, W. G. & CHAMPOUX, J. J. 1998. A model for the mechanism of human topoisomerase I. *Science*, 279, 1534-41.
- STINGELE, J., BELLELLI, R., ALTE, F., HEWITT, G., SAREK, G., MASLEN, S. L., TSUTAKAWA, S. E., BORG, A., KJAER, S., TAINER, J. A., SKEHEL, J. M., GROLL, M. & BOULTON, S. J. 2016.

- Mechanism and Regulation of DNA-Protein Crosslink Repair by the DNA-Dependent Metalloprotease SPRTN. *Mol Cell*, 64, 688-703.
- STINGELE, J., BELLELLI, R. & BOULTON, S. J. 2017. Mechanisms of DNA-protein crosslink repair. *Nat Rev Mol Cell Biol*, 18, 563-573.
- STINGELE, J., SCHWARZ, M. S., BLOEMEKE, N., WOLF, P. G. & JENTSCH, S. 2014. A DNA-dependent protease involved in DNA-protein crosslink repair. *Cell*, 158, 327-338.
- STORICI, F. & RESNICK, M. A. 2006. The delitto perfetto approach to in vivo site-directed mutagenesis and chromosome rearrangements with synthetic oligonucleotides in yeast. *Methods Enzymol*, 409, 329-45.
- SUN, D., GUO, K. & SHIN, Y. J. 2011. Evidence of the formation of G-quadruplex structures in the promoter region of the human vascular endothelial growth factor gene. *Nucleic Acids Res*, 39, 1256-65.
- SUN, D. & HURLEY, L. H. 2009. The importance of negative superhelicity in inducing the formation of G-quadruplex and i-motif structures in the c-Myc promoter: implications for drug targeting and control of gene expression. *J Med Chem*, 52, 2863-74.
- SUN, H., KAROW, J. K., HICKSON, I. D. & MAIZELS, N. 1998. The Bloom's syndrome helicase unwinds G4 DNA. *J Biol Chem*, 273, 27587-92.
- TADDEI, A. & GASSER, S. M. 2012. Structure and function in the budding yeast nucleus. *Genetics*, 192, 107-29.
- TAJRISHI, M. M., TUTEJA, R. & TUTEJA, N. 2011. Nucleolin: The most abundant multifunctional phosphoprotein of nucleolus. *Commun Integr Biol*, 4, 267-75.

- TAN, J., WANG, X., PHOON, L., YANG, H. & LAN, L. 2020. Resolution of ROS-induced G-quadruplexes and R-loops at transcriptionally active sites is dependent on BLM helicase. *FEBS Lett*, 594, 1359-1367.
- TENG, F. Y., JIANG, Z. Z., GUO, M., TAN, X. Z., CHEN, F., XI, X. G. & XU, Y. 2021a. G-quadruplex DNA: a novel target for drug design. *Cell Mol Life Sci*, 78, 6557-6583.
- TENG, Y. C., SUNDARESAN, A., O'HARA, R., GANT, V. U., LI, M., MARTIRE, S., WARSHAW, J. N., BASU, A. & BANASZYNSKI, L. A. 2021b. ATRX promotes heterochromatin formation to protect cells from G-quadruplex DNA-mediated stress. *Nat Commun*, 12, 3887.
- THIYAGARAJAN, M. M., WALDMAN, S. A., NOE, M. & KMIEC, E. B. 1998. Binding characteristics of *Ustilago maydis* topoisomerase I to DNA containing secondary structures. *Eur J Biochem*, 255, 347-55.
- TODD, A. K., JOHNSTON, M. & NEIDLE, S. 2005. Highly prevalent putative quadruplex sequence motifs in human DNA. *Nucleic Acids Res*, 33, 2901-7.
- TOMAZOU, E. M., SHEFFIELD, N. C., SCHMIDL, C., SCHUSTER, M., SCHONEGGER, A., DATLINGER, P., KUBICEK, S., BOCK, C. & KOVAR, H. 2015. Epigenome mapping reveals distinct modes of gene regulation and widespread enhancer reprogramming by the oncogenic fusion protein EWS-FLI1. *Cell Rep*, 10, 1082-95.
- TOSONI, E., FRASSON, I., SCALABRIN, M., PERRONE, R., BUTOVSKAYA, E., NADAI, M., PALU, G., FABRIS, D. & RICHTER, S. N. 2015. Nucleolin stabilizes G-quadruplex structures folded by the LTR promoter and silences HIV-1 viral transcription. *Nucleic Acids Res*, 43, 8884-97.

