

Stabilization of Topoisomerase 2 Mutants Initiates the  
Formation of Duplications in DNA  
by

Nicole Stantial

Department of Molecular Genetics and Microbiology  
Duke University

Date: \_\_\_\_\_

Approved:

\_\_\_\_\_  
Sue Jinks-Robertson, Supervisor

\_\_\_\_\_  
David M. MacAlpine

\_\_\_\_\_  
John H. McCusker

\_\_\_\_\_  
Thomas Petes

\_\_\_\_\_  
Beth A. Sullivan

Dissertation submitted in partial fulfillment of  
the requirements for the degree of Doctor of Philosophy in the Department of  
Molecular Genetics and Microbiology in the Graduate School  
of Duke University

2021

ABSTRACT

Stabilization of Topoisomerase 2 Mutants Initiates the  
Formation of Duplications in DNA  
by

Nicole Stantial

Department of Molecular Genetics and Microbiology  
Duke University

Date: \_\_\_\_\_

Approved:

\_\_\_\_\_  
Sue Jinks-Robertson, Supervisor

\_\_\_\_\_  
David M. MacAlpine

\_\_\_\_\_  
John H. McCusker

\_\_\_\_\_  
Thomas Petes

\_\_\_\_\_  
Beth A. Sullivan

An abstract of a dissertation submitted in partial  
fulfillment of the requirements for the degree of  
Doctor of Philosophy in the Department of  
Molecular Genetics and Microbiology in the Graduate School of  
Duke University

2021

Copyright by  
Nicole Stantial  
2021

## Abstract

Topoisomerase 2 (Top2) is an enzyme that helps maintain genome integrity by resolving topological structures that arise during cellular processes such as replication and transcription. To resolve these structures, a Top2 dimer creates a transient double-strand break (DSB) in the DNA. Each subunit forms a phosphotyrosyl bond with the 5' ends of the break, and this DNA-protein intermediate is called a Top2 cleavage complex (Top2cc). Following the passage of an intact duplex, Top2 re-ligates the DNA and is released to restore genome integrity. Top2cc stabilization by chemotherapeutic drugs such as etoposide leads to persistent and potentially toxic DSBs. This thesis characterizes two self-poisoning *top2* mutants that are associated with a mutation signature characterized by *de novo* duplications. These duplication events are dependent on clean removal of the Top2cc by tyrosyl-DNA phosphodiesterase 1 (TDP1) and DSB repair by nonhomologous end-joining. Removal of the Top2cc by Mre11 prevents the formation of the duplication events. The first mutant (*top2-FY,RG*) was identified through a screen for etoposide hypersensitivity, and it generates a stabilized cleavage intermediate *in vitro*. The second mutant (*top2-K720N*) is the yeast equivalent of a somatic mutation in human TOP2A identified in gastric cancers and cholangiocarcinomas that is also associated with a duplication mutation signature (ID17). Overall, the findings in this thesis are relevant for clinical use of chemotherapeutic drugs that target Top2 and have implications for genome evolution.

## **Dedication**

This thesis is dedicated to my late mother, Beth Stantial, who was always my greatest supporter. I wouldn't have gotten here without her. Love you always.

# Contents

Abstract .....	iv
List of Tables .....	xi
List of Figures .....	xii
List of Abbreviations .....	xiv
Acknowledgements .....	xvii
Chapter 1. Introduction .....	1
1.1 General Importance of Topoisomerases .....	1
1.2 Topoisomerase Biology .....	2
1.2.1 Classification of Topoisomerases .....	2
1.2.1.1 Type IA Topoisomerases .....	3
1.2.1.2 Type IB Topoisomerases .....	3
1.2.1.3 Type IC Topoisomerases .....	6
1.2.1.4 Type IIA Topoisomerases .....	6
1.2.1.5 Type IIB Topoisomerases .....	7
1.2.2 Topoisomerase II Catalytic Cycle .....	7
1.2.3 Topoisomerases in Replication and Chromosome Condensation .....	9
1.2.4 The Role of Topoisomerases in Transcription .....	13
1.3 Trapping the Top2 Cleavage Complex .....	18
1.3.1 Top2 Poisons .....	19
1.3.2 Endogenous Stabilization of the Top2 Cleavage Complex .....	24
1.4 Removal of the Stabilized TOP2 Cleavage Complex .....	27
1.4.1 Ubiquitination and SUMOylation of TOP2 Cleavage Complex for Proteasomal Degradation .....	27

1.4.2	Protease-dependent Removal of TOP2 Cleavage Complex .....	30
1.4.3	TDP1 and TDP2 .....	32
1.4.4	MRE11-dependent Repair of TOP2 Cleavage Complex .....	34
1.5	TOP2 Cleavage Consensus Sequence.....	35
1.6	TOP2-dependent Mutagenesis .....	39
1.6.1	TOP2 Associated Translocations .....	39
1.6.2	TOP2 and Secondary Structure-mediated DNA Fragility.....	39
1.6.3	Studies of Top2 Dependent Duplications Described in this Thesis.....	40
Chapter 2. Trapped topoisomerase II initiates formation of <i>de novo</i> duplications via the nonhomologous end-joining pathway in yeast .....		41
2.1	Introduction .....	41
2.2	Materials and Methods .....	44
2.2.1	Strains and Growth Conditions .....	44
2.2.2	Strain Constructions .....	44
2.1.3	<i>In vivo</i> Complex of Enzyme (ICE) Assay in Yeast .....	45
2.2.4	Western Blotting .....	47
2.2.5	Topoisomerase Biochemical Assays .....	48
2.2.6	Mutation and Recombination Rate Measurements.....	49
2.2.7	<i>CAN1</i> Mutation Spectra.....	50
2.2.8	<i>LYS2</i> Mutation Spectra.....	51
2.2.9	Recombination Experiments.....	52
2.2.10	Statistical Methods .....	52
2.3	Results .....	53
2.3.1	C-terminal <i>TOP2</i> Mutations Confer Hypersensitivity to Etoposide.....	53

2.3.2 Expression of <i>top2-RG</i> or <i>top2-FY,RG</i> is lethal in the absence of <i>RAD52</i> .....	54
2.3.3 Top2-FY,RG is trapped on yeast DNA in the absence of etoposide... 58	58
2.3.4 Relaxation and cleavage activities of mutant Top2 proteins.....	58
2.3.5 Top2-FY,RG is mutagenic and is associated with <i>de novo</i> duplications .....	63
2.3.6 Top2 overexpression elevates duplications in frameshift reversion assays .....	65
2.3.7 Top2cc-dependent duplications require the NHEJ pathway .....	67
2.3.8 Pathways for Top2cc removal affect duplication rates.....	68
2.3.9 Wss1 and Rad2 are not involved in the formation of Top2-dependent duplications .....	71
2.3.10 Mutagenic effects of etoposide .....	72
2.4 Discussion.....	73
Chapter 3. Characterization of Yeast Top2 Mutant Equivalent to a Somatic TOP2A Mutation Identified in Human Gastric Cancers and Cholangiocarcinomas .....	83
3.1 Introduction .....	83
3.2 Materials and Methods .....	86
3.2.1 Strains and growth conditions.....	86
3.2.2 Strain constructions .....	86
3.2.3 Mutation rate measurements.....	87
3.2.4 <i>CAN1</i> mutation spectrum .....	88
3.2.5 Generation of heat maps .....	89
3.2.6 RT-qPCR.....	90
3.2.7 Statistical methods .....	90

3.3 Results .....	92
3.3.1 Top2-K720N is associated with <i>de novo</i> duplications.....	92
3.3.2 Top2-K720N duplications have similar genetic requirements as Top2-FY,RG duplications.....	95
3.3.3 Distribution of the Top2-K720N duplications differs from Top2-FY,RG duplications .....	98
3.3.4 Top1 is required for Top1-FY,RG associated duplications but not Top2-K720N duplications.....	101
3.4 Discussion.....	103
3.4.1 Using binomial sampling to define duplication hotspots .....	104
3.4.2 Top1 involvement in Top2-FY,RG associated duplications .....	108
3.4.4 TOP2 p.K7432N is associated with 6- to 8-bp deletions not observed in Top2-K720N .....	109
3.4.5 Conclusion.....	110
4. Conclusions .....	111
Appendix A.....	123
Appendix B.....	125
Appendix C.....	126
Appendix D.....	127
Appendix E.....	128
Appendix F.....	129
Appendix G.....	130
Appendix H.....	131
Appendix I.....	132
Appendix J.....	133

Appendix K.....	134
Appendix L.....	135
Appendix M.....	136
Appendix N.....	137
Appendix O.....	138
References.....	139
Biography.....	173

## List of Tables

Table 1: Summary of the Topoisomerase Subfamilies .....	5
Table 2: Summary of the TOP2 Cleavage Consensus Sequence in the Presence of TOP2-targeting drugs .....	36
Table 3: TOP2 Cleavage Consensus Sequence Summary .....	38
Table 4: Non-essential Components of the 26S Proteasome.....	113
Table 5: List of Yeast Strains from Chapter 2.....	123
Table 6: Can-R Median Mutation Rates and Associated Mutation Types.....	127
Table 7: Can-R Median Frequencies and Associated Mutation Types in the Presence of Etoposide. ....	129
Table 8: List of Yeast Strains from Chapter 3.....	132
Table 9: Can-R Rates and Associated Mutation Types.....	134
Table 10: Can-R Rates and “Other” Mutation Types Associated with <i>top2-FY, RG</i> and <i>top2-K720N</i> .....	137
Table 11: >1 bp Deletions Observed in the <i>tdp1Δ</i> Strains Expressing <i>top2-K720N</i> .....	137

## List of Figures

Figure 1: Top2 Catalytic Cycle.....	8
Figure 2: The Roles of Top1 and Top2 during DNA Replication.....	9
Figure 3: The Roles of Top1 and Top2 During Transcription.....	14
Figure 4: Top2 poisons and their sequence preference. ....	21
Figure 5: Removal of the stabilized TOP2cc .....	29
Figure 6: Genetic Characterization of <i>top2</i> Mutants .....	56
Figure 7: Biochemical Activities of WT and Mutant Proteins. ....	60
Figure 8: Progression Through the Top2 Catalytic Cycle is not Required for Elevated Cleavage. ....	62
Figure 9: Top2-FY, RG Expression is Associated with <i>de novo</i> Duplications.....	66
Figure 10: Genetic Control of Mutations Associated with <i>top2-FY, RG</i> Expression. ....	70
Figure 11: Etoposide Stimulates Duplication Through its Interaction with Top2. ....	73
Figure 12: The F1025Y/R1128G Mutations Map to the Top2 C-gate. ....	75
Figure 13: Models for Top2-associated Duplications.....	77
Figure 14: ID17 Mutation Signature from COSMIC. ....	84
Figure 15: Top2-FY, RG and Top2-K720N are associated with <i>de novo</i> duplications.....	94
Figure 16: Genetic Requirements for Top2-K720N Associated Duplications. ....	97
Figure 17: Top2-FY, RG and Top2-K720N Duplication Distribution and Hotspots. ....	99
Figure 18: Effects of Top1 Loss on Top2-FY, RG and Top2-K720N Associated Duplications. ....	102
Figure 19: Structure of Yeast Top2.....	107
Figure 20: Potential Mechanisms for Top2cc Degradation and Removal. ....	116

Figure 21: <i>lys2</i> Reversion Assay with the Top2 Hotspots.....	119
Figure 22: <i>MLL</i> Translocation Assay. ....	121
Figure 23: Effects of <i>m-AMSA</i> on Viability and Top2 Cleavage Activity. ....	125
Figure 24: Quantitation of Top2 cleavage activity.....	126
Figure 25: Reversion Spectra for the <i>lys2DA746,NR -1</i> and <i>lys2DBgl1,NR +1</i> Frameshift Allele.....	128
Figure 26: Composite <i>can1</i> Spectrum of Duplications >1 bp in WT and <i>mre11-D56N</i> Strains Containing <i>pDED1-top2-FY,RG</i> .....	130
Figure 27: Sequence logo of 4-bp Duplications in WT and <i>mre11-D56N</i> strains containing <i>pDED1Top2-FY,RG</i> . ....	131
Figure 28: Using Binomial Sampling to Identify Duplication Hotspots in <i>can1</i> ..	133
Figure 29: Distribution of Duplication Sizes in the <i>top2-FY,RG</i> and <i>top2-K720N</i> Strains in the WT Background. ....	135
Figure 30: Composite of 2- to 5-bp Duplications in <i>top2-FY,RG</i> and <i>top2-K720N</i> strains.....	136
Figure 31: Composite of 4-bp Duplications in <i>top2-FY,RG</i> and <i>top2-K720N</i> strains.....	138

## List of Abbreviations

ALL	Acute Lymphocytic Leukemia
AML	Acute Myeloid Leukemia
bp	Base pairs
Can-R	Canavanine-resistant
CCS	Circular Consensus Sequence
ChIP-seq	Chromatin Immunoprecipitation Sequencing
CIs	Confidence Intervals
COSMIC	Catalogue of Somatic Mutations in Cancer
CTCF	CCCTC-Binding Factor
CTD	Carboxyl-terminal Domain
DPCs	DNA-protein Crosslinks
DSBs	Double-strand Break
DTT	Dithiothreitol
EV	Empty Vector
FRET	Förster Resonance Energy Transfer
G-segment	Gate-segment
HR	Homologous Recombination
ICE	<i>In vivo</i> Complex of Enzyme
ID	Insertion-deletion Signature
Kb	Kilobases

<i>MLL</i>	Mixed Lineage Leukemia
MMEJ	Microhomology-mediated End Joining
MRN(X)	MRE11-RAD50-NBS1(Xrs2)
<i>mre11-D56N</i>	Nuclease Dead Mre11
mRNA	Messenger RNA
N	Number
NHEJ	Non-homologous End Joining
Nt	Nucleotides
PCAWG-ICGC	Pan-Cancer Analysis of Whole Genomes Consortium of the International Cancer Genome Consortium
<i>pDED1</i>	<i>DED1</i> Promoter
PNKP	Polynucleotide Kinase Phosphatase
Pol4	Polymerase 4
PTS	Polyamine Transport System
RNAP II	RNA Polymerase II
rRNA	Ribosomal RNA
RSS	Root of the Square of the Sums
RT	Room Temperature
SBS	Single-base Substitution
SCAN1	Spinocerebellar Ataxia with Axonal Neuropathy
SCAR23	Spinocerebellar Ataxia Autosomal Recessive 23
SF2 helicase	Superfamily 2 Helicase

SPRTN	SprT-like Domain at the N-terminus
SUMOs	Small Ubiquitin-like MOdifiers
SC	Synthetic Complete
T-segment	Transfer-segment
TCGA	The Cancer Genome Atlas
TERs	Chromosomal Termination Sites
TOP2	Topoisomerase 2 - Mammals
Top2	Topoisomerase 2 – Yeast
Top2cc	Top2 Cleavage Complex
Top2-FY,RG	Top2-F1025Y,R1128G
Top2-RG	Top2-R1128G
<i>top2-ts</i>	<i>top2</i> Temperature Sensitive Alleles
Tdp1/2	Tyrosyl-phosphodiesterase 1/2
UPP	Ubiquitin-protease Pathway
WHD	Winged-helix Domain
Wss1	Weak Suppressor of Smt3
ZATT	Zinc Finger Protein Associated with TDP2 and TOP2

## Acknowledgements

I would like to thank my advisor, Sue Jinks-Robertson, for all the help and guidance she has given me. She has created an amazing lab environment that fosters growth and collaboration. I would also like to thank Tom Petes for all of his advice at Friday lab meetings, and the great conversations during Beer Hour. Additionally, I would like to thank my committee Beth Sullivan, David MacAlpine, and John McCusker for their guidance and probing questions. A special thank you goes to the collaborators I worked with on both my Top2 projects, especially John Nitiss and his lab for identifying the *top2-FY,RG* mutant and performing all the biochemistry, as well as Steve Rozen and Arnoud Boot for discovering the TOP2Ap.K743N mutation in the COSMIC database. The day to day activities in lab wouldn't have been as enjoyable without the past and present SJR lab members, including Sam, Dionna, Wei, Asiya, Eva Mei, Lisa, Jonny, and Eva H. I would also like to thank all of the past and present members of the Petes lab for being happy faces to see every day. I couldn't have made it through graduate school without my CMB cohort, especially Taylor Krebs and Hannah McMillan. I can't wait to see what is in store for us next. I want to thank my best friend Sam Kruger for always being there for me and my amoebas. My dad, sister, niece and future brother-in law have always been there for me in the ups and downs and are a great break from lab life. Finally, the biggest thank you goes to my fiance David and our cat Candy Cane. Their love and support have been my greatest motivation. I can't thank you enough!

# Chapter 1. Introduction

## 1.1 General Importance of Topoisomerases

The unwinding of DNA during cellular processes results in unique topological problems, such as DNA supercoils and catenanes. Supercoils form ahead and behind replication and transcription machineries, and catenanes result from newly synthesized DNA becoming entangled. Both of these structures are resolved by topoisomerases, which create transient single- or double-strand breaks in DNA to help restore genomic integrity. Topoisomerases have been identified in eukaryotes, archaeobacteria, and eubacteria, emphasizing their importance in cellular functions. Introducing breaks into the genome seems counterintuitive since unrepaired breaks can be lethal, but this action is crucial for cellular activities. Topoisomerases prevent rearrangements or recombination by creating transient covalent links with the DNA ends, which block the activation of DNA damage response pathways (1). Once the structures are resolved, topoisomerases re-ligate the DNA and restore its integrity. There are various types of topoisomerases, and the focus of this thesis will be on eukaryotic topoisomerase 2 (Top2 in yeast; TOP2 in mammals). To resolve topological structures, Top2 creates transient double-strand breaks (DSBs) in the genome.

A key intermediate of the Top2 catalytic cycle is the transient DNA-protein intermediate created when the enzyme cleaves DNA. Top2 acts as a homodimer and each subunit forms a phosphotyrosyl bond with the 5' end of a single strand break. This intermediate is called the Top2 cleavage complex (Top2cc), and it is

a target for numerous chemotherapeutics and antibiotics that are widely used in the clinic (2). The general mechanism of these drugs is to stabilize the Top2cc and increase the lifetime of the enzyme on the DNA. This leads to conflicts with replication and transcription machineries, creating permanent DSBs. Failure to repair the DSBs leads to cell death, which makes these drugs effective chemotherapeutic agents. Alternatively, cells can repair the trapped Top2cc. This is not always an error-free process and can result in DNA mutations. The mutagenic consequences of stabilizing the Top2cc intermediate is the focus of this thesis.

## **1.2 Topoisomerase Biology**

### **1.2.1 Classification of Topoisomerases**

Topoisomerases are classified by the number of DNA strands cleaved. Type I topoisomerases cleave one strand creating a nick, while type II topoisomerases cleave both strands to generate a DSB. Topoisomerases are further separated into subtypes A, B, or C. The subtypes are used to distinguish between enzyme families that have distinct amino acid sequences, as well as global structures. A summary of the subfamilies of topoisomerases can be found in Table 1. For all topoisomerases, DNA cleavage is associated with the formation of a transient phosphodiester bond between an active tyrosine and one end of the broken DNA. The formation of this DNA-protein intermediate allows for the DNA topology to be modified, and the enzyme is released upon DNA re-ligation.

### **1.2.1.1 Type IA Topoisomerases**

Type IA topoisomerases preferentially bind single-stranded DNA, which is produced when DNA is unwound, and resolve topological structures through a 'strand passage' mechanism (3-5). To do this, the enzyme cleaves a single strand of the DNA, and the active tyrosine forms a phosphotyrosyl link with the 5' end of the break. The formation of the nick allows for a second, intact strand to be passed through the gap. This is followed by resealing of the broken DNA strand. Type IA enzymes can only resolve one supercoil at a time (6). This process does not require ATP, with the only exception being reverse gyrase (7). There are three distinct subfamilies of type IA topoisomerases: bacterial Top IA, bacterial and eukaryotic Top III, and bacterial and archaeal reverse gyrase (8). Each subfamily of type IA topoisomerases has a different function. Bacterial Top IA resolves negative supercoils in DNA (9). Topo III mainly works to resolve single-stranded DNA entanglements that can arise during DNA replication and repair. Finally, reverse gyrase creates positive supercoils in DNA and renatures melted DNA strands through the action of an ATP-dependent superfamily 2 helicase (SF2 helicase) domain. The reverse gyrase is only found in thermophilic bacteria and archaea (9).

### **1.2.1.2 Type IB Topoisomerases**

In contrast to type IA topoisomerases, type IB enzymes form a phosphotyrosyl link with the 3' end of the nicked DNA. They additionally do not use the strand passage mechanism of type IA enzymes to resolve topological

structures, but rather allow one end to rotate around the intact phosphodiester bond on the other strand (10). Type IB enzymes use the friction between the DNA and the enzyme to control the rotation process. The friction also aids in the re-aligning of the broken ends prior to re-ligation (10). Type IB topoisomerases work to resolve positive and negative supercoils without the use of divalent ions or ATP (7). Not all the supercoils are resolved during one cleavage event, but rather the number of supercoils removed per step increases with the torque stored in the DNA and follows an exponential distribution (10). This class of topoisomerases is ubiquitous among eukaryotes and can also be found in some viruses and bacteria (9).

**Table 1: Summary of the Topoisomerase Subfamilies**

Subfamily	Mechanism	Structure	Activity	Representative Enzyme	Organism
Type IA	Strand Passage	5' monomer that creates a single-strand break	Relaxation of (-) supercoils	Bacterial DNA Top I	<i>Escherichia coli</i>
			Decatenation Introduce (+) supercoils	Bacterial DNA Top III Archaeal Reverse Gyrase	<i>E. coli</i> <i>Archaeoglobus fulgidus</i>
			Decatenation, resolve recombination intermediates with helicase Regulates transcription	Eukaryotic DNA Top3 Eukaryotic DNA TOP3A and TOP3B	<i>Saccharomyces cerevisiae</i> <i>Homo sapiens</i>
Type IB	DNA strand rotation	3' monomer or heterodimer that creates single-strand breaks	Relaxation of (+) and (-) supercoils	Poxvirus DNA Top I	<i>Vaccinia Virus</i>
				Trypanosoma DNA Top I	<i>Leishmania donovani</i>
				Eukaryotic DNA Top1/TOP1 Mitochondrial DNA TOP1	<i>S. cerevisiae/H. sapiens</i> <i>H. sapiens</i>
Type IC	DNA strand rotation	3' monomer that creates single strand breaks	Relaxation of (+) and (-) supercoils	Archaeal DNA Top V	<i>Methanopyrus kandleri</i>
Type IIA	Duplex DNA passage	5' homodimer or heterotetramer that creates double strand breaks	Introduction of (-) supercoils	Bacterial DNA gyrase**	<i>E. coli</i>
			Relaxation of (+) supercoils, Decatenation	Bacterial DNA Top IV	<i>E. coli</i>
			Relaxation of (+) and (-) supercoils, Decatenation	Eukaryotic DNA Top2 Eukaryotic DNA TOP2A and TOP2B	<i>S. cerevisiae</i> <i>H. sapiens</i>
Type IIB	Duplex DNA passage	5' homodimer or heterotetramer that creates double strand breaks	Relaxation of (+) and (-) supercoils	Archaeal DNA Top VI	<i>Sulfolobus shibatae</i>
			Decatenation	Plant DNA Top VI	<i>Arabidopsis thaliana</i>
			Weak relaxation/ decatenation	Bacterial DNA Top VIII	<i>Ammonifex degensii</i>

Adapted from (11)

### **1.2.1.3 Type IC Topoisomerases**

Type IC topoisomerases are functionally similar to type IB in that they nick a single strand of DNA to allow for strand rotation to resolve positive and negative supercoils (12, 13). However, the active site of type IC is not structurally similar to type IB enzymes, and these enzymes appear to have a different evolutionary lineage (12, 14). Type IC topoisomerases have only been identified in the archaeal genus *Methanopyrus* (14).