- TRIGUEROS, S. & ROCA, J. 2001. Circular minichromosomes become highly recombinogenic in topoisomerase-deficient yeast cells. *J Biol Chem*, 276, 2243-8.
- TRIGUEROS, S. & ROCA, J. 2002. Failure to relax negative supercoiling of DNA is a primary cause of mitotic hyper-recombination in topoisomerase-deficient yeast cells. *J Biol Chem*, 277, 37207-11.
- TSURUTANI, J., NITTA, T., HIRASHIMA, T., KOMIYA, T., UEJIMA, H., TADA, H., SYUNICHI, N., TOHDA, A., FUKUOKA, M. & NAKAGAWA, K. 2002. Point mutations in the topoisomerase I gene in patients with non-small cell lung cancer treated with irinotecan. *Lung Cancer*, 35, 299-304.
- TU, J., DUAN, M., LIU, W., LU, N., ZHOU, Y., SUN, X. & LU, Z. 2021. Direct genome-wide identification of G-quadruplex structures by whole-genome resequencing. *Nat Commun*, 12, 6014.
- TUBBS, A. & NUSSENZWEIG, A. 2017. Endogenous DNA Damage as a Source of Genomic Instability in Cancer. *Cell*, 168, 644-656.
- TUESUWAN, B., KERN, J. T., THOMAS, P. W., RODRIGUEZ, M., LI, J., DAVID, W. M. & KERWIN, S. M. 2008. Simian virus 40 large T-antigen G-quadruplex DNA helicase inhibition by G-quadruplex DNA-interactive agents. *Biochemistry*, 47, 1896-909.
- UMOH, M. E., DAMMER, E. B., DAI, J., DUONG, D. M., LAH, J. J., LEVEY, A. I., GEARING, M., GLASS, J. D. & SEYFRIED, N. T. 2018. A proteomic network approach across the ALS-FTD disease spectrum resolves clinical phenotypes and genetic vulnerability in human brain. *EMBO Mol Med*, 10, 48-62.

- VAN ATTIKUM, H., FRITSCH, O. & GASSER, S. M. 2007. Distinct roles for SWR1 and INO80 chromatin remodeling complexes at chromosomal double-strand breaks. *EMBO J*, 26, 4113-25.
- VAN HOLDE, K. & ZLATANOVA, J. 1994. Unusual DNA structures, chromatin and transcription. *Bioessays*, 16, 59-68.
- VAN WIETMARSCHEN, N., MERZOUK, S., HALSEMA, N., SPIERINGS, D. C. J., GURYEV, V. & LANSDORP, P. M. 2018. BLM helicase suppresses recombination at G-quadruplex motifs in transcribed genes. *Nat Commun*, 9, 271.
- VARIZHUK, A., ISAAKOVA, E. & POZMOGOVA, G. 2019. DNA G-Quadruplexes (G4s) Modulate Epigenetic (Re)Programming and Chromatin Remodeling: Transient Genomic G4s Assist in the Establishment and Maintenance of Epigenetic Marks, While Persistent G4s May Erase Epigenetic Marks. *Bioessays*, 41, e1900091.
- VAZ, B., POPOVIC, M., NEWMAN, J. A., FIELDEN, J., AITKENHEAD, H., HALDER, S., SINGH, A. N., VENDRELL, I., FISCHER, R., TORRECILLA, I., DROBNITZKY, N., FREIRE, R., AMOR, D. J., LOCKHART, P. J., KESSLER, B. M., MCKENNA, G. W., GILEADI, O. & RAMADAN, K. 2016. Metalloprotease SPRTN/DVC1 Orchestrates Replication-Coupled DNA-Protein Crosslink Repair. *Mol Cell*, 64, 704-719.
- WALLIS, J. W., CHREBET, G., BRODSKY, G., ROLFE, M. & ROTHSTEIN, R. 1989. A hyper-recombination mutation in *S. cerevisiae* identifies a novel eukaryotic topoisomerase. *Cell*, 58, 409-19.
- WANG, G. & VASQUEZ, K. M. 2014. Impact of alternative DNA structures on DNA damage, DNA repair, and genetic instability. *DNA Repair (Amst)*, 19, 143-51.