### **1.2.1.4 Type IIA Topoisomerases**

Like type IA topoisomerases, type IIA enzymes utilize a strand passage mechanism to resolve topological structures. However, instead of creating a single-strand break in the DNA, type IIA enzymes cleave both DNA strands to create a transient DSB (15-17). An intact duplex is then passed through the DSB and this process requires ATP. Type IIA enzymes can resolve positive and negative supercoils, as well as untangle intertwined chromosomes and DNA catenanes (18, 19). Type IIA topoisomerases are present in all organisms and can be further categorized into subfamilies based on functional properties. These subfamilies are eukaryotic Top II, bacterial Top IV, and bacterial and archaeal gyrase (9).

The majority of eukaryotic organisms express a single form of Top2. However, vertebrates have two isoforms, TOP2A and TOP2B (20). TOP2A activity is essential for cell viability, and it is primarily active during replication and mitosis (21). TOP2A levels increase in mid-S phase through mitosis and levels

rapidly decrease at the end of mitosis (22, 23). In contrast, TOP2B levels remain constant throughout the cell cycle and this enzyme plays a role in regulating gene expression (23-26).

#### **1.2.1.5 Type IIB Topoisomerases**

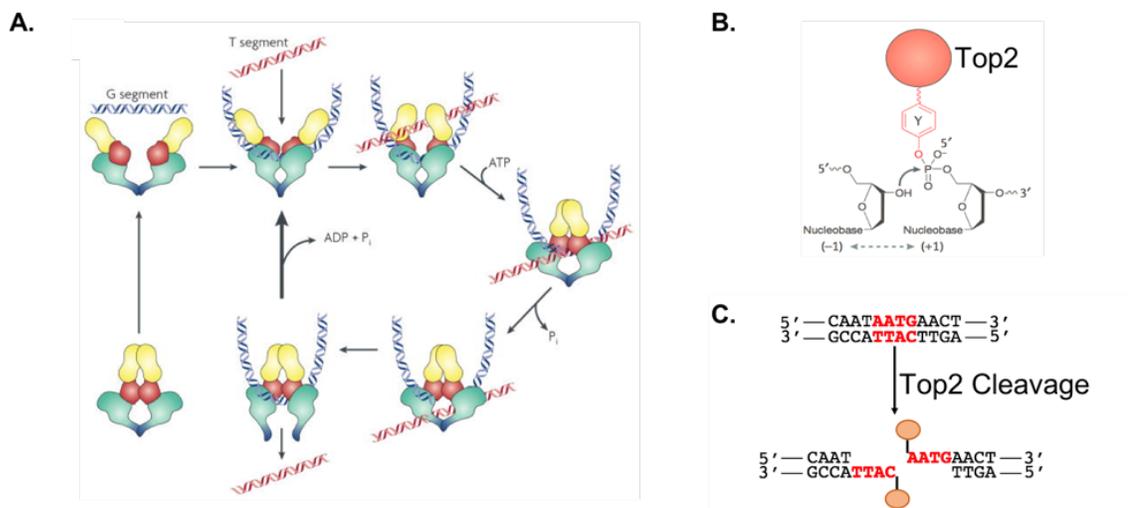
Type IIB topoisomerases unlink tangled DNA duplexes and relax positive and negative supercoils through a strand passage mechanism, similar to Type IIA enzymes (27). However, the relative arrangements of the ATPase and DNA-cleavage domains, as well as the primary sequence, differ greatly from the type IIA enzymes (9). This class of topoisomerases is found in some archaea, plants, bacteria, protists, and algal species (28, 29).

#### **1.2.2 Topoisomerase II Catalytic Cycle**

During its catalytic cycle, eukaryotic Top2 acts as a homodimer and follows the general principles of a two-gate model (30). The Top2 catalytic cycle is summarized in Fig. 1. In the presence of  $Mg^{+2}$ , each Top2 subunit cleaves the Gate- or G-segment of double-stranded DNA 4-basepairs (bp) apart, and the active site of each subunit forms a covalent phosphotyrosyl bond with the 5' end of the DNA, creating a DSB. This covalent DNA-protein intermediate is referred to as a Top2cc. Association of the enzyme with ATP results in the closure the N-terminus gate around a second DNA duplex, termed the Transfer- or T-segment. Hydrolysis of the ATP allows for the T-segment to be propelled through the DSB, and then the G-segment is re-sealed. Finally, the T-segment is released through the C-terminal gate of the homodimer. Hydrolysis of a second ATP re-opens the

N-terminus gate in order to reset the enzyme for either another round of strand passage or release from the DNA (30).

Type II topoisomerases generate two coordinated nicks on the opposite strands of the double helix as opposed to one unified double-stranded DNA break. Evidence for this comes from findings that levels of scission on the two strands of a cleavage site are not equal and that the enzyme re-ligates the two strands of cleaved molecules in an independent fashion (31, 32). Once Top2 cleaves the first strand, it cuts the second strand at a rate that is ~10-fold faster (33, 34). The mechanism that underlies the enhanced cleavage is not known.



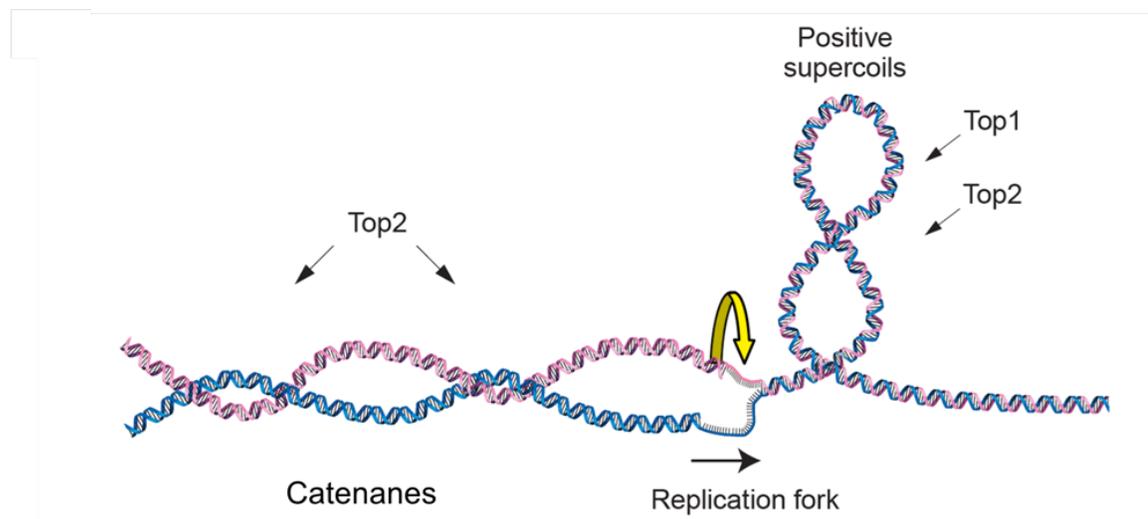
**Figure 1: Top2 Catalytic Cycle**

A) Top2 acts as a homodimer and its catalytic cycle involves two duplex DNA molecules. The first duplex called the G-segment is cleaved by Top2 to create a DSB in the presence of  $Mg^{+2}$ . Top2 then engages the second, intact T-segment DNA duplex. ATP hydrolysis allows for the enzyme to form a closed clamp. The T-segment can then pass through the DSB and subsequently exit through the C-terminus of the Top2 homodimer. A second ATP hydrolysis event allows for the clamp to re-open for the release of the G-segment after the DNA has been re-ligated. Alternatively, the G-segment can remain engaged for an additional round through the catalytic cycle. B) During the cleavage event, each subunit of Top2

creates a 5' phosphotyrosyl bond with one end of the break. C) The Top2 subunits cleave complementary strands 4-bp apart to create 5' overhangs. Adapted from (1, 35).

### 1.2.3 Topoisomerases in Replication and Chromosome Condensation

Unwinding of the DNA duplex during replication results in topological changes in the DNA. These topological changes include the over-winding of unreplicated DNA ahead of the replication fork, creating positive supercoils, as well as the intertwining of newly synthesized DNA behind the replication fork to form catenanes (Fig. 2) (36, 37). If these structures are left unresolved, they can result in replication fork collapse, DSB formation, or missegregation of sister chromatids. These structures can be resolved by type IB and type IIA topoisomerases (38-40). In eukaryotes, Top1 and Top2 resolve replication-induced supercoils, while Top2 resolves catenanes (Fig. 2).



**Figure 2: The Roles of Top1 and Top2 during DNA Replication**

As the replication fork progresses, the parental DNA duplex is unwound, resulting in the formation of positive supercoils ahead of the replication machinery. Positive supercoils can be resolved by either Top1 or Top2. Behind the replication machinery, newly synthesized sister chromatids become intertwined to create catenanes. These structures are exclusively resolved by Top2. Adapted from (41).

Our initial understanding of the role of Top2 in replication was determined using yeast as a model organism. *TOP2* is an essential gene in yeast, so temperature-sensitive *top2* alleles (*top2-ts*) were isolated to study the functions of Top2 *in vivo* (42, 43). At non-permissive temperatures, *top2-ts* mutants fully replicate their DNA, but cells fail to properly segregate sister chromatids during mitosis (42, 44). Furthermore, there is a 10-fold increase in nondisjunction in cells lacking Top2 activity, indicating that the primary cause of the high lethality in *top2-ts* mutants is aneuploidy (45). These findings indicated that Top2 is required for successful segregation of sister chromatids and completion of mitosis.

The majority of sister-chromatid decatenation is complete before cells enter into mitosis. However, specific chromosomal regions such as centromeres, telomeres, and rDNA remain catenated through anaphase (46-49). This keeps sister chromatids in close proximity to one another to aid in homologous recombination (HR) (50). Top2 is also involved in maintaining the catenated state of DNA in order to assist in chromosome condensation and cohesion prior to segregation. Top2 is required for condensation and works together with condensin to maintain the state of chromosome condensation (51, 52). This cooperative action between Top2 and condensin is called the Goldilocks

equilibrium (52), in which condensin generates positive supercoils in an ATP-dependent manner and Top2 acts on these positive supercoils to maintain the appropriate tension required for chromosome condensation (53, 54). Once chromosome condensation is complete, Top2 binding at centromeres is enhanced (55-57). Cohesin at the centromeres prevents Top2 activity until cohesin is cleaved by separase (46, 58, 59). This allows for the timely decatenation of the centromeres to allow for proper chromosome segregation. Loss of Top2 can lead to the formation of inter-chromosomal linkages known as anaphase bridges, which promote DNA damage if left unresolved (46, 58-60).

Unlike Top2, Top1 is not an essential gene in yeast (61, 62). Cells lacking *TOP1* have a temporary delay in DNA elongation in early S-phase, but otherwise grow normally (63). In the absence of both Top1 and Top2 activity, positive supercoils build up in front of the replication machinery, and further elongation of the newly synthesized DNA is prohibited (63, 64). However, these supercoils are not observed if either Top1 or Top2 is active, indicating that both enzymes can act as a DNA swivel and resolve positive supercoils (63, 64). The only known function of Top2 that Top1 cannot perform is the decatenation of intact DNA molecules, which is required for the segregation of sister chromatids in mitosis. Further studies attempting to separate the functions of Top1 and Top2 found that Top1 normally acts in front of the replication fork to relieve positive supercoils, while Top2 resolves precatenanes that form in the newly replicated DNA (40, 65-67).

Although Top2 is best known for its essential role in resolving catenanes to ensure proper segregation of sister chromatids, more specialized roles of Top2 during replication have been discovered since the initial experiments in yeast. To prevent the accrual of topological stress, Top1 and Top2 have been shown to localize to replication origins during the initiation process (68-71). There is no evidence in yeast or mammalian cells that Top2 plays a direct role in replication initiation (67, 72). However, studies performed in *Xenopus* extracts suggest a pre-replicative role of TOP2A in which it negatively regulates loading of the MCM2-7 helicase to establish sequential firing of origin clusters (66).

Topoisomerases may need to remain inactive at replication origins in order to maintain the negative supercoiling necessary for replication initiation (35).

In eukaryotes, replication forks move bidirectionally until they fuse with forks from adjacent origins (73). As the fork approaches the termination zone, Top1 is no longer able to resolve positive supercoils accumulating ahead of the fork and progression relies on Top2 activity (74). Top2 alone is insufficient for complete replication termination, but Top2 does facilitate fork fusion and resolves topological structures at chromosomal termination sites (TERs) (75-78).

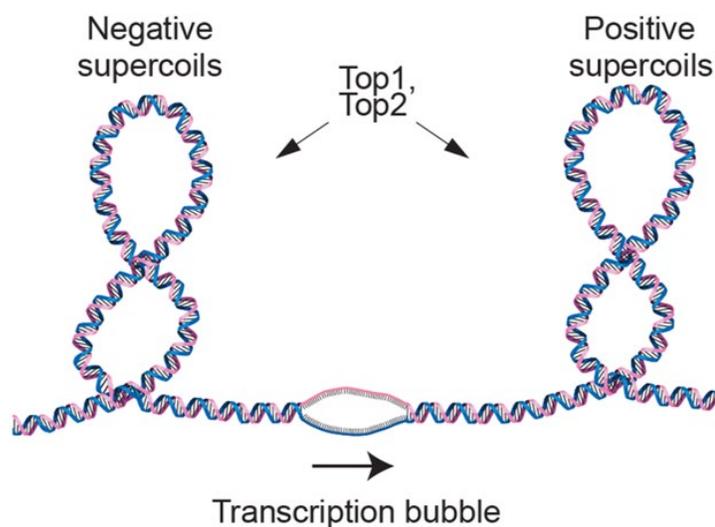
Catalytically dead Top2 has been shown to interfere with replication termination, and in the absence of Top2, TERs accumulate chromosomal breaks and rearrangements in yeast (72, 75). Although Top2 is not generally required for replication termination, it is crucial for termination in regions prone to replication stress (79). For example, when replication forks stall in telomeric regions, Top2

mediated cleavage and NHEJ-dependent re-ligation aid in the removal of the replication forks and the re-starting of replication. To achieve this, Top2 cleaves the DNA to release the stalled replication fork and creates a *t*-circle-tail. (80). This is followed by DNA ligation and the initiation of a new replication fork to allow for efficient replication of telomeric DNA (80). Top2 activity during replication termination and in regions prone to replication stress preserves genome integrity by preventing chromosome breakage during cell division.

#### **1.2.4 The Role of Topoisomerases in Transcription**

The twin-supercoiled-domain model describes the DNA supercoiling associated with movement of the transcription machinery (Fig. 3). The domains consist of positive and negative supercoils that form in front of and behind the advancing transcription machinery, respectively. This model is supported by *in vitro* and *in vivo* evidence and topoisomerase activity is important for managing transcription-associated supercoils (63, 81-83). Early studies in *S. cerevisiae* revealed that the loss of either Top1 or Top2 does not affect transcription levels, but the loss of both enzymes leads to a significant decrease in ribosomal RNA (rRNA) and messenger RNA (mRNA) synthesis (63). Genome-wide analysis further revealed a reduction of global RNA synthesis in the absence of Top1 and Top2 (84). Topoisomerases have been shown to play a role in transcription initiation and elongation, but each enzyme's exact role can depend on the species or even the particular genes under investigation.

Topoisomerases are critical for regulated, gene-specific transcription initiation, as they help maintain a delicate balance in the superhelicity of DNA within promoter regions. Top1 and Top2 have been shown to localize to promoter regions of highly transcribed genes, and evidence of topoisomerase activity at promoter regions has been documented in several model organisms (84-86). The effects of this activity are not straightforward. In *S. cerevisiae*, Top1 inhibits transcription initiation of stress-inducible genes located in subtelomeric regions (87). Conversely, requirements for topoisomerase activity in the activation of highly regulated promoters, such as the *GAL* genes, have been reported in budding yeast (88-90). In mammalian cells, Top1 can act as a repressor of basal transcription and an enhancer of active transcription from TATA-containing promoters (91).



**Figure 3: The Roles of Top1 and Top2 During Transcription**

The unwinding of DNA during transcription creates two domains of supercoiling. Positive supercoils accumulate ahead of the transcription bubble, while negative supercoils form behind the machinery. Both positive and negative supercoils can be resolved by Top1 or Top2. Adapted from (41).

In line with roles of topoisomerases during transcription initiation, Top1 and Top2 have been shown to redundantly aid in the recruitment of RNA polymerase II (RNAP II) to actively transcribed genes in both *S. cerevisiae* and *Schizosaccharomyces pombe* (84-86, 90). This has also been observed in higher eukaryotes (92, 93). In human cells, the activation of Top1 is coupled with the phosphorylation of the carboxyl-terminal domain (CTD) of RNAP II by BRD4 to regulate RNAP II-pause-release (94). This coordinated action helps overcome the torsional stress opposing transcription and allows elongation to begin. In addition to Top1 activity in the initiation process, Top1 activity is not constitutive but can be regulated by different factors (94).

Early studies showed that TOP2B is required for the expression of developmentally regulated genes in neurons of mammalian cells (24). More specifically, DSBs induced by TOP2B are required for regulated gene transcription of a number of genes, including the activation of early response genes in neurons following stimulation (25, 26). Furthermore, single nucleotide mapping of DSBs in human cells found an enrichment around transcription start sites of highly transcribed genes that contain RNAP II pause sites (95). Inhibition of TOP1 and TOP2 by camptothecin or etoposide, respectively, resulted in an increase in DSBs at pause sites, indicating that the topoisomerases are involved in creating these DSBs and contribute to transcription activation. TOP1 and

TOP2 chromatin immunoprecipitation sequencing (ChIP-seq) data further implicated TOP2B and to a lesser extent TOP1, in DSB formation at the pause site and loss of TOP2B diminishes DSBs at pause sites (95). Previous studies support this finding, showing that TOP2B is recruited to pause-sites, and loss of TOP2B catalytic activity results in the retention of RNAP II at transcription start sites (96). The coupling of transcription initiation and topoisomerase activity emphasizes the importance of supercoiling dynamics in this process.

The task of transcription initiation is complicated by the fact that DNA is organized into chromatin, which is composed of an array of histone-DNA complexes called nucleosomes. The positions of the nucleosomes play a critical role in gene expression. When tightly packed into chromatin, DNA is not accessible to RNA polymerases or other transcription factors. Several studies have investigated the role of topoisomerases in nucleosome remodeling. In *S. pombe*, Top1 is required for nucleosome disassembly, which allows for recruitment of RNAP II and efficient transcription (86). However, other studies in *S. cerevisiae* and higher eukaryotes have found that the loss of either Top1 or Top2 does not affect nucleosome density, suggesting that these enzymes are not required for nucleosome disassembly (51, 84, 97). Although topoisomerases may not be directly involved in nucleosome remodeling, studies investigating the mammalian TOP2B have shown that it is part of a signaling cascade necessary for transcription initiation. In mammalian cells, DSB formation by TOP2B in the

promoter region of an inducible gene can signal for the recruitment of downstream histone regulators (25).

As transcription proceeds through elongation, positive supercoils accumulate ahead of the advancing machinery. *In vitro*, human TOP1, human TOP2A, and yeast Top2 are required to synthesize transcripts 200 nucleotides (nt) and longer (92, 98). This is likely due to the acquisition of high levels of supercoils that block elongation and highlights the importance of topoisomerase activity during this stage of transcription, as many RNAP II transcripts exceed 200 nt in length. *In vivo* studies further support these results. Studies performed in budding yeast found that the loss of Top2 results in an abrupt loss in transcripts greater than 3 kilobases (kb) in length, irrespective of function (99). The loss of transcripts was caused by elongation blockage and was not observed in Top1 mutants. These results indicated that Top1 and Top2 activity is not interchangeable during transcription elongation. Furthermore, there is a critical transcript length beyond which Top1 can no longer efficiently resolve positive supercoils and Top2 becomes responsible for resolving these structures (99). This switch is likely due to conformational changes in the DNA. Previous studies have shown that nucleosomal DNA can act as a buffer for torsional stress that accumulates during elongation, but once the buffering capacity is reached, positive supercoils begin to form (100-102). The formation of positive supercoils in the DNA obstructs the strand-rotation mechanism of Top1 but facilitates the strand-passage mechanism of Top2 (103). A similar separation of Top1 and

Top2 activity is seen in budding yeast during transcription of rDNA, in which Top1 and Top2 predominantly relax negative and positive supercoils, respectively (104).

In support of these findings, high-resolution mapping of supercoil dynamics in Burkitt's lymphoma cells found that high-output promoters sustain high levels of supercoiling that is counterbalanced by topoisomerase recruitment (105). TOP1 adequately handles torsional strain generated at low-output promoters but supercoils generated at high-output promoters required TOP2 activity. A length-dependent defect in the expression of long genes (>2 kb) in the absence of TOP1 and TOP2 was also observed in mouse and human neurons (106). Mapping of RNAP II density genome-wide in neurons showed that this length-dependent defect was due to impaired transcription elongation. The defects seen in the transcription of long genes has led to speculation that mutations in topoisomerases may contribute to autism and other neurodevelopmental disorders.