- WANG, Y., YANG, J., WILD, A. T., WU, W. H., SHAH, R., DANUSSI, C., RIGGINS, G. J., KANNAN, K., SULMAN, E. P., CHAN, T. A. & HUSE, J. T. 2019. G-quadruplex DNA drives genomic instability and represents a targetable molecular abnormality in ATRX-deficient malignant glioma. *Nat Commun*, 10, 943.
- WATT, P. M., HICKSON, I. D., BORTS, R. H. & LOUIS, E. J. 1996. SGS1, a homologue of the Bloom's and Werner's syndrome genes, is required for maintenance of genome stability in *Saccharomyces cerevisiae*. *Genetics*, 144, 935-45.
- WEITZMANN, M. N., WOODFORD, K. J. & USDIN, K. 1996. The development and use of a DNA polymerase arrest assay for the evaluation of parameters affecting intrastrand tetraplex formation. *J Biol Chem*, 271, 20958-64.
- WELLS, R. D. 2007. Non-B DNA conformations, mutagenesis and disease. *Trends Biochem Sci*, 32, 271-8.
- WILLIAMS, J. D., FLEETWOOD, S., BERROYER, A., KIM, N. & LARSON, E. D. 2015. Sites of instability in the human TCF3 (E2A) gene adopt G-quadruplex DNA structures in vitro. *Front Genet*, 6, 177.
- WILLIAMS, J. D., HOUSEROVA, D., JOHNSON, B. R., DYNIEWSKI, B., BERROYER, A., FRENCH, H., BARCHIE, A. A., BILBREY, D. D., DEMEIS, J. D., GHEE, K. R., HUGHES, A. G., KREITZ, N. W., MCINNIS, C. H., PUDNER, S. C., REEVES, M. N., STAHLY, A. N., TURCU, A., WATTERS, B. C., DALY, G. T., LANGLEY, R. J., GILLESPIE, M. N., PRAKASH, A., LARSON, E. D., KASUKURTHI, M. V., HUANG, J., JINKS-ROBERTSON, S. & BORCHERT, G. M. 2020. Characterization of long G4-rich enhancer-associated genomic regions

- engaging in a novel loop:loop 'G4 Kissing' interaction. *Nucleic Acids Res*, 48, 5907-5925.
- WILLIAMS, J. S., SMITH, D. J., MARJAVAARA, L., LUJAN, S. A., CHABES, A. & KUNKEL, T. A. 2013. Topoisomerase 1-mediated removal of ribonucleotides from nascent leading-strand DNA. *Mol Cell*, 49, 1010-5.
- WONG, L. H., MCGHIE, J. D., SIM, M., ANDERSON, M. A., AHN, S., HANNAN, R. D., GEORGE, A. J., MORGAN, K. A., MANN, J. R. & CHOO, K. H. 2010. ATRX interacts with H3.3 in maintaining telomere structural integrity in pluripotent embryonic stem cells. *Genome Res*, 20, 351-60.
- 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. *J Biol Chem*, 269, 27029-35.
- WRIGHT, C. M., VAN DER MERWE, M., DEBROT, A. H. & BJORNSTI, M. A. 2015. DNA topoisomerase I domain interactions impact enzyme activity and sensitivity to camptothecin. *J Biol Chem*, 290, 12068-78.
- WU, Y., SHIN-YA, K. & BROSH, R. M., JR. 2008. FANCI helicase defective in Fanconi anemia and breast cancer unwinds G-quadruplex DNA to defend genomic stability. *Mol Cell Biol*, 28, 4116-28.
- 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 Genet*, 10, e1004839.