### **1.3 Trapping the Top2 Cleavage Complex**

The catalytic cycle of Top2 is normally an error-free process, but conditions that significantly increase either the concentration or lifetime of these breaks can lead to activation of the DNA damage response and cell-cycle arrest, senescence, or apoptosis (107-109). Other deleterious side effects can occur, including insertions, deletions, and chromosomal aberrations (110-113). Drugs that inhibit Top2 have been the most common and effective anticancer and

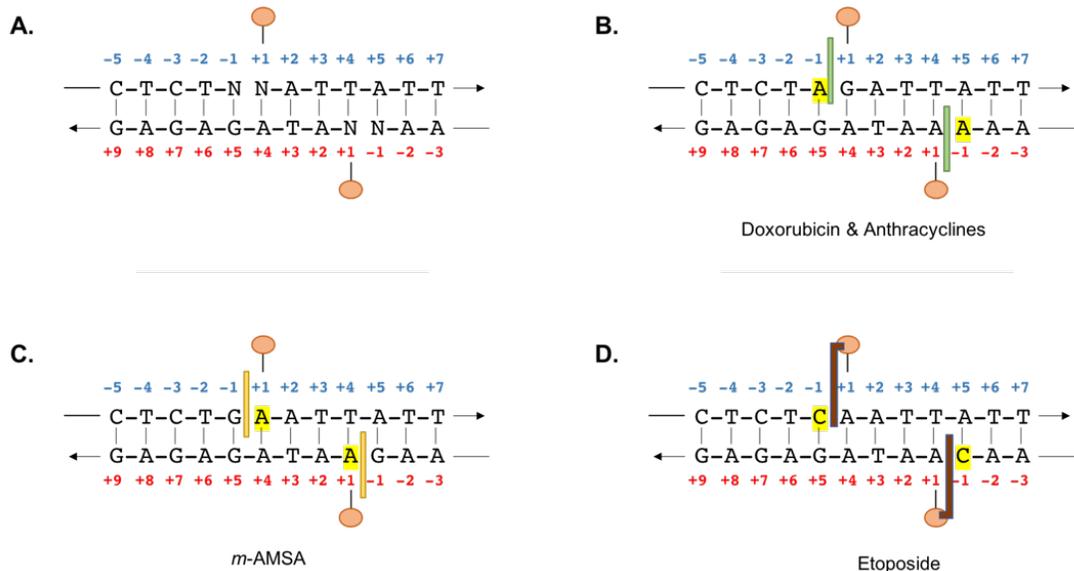
antibacterial drugs (1, 2). There are two classes of small molecules that target Top2: poisons and catalytic inhibitors. Top2 poisons increase the steady-state levels of the protein-DNA intermediates, either by increasing Top2 cleavage or stabilizing the Top2cc on DNA (114). Collisions between the trapped Top2cc and the replication or transcription machinery creates permanent DSBs. Such DSBs destabilize the genome and lead to programmed cell death, making Top2 poisons an effective chemotherapeutic agent. By contrast, Top2 catalytic inhibitors eliminate the essential enzymatic activity of Top2, which can also lead to cell death. Catalytic inhibitors do not increase levels of Top2cc and are generally not used as anticancer agents (114). This section will first focus on Top2 poisons commonly used as chemotherapeutic agents, and then will discuss endogenous Top2 poisons.

### **1.3.1 Top2 Poisons**

Top2 poisons are a diverse group of drugs that differ in chemical composition, potency, sequence selectivity, and their ability to increase the steady-state level of Top2cc. This class of drugs can be divided into two broad categories: intercalating and non-intercalating agents. Intercalating compounds include doxorubicin and other anthracyclines, mitoxantrone, and mAMSA (1).

To quickly identify bases relative to the Top2 cleavage site, a numbering scheme has been developed. Top2 forms a phosphotyrosyl bond with the +1 base, and all subsequent bases in the 3' direction are numbered with increasing values. Conversely, the bases in the 5' direction are numbered -1, -2, etc. Top2-

targeting intercalators insert between the between the -1 and +1 bases and disrupt the geometry required for Top2-mediated re-ligation (Fig. 4A) (1). Doxorubicin and other anthracyclines are most effective when adenine is present at the -1 position, while *m*-AMSA prefers an adenine at the +1 position (Fig. 4B and 4C) (115, 116). Upon removal of the drug, reversal of the Top2cc is slow, which contributes to the potency of anthracyclines as Top2 poisons. Intercalating agents that target Top2, such as doxorubicin, also have a variety of notable side effects, including bone marrow suppression and cardiotoxicity. The cardiotoxicity has been attributed to the generation of reactive oxygen species, as well as TOP2B targeting that leads to mitochondrial damage (117, 118). Mitoxantrone was developed as a synthetic analog of anthracyclines, and its reduced potential to undergo redox reactions limits its cardiotoxicity (2). Mitoxantrone is a first-line therapy for pediatric and adult acute leukemia and a second-line therapy for breast cancer, prostate cancer, and hematological malignancies (2).



**Figure 4: Top2 poisons and their sequence preference.**

A) Each Top2 subunit creates a 5' phosphotyrosyl bond with a single strand of the duplex DNA. Top2 forms a phosphotyrosyl bond with the +1 base and all subsequent bases in the 3' direction is numbered with increasing values. The base immediately 5' of the phosphotyrosyl bond is numbered -1, and all subsequent bases in the 5' direction are numbered with decreasing values. B) Doxorubicin (green) and other anthracyclines intercalate between the -1 and +1 when an adenine is at the -1 (highlighted in yellow). C) *m*-AMSA (orange) intercalates between the -1 and +1 positions when there is an adenine at the +1 (highlighted in yellow). D) Etoposide (maroon) is a non-intercalating agent that preferentially interacts with Top2 to prevent the re-ligation reaction when a cytosine is located at the -1 position (yellow highlight). The arrows indicate the 3' end of the DNA.

The most well-known example of a non-intercalating Top2 poison is etoposide, commonly used as a first-line therapy for a variety of malignancies, including leukemias, lymphomas, and several solid tumors (2). High-resolution imaging of the ternary complex between Top2, DNA, and etoposide revealed that interactions with specific amino acids of the enzyme are critical for etoposide to enter the Top2-DNA complex and inhibit re-ligation (119, 120). Like the

intercalating agents, etoposide displays sequence-specific trapping of the Top2cc and prefers a cytosine at the -1 position (Fig. 4D) (2). The Top2cc trapped by etoposide frequently uncouples the re-ligation reactions, creating single-strand breaks (SSBs) rather than DSBs in DNA (121, 122). This indicates that etoposide traps the Top2 subunits asymmetrically. Etoposide produces a higher concentration of Top2cc when compared to the intercalating agents, and the trapped Top2cc are readily reversible upon drug washout (116, 123).

Effective etoposide treatment results in the accumulation of irreparable DNA damage, culminating in cell death by apoptosis. However, if the concentration of Top2-dependent DSBs is too low, cells may be able to evade death but can acquire other mutagenic events, such as translocations. Treatment with etoposide is associated with a balanced translocation between the mixed lineage leukemia (*MLL*) gene on chromosome 11q23 and several partners. This translocation is often associated with the initiation of specific types of acute leukemias, such as acute lymphocytic leukemia (ALL) and acute myeloid leukemia (AML) (124, 125). Over 50% of patients with therapy-induced leukemia display translocations within an 8.3 kb breakpoint cluster region in the *MLL* gene, and these DNA breakpoints are in close proximity to TOP2 cleavage sites (124-126). A previous study suggested that the balanced translocations observed with the *MLL* gene are dependent on TOP2B and not TOP2A (127). A model was proposed in which DSBs induced by TOP2B in pairs of genes undergoing transcription within a common transcription factory are stabilized by etoposide,

with the close proximity of the genes promoting illegitimate end joining and translocation. Although it is widely accepted that TOP2-mediated cleavage is involved in the development of *MLL* translocations, the mechanism for these translocations is still controversial. An alternative mechanism indirectly involves TOP2-induced cleavage, but it is ultimately apoptotic nucleases that are implicated in the breaks that lead to the translocations. Evidence for a major apoptotic cleavage site in the breakpoint cluster region of the *MLL* gene supports this model (128-130). Both mechanisms for the formation of the balanced *MLL* translocations require the non-homologous end joining (NHEJ) DSB repair pathway (126, 128). Both mechanisms may be relevant to the formation of the translocations as this is a complex process.

TOP2 poisons are effective chemotherapeutic agents, but they have two significant limitations. The first is that those currently used in the clinic equally target TOP2A and TOP2B, and many of the harmful side-effects have been shown to be caused by inhibition of TOP2B (117, 118, 127). Secondly, available TOP2 poisons do not specifically target tumor cells. The ideal poison would specifically target only cancer cells and TOP2A. One compound, F14512, has these characteristics (131). F14512 is a polyamine-vectorized Top2-inhibitor that was derived from etoposide and is selectively taken up by tumor cells via the polyamine transport system (PTS), which is upregulated in cancer cells. Additionally, F14512 increases the stability of the Top2 $\alpha$ cc without affecting TOP2B, and is more potent compared to etoposide (131-133). Although F14512

had antitumor activity in mouse and canine lymphoma models and phase I trials were completed in human hematologic pathologies (134-136), phase II trials ended in 2017 due to the discontinuation F14512 development in the indication for AML. It will be interesting to see if this research is continued or other drugs with similar characteristics are developed in the future.

### **1.3.2 Endogenous Stabilization of the Top2 Cleavage Complex**

Genomes are constantly subject to a variety of endogenous and environmental insults that can generate numerous forms of DNA damage. This damage includes abasic sites, mismatches, alkylated bases, and oxidized lesions. In addition to interfering with fundamental processes such as DNA replication and transcription, some lesions also act as Top2 poisons. The search for endogenous Top2 poisons began with the observation that a majority of primary infant leukemias had a chromosomal translocation involving the *MLL* gene on chromosome 11 (137-140). As mentioned previously, this translocation had previously been observed in adult patients that developed secondary therapy-related leukemias after being treated with etoposide-based regimens (137, 138, 141). *In vitro* studies showed that the 11q23 translocation breakpoints are Top2 cleavage sites (124) and, in conjunction with the nonrandom association of etoposide treatment with these translocations, suggested that Top2 may be involved in initiating these events. Furthermore, the observation of the 11q23 translocation in drug-naïve primary infant leukemias suggested the possibility of endogenous Top2 poisons.

In terms of endogenous Top2 poisons, DNA lesions that distort DNA and disrupt base pairing, such as abasic sites, base mismatches, and exocyclic DNA adducts (1,N<sup>6</sup>-ethenoadenine), increase Top2 cleavage (142-146). The effects of DNA lesions on Top2 cleavage are dependent on their position within the Top2 cleavage site. For example, Top2 cleavage is stimulated when an abasic site is located within the 4 bp between the scissile bonds (positions +1, +2, +3, and +4), but if the lesion is directly outside this region, it has an inhibitory effect on cleavage (143, 144). Similarly, base mismatches that occur at the +1 and +2 positions enhance Top2 cleavage, while mismatches at the -4, -3, -2, and -1 positions have an inhibitory effect on cleavage (147). These observations indicate that proper base pairing at the 3' terminus is required for Top2 cleavage. The structural variation of an apurinic site within a Top2 cleavage site has been examined by NMR to gain insight into how these sites increase Top2 cleavage (148). An apurinic site at the +2 position of a Top2 cleavage site leads to loss of base stacking at the lesion and collapse of the major groove, which reduces the distance between the two scissile phosphodiester bonds. Additionally, the apurinic lesion induced a bend centered around the Top2 cleavage site, and the base immediately opposite the lesion was extrahelical and relocated to the minor groove. All these structural variations have the potential to influence the interactions between Top2 and DNA and affect Top2 cleavage efficiency (148).

Nicks in the DNA are the simplest type of lesion. They contain a clean 3'-hydroxyl group that can undergo ligation or act as a primer for new DNA

synthesis. DNA nicks can be generated either directly by DNA damage or as intermediates in essential nuclear processes and DNA repair pathways (149). Tens of thousands of nicks form in human cells each day. A nick at one scissile bond in DNA has also been shown to greatly enhance Top2 cleavage at the scissile bond on the opposite strand, indicating that nicks are Top2 poisons (150). The enhanced cleavage is likely due to enhanced helix flexibility associated with a nick, which allows DNA to achieve the bent transition state required for efficient cleavage. The presence of a nick does not affect Top2 re-ligation (150).

Like DNA nicks, misincorporated ribonucleotides (ribos) are prominent in the genome. It is estimated that 10,000 – 15,000 ribos are inserted into the 12 Mb yeast genome during each round of replication (151). A similar level of ribo incorporation has been estimated in cultured mouse cells (152, 153). The presence of a single ribo at the scissile phosphodiester bond on one strand of DNA enhances Top2 cleavage *in vitro*, indicating that misincorporated ribos are Top2 poisons (154). Similar to other endogenous Top2 poisons, cleavage at a ribo does not affect the re-ligation step of the catalytic cycle. Additionally, ribos located between the Top2 cleavage sites stimulate cleavage, while ribos outside the cleavage sites have an inhibitory effect (155). Top2 cannot cleave RNA/DNA hybrids, however, in which one strand is entirely composed of RNA (155). All of the endogenous Top2 poisons identified to date enhance the Top2 cleavage rather than inhibit the re-ligation stage of the catalytic cycle.

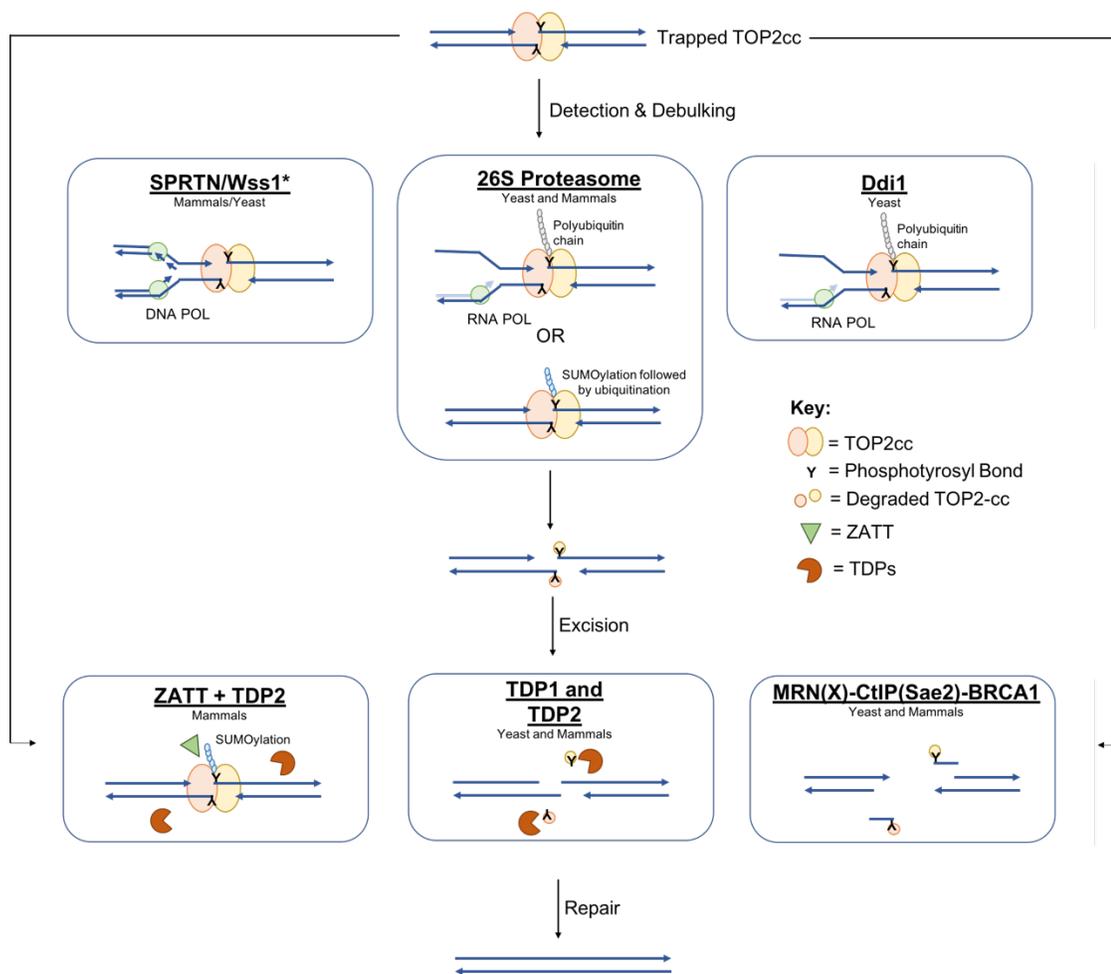
## **1.4 Removal of the Stabilized TOP2 Cleavage Complex**

Topoisomerase activity is critical for a number of cellular processes, including replication and transcription. As mentioned in the previous section, failure of topoisomerases to complete their catalytic cycle can result in their stabilization on DNA. The frequency and cytotoxicity of these events have led to the evolution of multiple pathways that can degrade and then excise the TOP2cc prior to DNA repair by canonical pathways. The 26S proteasome (156) and proteases, such as mammalian SPRTN (157, 158) and yeast Ddi1 (159), debulk the Top2/TOP2cc. This is generally followed by either hydrolysis of the phosphotyrosyl bond by tyrosyl DNA phosphodiesterases TDP1 and TDP2, or nuclease-dependent repair by MRE11-dependent pathways (160). It should be noted that proteolysis-dependent repair is important for the removal of additional DNA-protein crosslinks (DPCs) such as formaldehyde-induced crosslinks (161). Fig. 5 summarizes the various pathways that are responsible for removing stabilized TOP2cc.

### **1.4.1 Ubiquitination and SUMOylation of TOP2 Cleavage Complex for Proteasomal Degradation**

Stabilized TOP2cc can be degraded by the 26S proteasome, a large protease complex that degrades intracellular proteins (162). Proteins are targeted for proteasomal degradation by the addition of polyubiquitin chains; this pathway is referred to as the ubiquitin-protease pathway (UPP). The degradation of TOP2cc by UPP was discovered through observation that inhibition of the

proteasome in solid tumors attenuated resistance to etoposide treatment by blocking the degradation of TOP2cc (156). Further studies performed in HeLa and other cancer cell lines found that the proteasome preferentially degrades TOP2B, and inhibition of transcription, but not replication, blocks TOP2B degradation (163). This suggests that obstruction of the transcription machinery by trapped TOP2Bcc signals for ubiquitination and subsequent degradation (164). Even though TOP2A is targeted for proteasomal degradation to a lesser extent than TOP2B, TOP2A degradation is also dependent on transcription (165). To further highlight the importance of transcription in UPP-dependent TOP2cc degradation, a study that employed a variation of END-seq to distinguish between unprocessed and processed TOP2cc found that transcription promotes UPP processing of TOP2cc (166). Combining proteasomal inhibitors and TOP2-targeting drugs has promising clinical implications as this increases the likelihood of irreparable DNA damage and tumor cell death (156, 162, 167, 168).



**Figure 5: Removal of the stabilized TOP2cc**

Once TOP2 becomes trapped on DNA there are multiple degradation and removal pathways present in yeast and mammalian cells. Degradation by the metalloprotease SPRTN is dependent on replication. Wss1(\*) is a functional homolog in yeast that may play a role in the Top2cc removal. Degradation by the 26S proteasome requires either polyubiquitination of the TOP2cc upon collision with the transcription machinery or SUMOylation of persistent TOP2cc, which is subsequently followed by polyubiquitination. The aspartic protease Ddi1 can also degrade polyubiquitinated Top2 in a transcription-dependent manner. Once the TOP2cc is debulked, it can be removed from the DNA by either TDP1 or TDP2 that directly cleaves the TOP2 phosphodiester bond or MRN(X)-CtIP(Sae2)-BRCA1 that releases the TOP2cc as part of a short oligonucleotide. ZATT, a SUMO ligase, can mediate the direct removal of the TOP2cc using TDP2 without degradation of the protein. Yeast cells lack ZATT. Once TOP2 is removed, the

DSB can be repaired by canonical pathways including NHEJ and HR. The arrows indicate the 3' ends of the DNA.

Small Ubiquitin-like MOdifiers (SUMOs) are ubiquitin-like proteins that regulate a number of cellular processes, including subcellular transport, protein-protein interactions, and DNA damage response pathways (162). Like ubiquitination, SUMOylation of human TOP2cc was observed in response to treatment with teniposide, an etoposide analog (169). SUMOylation of a stabilized TOP2cc is followed by ubiquitination and subsequent proteasomal degradation (162). This process is independent of transcription and replication, which leads to the current model in which prolonged chromatin occupancy by TOP2 engages the SUMO system for repair (162).

#### **1.4.2 Protease-dependent Removal of TOP2 Cleavage Complex**

The first non-proteasomal metalloprotease to be identified was Wss1 (Weak suppressor of *smt3*) in yeast. Wss1 cleaves Top1 directly in a DNA-dependent manner, and yeast strains lacking Wss1 accumulate Top1cc and gross chromosomal rearrangements (170). Wss1 targets SUMOylated substrates in yeast for degradation (171, 172). Although Wss1 removes trapped Top1cc, there is no evidence that it can degrade Top2cc in yeast.

SPRTN (SprT-like domain at the N-terminus) is a replication-coupled protease that degrades DPCs in higher eukaryotes, and it is a functional homolog to Wss1 (157, 158, 173-177). SPRTN can degrade TOP1 and TOP2 in a DNA-dependent manner, and cells that are deficient in SPRTN are hypersensitive to

camptothecin, a TOP1 inhibitor, and etoposide (157, 158). SPRTN degrades TOP1cc, TOP2cc and other DPCs to a short peptide, facilitating bypass by the translesion synthesis polymerase complex REV1- Pol $\zeta$  (174).

The protease activity of SPRTN is tightly regulated. Monoubiquitinated SPRTN is unable to degrade DPCs (177) and its deubiquitination promotes DNA binding and activation of degradation activity. Evidence suggests that SPRTN activity is uniquely activated by binding to single-stranded DNA (177). Additionally, SPRTN has autocleavage activity that may be involved in regulating and preventing unnecessary degradation activity (157, 158, 177). None of these regulatory mechanisms explain how SPRTN is specifically recruited to DPCs during DNA replication. Proteasomal degradation of DPCs during replication is promoted by the ubiquitination of the DPC (178). In contrast, SPRTN degradation does not require ubiquitination of the DPC but rather the extension of a nascent strand to the DPC (178). These observations indicate that SPRTN and proteasome activity are coupled to DNA replication by distinct mechanisms. Mutations in SPRTN that compromise its protease activity cause Ruijs-Aalfs syndrome, which is a human autosomal recessive disorder characterized by genomic instability, progeria, and early-onset hepatocellular carcinoma (179).

A novel aspartate protease has recently been identified in yeast called Ddi1 (DNA-damage inducible 1) (159, 180). Yeast strains lacking both Wss1 and Ddi1 are hypersensitive to camptothecin, and a *ddi1* single mutant is hypersensitive to etoposide treatment (159). This indicates that Ddi1 is involved

in the removal of trapped Top1 and Top2 in yeast. Ddi1 seems to work in parallel to Wss1 as the additional loss of Wss1 in *ddi1* yeast enhances hypersensitivity to etoposide. The protease activity of Ddi1 appears to be replication-dependent and a recent study reported that Ddi1 also proteolyzes polyubiquitinated substrates (159, 181). This raises the possibility that in yeast, Wss1 targets SUMOylated substrates while Ddi1 is responsible for ubiquitinated DPCs. Further studies are needed to determine if the human homologs of Ddi1, DDI1 and DDI2, similarly act as proteases for TOP1cc and/or TOP2cc.

#### **1.4.3 TDP1 and TDP2**

Tyrosyl-phosphodiesterase 1 and 2 (Tdp1/TDP1; TDP2) directly hydrolyze the covalent, phosphotyrosyl bond between topoisomerases and DNA. Tdp1 was initially identified in *S. cerevisiae* for its ability to hydrolyze phosphotyrosyl bonds at the 3' end of DNA and was later shown to be involved in the repair of stabilized Top1cc (182-184). TDP1 is conserved in eukaryotes, and cells lacking TDP1 are hypersensitive to TOP1 inhibitors (161). In contrast, mammalian-specific TDP2 targets 5'-termini ends and removes stabilized TOP2cc (185, 186). Cells deficient in TDP2 are hypersensitive to TOP2 inhibitors (185-188). There is no known orthologue of TDP2 in yeast (189). However, yeast cells that lack Tdp1 are hypersensitive to Top2 targeting agents (190). Similarly, proliferating human cells deficient in TDP1 are more susceptible to TOP2 poisons, and TOP2Acc accumulate in these cells (191). These results indicate that TDP1 can also hydrolyze 5' phosphotyrosyl linkages and aid in the repair of stabilized TOP2cc.