- YADAV, P., KIM, N., KUMARI, M., VERMA, S., SHARMA, T. K., YADAV, V. & KUMAR, A. 2021. G-Quadruplex Structures in Bacteria: Biological Relevance and Potential as an Antimicrobial Target. *J Bacteriol*, 203, e0057720.
- 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 Res*, 44, 718-29.
- YAN, W., SCHILKE, B., PFUND, C., WALTER, W., KIM, S. & CRAIG, E. A. 1998. Zuotin, a ribosome-associated DnaJ molecular chaperone. *EMBO J*, 17, 4809-17.
- YANG, X., LI, Y., GAO, Z., LI, Z., XU, J., WANG, W. & DONG, Y. 2017. Structural analysis of Wss1 protein from *saccharomyces cerevisiae*. *Sci Rep*, 7, 8270.
- YANG, Z., CAREY, J. F. & CHAMPOUX, J. J. 2009. Mutational analysis of the preferential binding of human topoisomerase I to supercoiled DNA. *FEBS J*, 276, 5906-19.
- YIN, Y., MORGUNOVA, E., JOLMA, A., KAASINEN, E., SAHU, B., KHUND-SAYEED, S., DAS, P. K., KIVIOJA, T., DAVE, K., ZHONG, F., NITTA, K. R., TAIPALE, M., POPOV, A., GINNO, P. A., DOMCKE, S., YAN, J., SCHUBELER, D., VINSON, C. & TAIPALE, J. 2017. Impact of cytosine methylation on DNA binding specificities of human transcription factors. *Science*, 356.
- ZHANG, H., XIONG, Y. & CHEN, J. 2020. DNA-protein cross-link repair: what do we know now? *Cell Biosci*, 10, 3.



- ZHANG, S., SUN, H., WANG, L., LIU, Y., CHEN, H., LI, Q., GUAN, A., LIU, M. & TANG, Y. 2018. Real-time monitoring of DNA G-quadruplexes in living cells with a small-molecule fluorescent probe. *Nucleic Acids Res*, 46, 7522-7532.
- ZHANG, T., WALLIS, M., PETROVIC, V., CHALLIS, J., KALITSIS, P. & HUDSON, D. F. 2019. Loss of TOP3B leads to increased R-loop formation and genome instability. *Open Biol*, 9, 190222.
- ZHAO, J., BACOLLA, A., WANG, G. & VASQUEZ, K. M. 2010. Non-B DNA structure-induced genetic instability and evolution. *Cell Mol Life Sci*, 67, 43-62.
- ZHENG, D. Q., ZHANG, K., WU, X. C., MIECZKOWSKI, P. A. & PETES, T. D. 2016. Global analysis of genomic instability caused by DNA replication stress in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A*, 113, E8114-E8121.
- ZHENG, K. W., XIAO, S., LIU, J. Q., ZHANG, J. Y., HAO, Y. H. & TAN, Z. 2013. Co-transcriptional formation of DNA:RNA hybrid G-quadruplex and potential function as constitutional cis element for transcription control. *Nucleic Acids Res*, 41, 5533-41.
- ZUCO, V., SUPINO, R., FAVINI, E., TORTORETO, M., CINCINELLI, R., CROCE, A. C., BUCCI, F., PISANO, C. & ZUNINO, F. 2010. Efficacy of ST1968 (namitecan) on a topotecan-resistant squamous cell carcinoma. *Biochem Pharmacol*, 79, 535-41.
- ZYNER, K. G., MULHEARN, D. S., ADHIKARI, S., MARTINEZ CUESTA, S., DI ANTONIO, M., ERARD, N., HANNON, G. J., TANNAHILL, D. & BALASUBRAMANIAN, S. 2019. Genetic interactions of G-quadruplexes in humans. *Elife*, 8.

## **Vita**

Alexandra Berroyer is the daughter of Timothy Berroyer and Susan Fleming. After finishing high school in Nokomis, Illinois in 2010, she attended college at Millikin University in Decatur, Illinois. In 2014, she received a Bachelor of Science degree with a major in biology and a minor in chemistry. Immediately after completing her undergraduate studies, she attended Illinois State University for 2 years to earn a master's degree in biotechnology. In August of 2016, she began working towards a Ph.D. at The University of Texas MD Anderson Cancer Center UTHHealth Graduate School of Biomedical Sciences