TDP1 and TDP2 have distinct catalytic cycles for the hydrolysis of phosphotyrosyl bonds. TDP1 acts as a monomer and forms a transient covalent bond with the DNA during the hydrolysis of the phosphodiester bond between the tyrosine residue and DNA 3' phosphate (192, 193). After hydrolysis, the DNA has a 3'-phosphate end that is further processed by polynucleotide kinase phosphatase (PKNP) to generate a 3'-hydroxyl end that can be extended by DNA polymerases or used for ligation (189). A homozygous mutation within the catalytic domain of TDP1 (H493R) causes TDP1 to become trapped on DNA (194). The accumulation of this TDP1-DNA covalent intermediate is associated with the autosomal recessive disorder called SCAN1 (spinocerebellar ataxia with axonal neuropathy), a neurodegenerative syndrome (194). TDP2 catalytic activity requires two divalent metal ions, and TDP2 does not form a transient covalent intermediate (185, 195, 196). Additionally, removal of the TOP2cc generates a 5'-phosphate that can be directly extended by polymerases and used for ligation by NHEJ. Similar to mutations in *TDP1*, homozygous mutations in *TDP2* are associated with SCAR23 (spinocerebellar ataxia autosomal recessive 23), a disease characterized by intellectual disability, seizures, and ataxia (161).

Proteasomal degradation of the TOP2cc is required before removal by TDP2 since the phosphotyrosyl bond is hidden within the cleavage complex (197). However, a recent study identified a SUMO ligase called ZATT (zinc finger protein associated with TDP2 and TOP2; previously known as ZNF451) that can mediate the direct resolution of TOP2cc by TDP2 in a proteasomal-independent

manner (198). ZATT directly binds to and participates in the repair of the TOP2cc in two ways. First, ZATT SUMOylates the TOP2cc, which recruits TDP2. Secondly, ZATT remodels the TOP2cc so that TDP2 can access the phosphotyrosyl bond (198). It is important to note that the TDP2- and ZATT-dependent removal of the TOP2cc was only observed in the presence of proteasomal inhibitors, indicating that proteasomal removal of the TOP2cc is the primary pathway for repair.

#### **1.4.4 MRE11-dependent Repair of TOP2 Cleavage Complex**

MRE11 is an endo- and exonuclease ( $3' \rightarrow 5'$ ) that is a member of the MRN complex, comprised of MRE11, RAD50 and NBS1 (Xrs2 in yeast) in mammalian cells. This complex is responsible for recognizing and repairing DSBs as well as signaling their presence and is a major component of HR. During HR, the MRE11 nuclease activity generates 3' single-stranded tails at the DSB and recruits long-range nucleases, such as EXO1 and DNA2 (199, 200). MRN also removes stabilized TOP2cc. The presence of a protein adduct at the end of a DSB strongly stimulates endonucleolytic cleavage by MRN ~15-20 nucleotides away from the blocked DNA ends (201, 202). The endonucleolytic cleavage is promoted by CtIP in mammalian cells and Sae2 in yeast cells. The cleavage can occur on either DNA strand, promotes  $3' \rightarrow 5'$  resection, and results in the release of one Top2cc subunit as part of a short oligo. A second cleavage event on the complementary strand releases the second protein adduct subunit and creates clean ends that can be repaired by HR or NHEJ (201, 202).

Genetic studies have shown that multiple components are required for MRN removal of TOP2cc. In yeast, chicken, and human cells, CtIP (Sae2 in yeast) activates the endonucleolytic activity of Mre11, and loss of MRE11 or CtIP (Sae2 in yeast) results in the accumulation of TOP2cc and hypersensitivity to TOP2-targeting drugs (203-208). Additionally, BRCA1, which interacts with CtIP, is required for removal of TOP2cc in *Xenopus laevis* oocytes and mammalian cells (205, 206, 209). In the absence of BRCA1, MRE11 is not efficiently recruited to TOP2cc sites, which results in the accumulation of these cleavage complexes. The accumulation of TOP2cc is accompanied by high levels of mitotic chromosomal breaks and loss of viability (208). Importantly, overexpression of TDP2 in cells lacking MRE11 restores genome stability, providing more evidence that functionally redundant pathways are responsible for removing stabilized TOP2 (208). Once TOP2cc is removed, the DNA ends can be repaired by HR or NHEJ. In yeast cells, most Top2-generated DSBs are repaired by HR (204, 210). However, in higher eukaryotes, TOP2-generated breaks are primarily repaired by NHEJ (208, 211, 212).

## **1.5 TOP2 Cleavage Consensus Sequence**

TOP2 activity is required at multiple locations throughout the genome, and its cleavage is not randomly distributed. This indicates that the enzyme can recognize specific characteristics of DNA to identify the correct binding and cleavage locations. Various groups have analyzed sequence selectivity to determine a consensus sequence for TOP2 binding and cleavage. Since the

TOP2cc is a fleeting intermediate, many studies utilized TOP2-targeting drugs to stabilize the TOP2cc and identify the consensus sequence. The DNA cleavage sites seem to be influenced by which class of TOP2-targeting drugs is used, and the sequence requirements identified in each study vary (115, 116, 213). A summary of the cleavage requirements for different TOP2-targeting drugs can be found in Table 2. Defined cleavage sites have been identified in plasmid and genomic DNA, and a degenerate consensus sequence of up to 20 bp has been derived for TOP2 using data from several organisms (115, 116, 213-219). A summary of the consensus sequence divided by species for TOP2 is in Table 3.

**Table 2: Summary of the TOP2 Cleavage Consensus Sequence in the Presence of TOP2-targeting drugs**

Top2 Poison	1st Base Preference		2nd Base Preference	
	Position	Base	Position	Base
Doxorubicin	-1	A	-2	G
Etoposide (VP-16)	-1	C	-	-
Teniposide (VM-26)	-1	C	-	-
m-AMSA	+1	T	-1	T

A recent study that mapped etoposide-induced TOP2 cleavage with strand-specific nucleotide resolution across both the *S. cerevisiae* and human genomes found a preference for cytosine 5' of the TOP2 cleavage site (220). This finding is in line with previous reports (116). Importantly, this study demonstrated the TOP2cc maps can be generated using this technique and can potentially be utilized to map TOP2 activity in the absence of a TOP2 targeting drug.

More recent studies have moved away from identifying a consensus sequence to investigating other factors that may affect TOP2 binding and cleavage. Structural and biochemical studies have demonstrated that DNA bending is critical for cleavage and enzyme-mediated bending precedes cleavage (221). Additionally, single-molecule Förster resonance energy transfer (FRET) was used to examine the fundamental steps preceding and during TOP2 cleavage and re-ligation in the absence of TOP2-targeting drugs (222). It was found that the cleavage site is selected during the DNA bending steps, indicating that TOP2 cleavage does not depend on an intrinsic property of DNA, but is determined by the protein-DNA interactions (222). Further investigation of the requirements for TOP2 activity is critical for a full understanding of this process.

**Table 3: TOP2 Cleavage Consensus Sequence Summary**

	Nucleotide Position													
	-5	-4	-3	-2	-1	+1	+2	+3	+4	+5	+6	+7	+8	+9
<b>Human TOP2A</b>	no A	no T	A, no C	-	C, no A	-	-	-	-	no T	-	T, no G	C, no A	-
<b>Human TOP2B</b>	-	-	no C	G	-	T	-	-	-	no T	-	T, no G	C	-
<b>Drosophila</b>	T	-	A, T	A	C, T	A	T	T	-	A	T	-	-	G
<b>Murine</b>	C	G	no G	no T	no A	A	-	-	-	no T	T	A	C	-

## **1.6 TOP2-dependent Mutagenesis**

### **1.6.1 TOP2 Associated Translocations**

The transient cleavage complex is essential to the TOP2 catalytic cycle, but it also poses a threat to genome stability. If a TOP2cc is stabilized on DNA, it can encounter the replication or transcription machinery, which converts the transient intermediate into a DSB (223). Repair of the DSB is not always an error-free process and can result in different types of mutations. The major mutagenic event that is associated with aberrant TOP2 activity in mammals is translocations. As mentioned previously, one of the most well-studied translocations associated with TOP2 stabilization involves the *MLL* gene on chromosome 11 and several gene partners (124-126). This translocation is often linked with specific types of acute leukemias (124-126). Recently, transcription has been implicated as a driver of TOP2-induced DSBs (166, 220, 224, 225) and TDP2 removal of TOP2cc protects cells against transcription-associated translocations (224). However, errors can still occur, and NHEJ has been found to facilitate the formation of translocation events (225).

### **1.6.2 TOP2 and Secondary Structure-mediated DNA Fragility**

Analyses of the genomic distribution of TOP2-induced DSBs has shown enrichment around transcription start sites and CTCF (CCCTC-Binding factor)-binding motifs, which are important for transcription and chromatin architecture regulation (220). TOP2-induced DSBs were also localized to regions predicted to form highly stable DNA secondary structures in yeast and humans, such as

hairpins, G-quadruplex, and multiple stem loop structures (226). Loss of TOP2 results in a significant reduction of DSBs in these areas, indicating TOP2 plays a direct role in generating secondary structure-mediated DNA fragility (226).

### **1.6.3 Studies of Top2 Dependent Duplications Described in this Thesis**

Improper repair of stabilized Top2cc can lead to genome instability and cell death. However, the only mutagenic event that has previously been associated with TOP2-induced DSBs is translocations (227). This thesis focuses on the discovery of a novel Top2-dependent mutation type, *de novo* duplications. Chapter 2 characterizes a yeast *top2* mutant, *top2-F1025Y,R1128G* (*top2-FY,RG*). After cleavage, Top2-FY,RG becomes stabilized on DNA, even in the absence of etoposide. Using a forward mutation assay, we found that stabilization of Top2-FY,RG results in the formation of NHEJ-dependent *de novo* duplications. Chapter 3 discusses the discovery of a *TOP2* mutant, *top2-K743N* (yeast equivalent: *top2-K720N*), which is found in gastric cancers and cholangiocarcinoma and is also associated with the formation of *de novo* duplications in yeast. Finally, Chapter 4 focuses on future directions of this work.

## **Chapter 2. Trapped topoisomerase II initiates formation of *de novo* duplications via the nonhomologous end-joining pathway in yeast**

Chapter 2 was previously published (228), with the exception of the Rad2 and Wss1 results. This chapter represents a collaborative effort with the lab of Dr. John Nitiss at the University of Illinois at Chicago.

### **2.1 Introduction**

Topoisomerases are enzymes that transiently cut DNA in a highly regulated fashion to carry out topological alterations. Type I topoisomerases are typically monomers that make single-strand breaks in DNA while type II enzymes are homodimers that cleave both DNA strands to create double-strand breaks (DSBs) (9, 229, 230). The ability to cut DNA is absolutely required to change DNA topology and cleavage occurs through formation of a transient, covalent phosphotyrosyl linkage with DNA. Type IA and IB enzymes create 5'- and 3'-phosphotyrosyl links, respectively, while type II enzymes form a 5'-phosphotyrosyl link. The resulting single- or double-strand break alters DNA winding using a swiveling (Type IB) or strand passage mechanism (Type IA and Type II), after which the phosphotyrosyl bond is reversed to restore the phosphodiester backbone of the DNA. While the use of a covalent enzyme-DNA intermediate makes cleavage and rejoining a relatively error-free process, any DNA breakage is potentially dangerous (231). If DNA is cut and the break persists, a DNA damage response is activated that leads to cell cycle arrest,

senescence or apoptosis (107-109). Stabilization of covalent DNA-topoisomerase intermediates has been exploited to identify and develop therapeutic small molecules such as fluoroquinolone antibiotics and a variety of anti-cancer drugs (2, 232). These inhibitory molecules have also been very useful as probes for enzyme mechanism (1, 233, 234).

In addition to small molecules that trap topoisomerases on DNA, a variety of DNA structural alterations can affect cleavage and re-ligation. For type I enzymes, DNA lesions that include abasic sites, nicks, base-base mismatches and base alterations can lead to enhanced DNA cleavage *in vitro* and *in vivo* (235-237). Notably, single ribonucleotides embedded in duplex DNA lead to elevated cleavage by type IB topoisomerases *in vitro* (238) and are associated with a distinctive mutation signature *in vivo* (239). In the yeast *Saccharomyces cerevisiae*, this signature is composed of deletions that remove a single unit from a tandem repeat and reflects sequential cleavage of the same DNA strand by Top1 (240-242). Type II topoisomerases can also be trapped on DNA by structural alterations, although the range of lesions appears more limited than for type I enzymes (243). Lesions that can trap eukaryotic type II topoisomerases include abasic sites (142) and mis-incorporated ribonucleotides (154). The trapping of both subunits of Top2 results in a DSB, while the trapping of only a single subunit leads to a persistent single-strand nick.

The study of cytotoxic and mutagenic mechanisms of elevated topoisomerase cleavage has been facilitated by the identification of mutants that

are proficient for DNA cleavage but defective in re-ligation (244-246). While mutations that result in stabilized cleavage intermediates have been readily obtained in yeast *TOP1*, few examples have been identified for type II topoisomerases. The sole exception is a mutant form of human TOP2 $\alpha$  (Asp48 changed to Asn) identified in a screen for mutations affecting the action of bisdioxopiperazines (247). Similar to yeast *top1* mutants, the human TOP2 $\alpha$ <sup>D48N</sup> mutant exhibits elevated DNA cleavage *in vitro* and cannot be expressed in recombination-defective (*rad52* $\Delta$ ) yeast cells. We describe here a mutant yeast Top2 protein (Top2-F1025Y,R1128G; abbreviated Top2-FY,RG) that similarly is lethal when overexpressed in a *TOP2 rad52* $\Delta$  background and leads to elevated DNA cleavage *in vitro*. The corresponding amino acid changes, however, are in the C-terminal dimerization domain of Top2, thereby implicating this domain in the regulation of DNA cleavage and/or re-ligation by the enzyme. Top2-FY,RG overexpression elevates homologous recombination and spontaneous mutagenesis, and is associated with a distinctive mutation signature (*de novo* duplications) that is dependent on the nonhomologous end-joining (NHEJ) pathway of DSB repair. Similar *de novo* duplications were observed following treatment of WT cells with etoposide, a chemotherapeutic drug that stabilizes the covalent Top2 cleavage intermediate. Finally, we implicate WT Top2 as the source of rare *de novo* duplications observed previously in frameshift reversion assays (248). These data suggest important roles for Top2-dependent

mutagenesis in genome evolution as well as in genetic stability following chemotherapy.

## **2.2 Materials and Methods**

### **2.2.1 Strains and Growth Conditions**

YPD (1% yeast extract, 2% Bacto-peptone, 2% dextrose, 250 mg/liter adenine; 2% agar for plates) was used for non-selective growth. Synthetic complete (SC) medium contained 0.15% yeast nitrogen base, 0.5% ammonium sulfate and 2% dextrose (2% agar added for plates) and was supplemented with all amino acids plus adenine and uracil. Drop-out plates missing one amino acid or base (e.g., SC-Ura medium contained no uracil) were used for selective growth. For the mitotic recombination experiments, diploid strains were grown in SC/peptone-Ura (SCP-Ura; 2% tryptone, 0.17% yeast nitrogen base lacking  $\text{NH}_4\text{SO}_4$ , 250  $\mu\text{g/ml}$  adenine; 2% agar for plates). Canavanine-resistant (Can-R) mutants were selected on SC-Arg plates containing 60 $\mu\text{g/ml}$  L-canavanine sulfate. All growth at 30° unless otherwise indicated.

### **2.2.2 Strain Constructions**

A list of all yeast strains is provided in Appendix A. Haploid strains used for mutation analyses were *RAD5* derivatives of *W303* [*ade2-1 his3-11,15 ura3-1 leu2-3,112 trp1-1 can1-100 rad5-G535R*; (249)]. *DNL4*, *YKU70*, *POL4*, *TDP1*, *WSS1*, or *RAD2* was deleted by one-step allele replacement using PCR fragments amplified from a plasmid containing a selectable drug resistance

marker. The *natMX4* cassette was amplified from pAG25 (250), *loxP-hph-loxP* from pSR955 (251), and the *loxP-G418-loxP* from pSR941 (252). The nuclease-dead *mre11-D56N* allele was introduced using two-step allele replacement following transformation with the *SphI*-digested pSM444 (253).

To increase intracellular accumulation of etoposide, the *PDR1* gene, which regulates the expression of drug efflux proteins, was replaced with a *LEU2*-marked construct containing the DNA binding domain of Pdr1 fused to the repressor domain of Cyc8 (254). Prior to transformation with an appropriate *pdr1DBD-CYC8::LEU2* fragment, the *LEU2* locus was replaced with a *loxP-TRP1-loxP* cassette amplified from pSR954 (255) to prevent recombination with the introduced fragment. The *TRP1* marker was subsequently removed by expression of the Cre recombinase from plasmid pSH63 (256). The *top2-5* temperature-sensitive allele was introduced by two-step allele replacement following transformation with *KpnI*-digested YIp5top2-5 (257) and selecting Ura<sup>+</sup> colonies. Integration of this plasmid truncates the endogenous *TOP2* allele so that only *top2-5* is expressed.

### **2.1.3 *In vivo* Complex of Enzyme (ICE) Assay in Yeast**

A C-terminally, HA-tagged mutant Top2 was generated by amplifying a fragment of pDED1Top2-FY,RG containing the *top2-FY,RG* mutations (forward and reverse primers 5'-TGTCAACTGAACCGGTAAGCGCCTCTG and 5'-CTTTGTCTCCTTGATCGTTGTGGT, respectively) and the backbone of pDED1Top2WT-3HA containing the HA-tag (forward and reverse primers 5'-

CCTAAATTGGCCAAGAAGCCAGTCAGGAA and 5'-AGCACCATAACCGTTTCTACCACCAGT). The two fragments were co-transformed into yeast and assembled by homologous recombination.

YMM10t2-4 containing the HA-tagged WT or mutant plasmid was inoculated into SC-Ura and incubated overnight at 34°. Cells were diluted to OD<sub>600</sub>=0.8 in 20 ml SC-Ura and after re-entering log phase (after approximately 2 h), DMSO or 20 µg/ml etoposide was added. After incubation for an additional 1 h at 34°, cells were harvested by centrifugation and washed in lysis buffer (6 M guanidinium thiocyanate, 1% sarkosyl, 4% Triton X-100, 1X TE, pH 7.5). A yeast protease inhibitor cocktail (Sigma-Aldrich; 50 µl per g of yeast cells) and dithiothreitol (DTT; final concentration 1%) were added prior to lysis. Cells were lysed at 4° using a Bead Beater (Biospec) for two 50-sec rounds. Lysates were incubated at 65° for 15 min, followed by a 15 min, 14,000 rpm centrifugation to remove cell debris. 400 µl of the lysates were mixed with 700 µl of 1% sarkosyl and the centrifugation was repeated. 1 ml of each lysate was mixed with 2 ml of 1% sarkosyl and then layered on 2 ml of 150% (w/v) cesium chloride (in 1X TE, pH 7.5) in 4.9 ml OptiSeal tubes (Beckman Coulter, cat #362185). Samples were centrifuged for 18 h at 42,000 rpm in an NVT 65.2 rotor (Beckman Coulter) at 25°. After removing the supernatant, pellets were washed with 500 µl of 95% ethanol and air-dried for 5 min. Pellets were resuspended in 200 µl of 1X TE and incubated at 65° for 5 min followed by a 1-h incubation at room temperature while shaking. RNase A was added to a final concentration of 100 µg/ml and samples

were incubated for 1 h at 37°. Samples were ethanol precipitated, washed with 70% ethanol and air dried for 5 min. Pellets were dissolved in 50-100 µl of H<sub>2</sub>O, incubated at 65° for 5 min and then incubated for 1 h at room temperature incubation while shaking. DNA was quantitated using a UV spectrophotometer. 10 µg aliquots were digested with 0.5 µl micrococcal nuclease for 30 min at 37° micrococcal nuclease in 500 mM Tris-Cl pH 7.9, 50 mM CaCl<sub>2</sub>. The reaction was stopped by adding 4X Laemmli buffer, samples were boiled for 5 min and equal volumes were loaded onto a 4-15% gradient gel (Bio-Rad) for SDS-PAGE. After transfer to a PVDF membrane, Top2 covalent complexes were detected using an anti-HA (Santa Cruz, sc-805) or anti-Top2 (TopoGen, #TG2014) antibody. In addition, 2 µg were transferred to nitrocellulose using a slot-blot apparatus (Bio-Rad) and an anti-dsDNA antibody (Abcam, #27156) was used to confirm equal amounts of DNA in each sample. After incubation with the corresponding secondary anti-rabbit antibody (GE Healthcare, NA934) or anti-mouse antibody (GE Healthcare, NA931), blots were visualized using SuperSignal™ West Femto ECL substrate (ThermoFisher Scientific, #34095).

#### **2.2.4 Western Blotting**

Yeast carrying either pDED1Top2-3HA or pDED1Top2-FY,RG-3HA were grown and diluted as described previously. Log-phase cells were pelleted and lysed using an alkaline lysis/glass bead procedure (258). Following neutralization, samples were digested with for 1 h on ice with micrococcal nuclease to release Top2 from DNA. Protein concentrations were determined

using the Bradford assay (Bio-Rad). Lysates were mixed with 4X Laemmli buffer, boiled for 5 min, and 30  $\mu$ g were loaded onto a 4-15% gradient gel (Bio-Rad) for SDS-PAGE. Separated proteins were transferred to a PVDF membrane and the membrane was incubated with antibodies to the HA epitope tag (Santa Cruz Biotechnology, sc-805) and  $\alpha$  tubulin (Santa Cruz Biotechnology, sc-53030), followed by incubation with the corresponding secondary anti-rabbit antibody (GE Healthcare, NA934) or anti-rat antibody (Santa Cruz Biotechnology, sc-2006). Blots were visualized as described above.

### **2.2.5 Topoisomerase Biochemical Assays**

Wild-type and mutant Top2 proteins were overexpressed from YEpGALyTop2 (259) in yeast strain JEL1t1 (260) and were purified as previously described (261). Top2 relaxation assays were carried out as previously described (261). Briefly, reactions were carried out in a final volume of 20  $\mu$ l containing 50 mM Tris-HCl, pH-8.0, 100 mM KCl, 1 mM EDTA, 8 mM MgCl<sub>2</sub>, 2% glycerol, and 1 mM ATP; purified Top2 as indicated; and 200 ng of pUC18 DNA. pUC18 was purified from *E. coli* using a Qiagen plasmid Midi kit. After incubation at 30° for 15 min, reactions were quenched with 10mM EDTA and analyzed on 1% agarose gels in 1X TAE. Conditions for Top2 DNA cleavage reactions were the same as for Top2 relaxation assays, but reactions were quenched with SDS. Following treatment with proteinase K for 2 h at 50°, samples were extracted sequentially with phenol, phenol-chloroform (50:1) and chloroform. DNA was collected by ethanol

precipitation and analyzed by agarose gel electrophoresis as above. pUC18 DNA linearized with *EcoRI* was used as a standard in cleavage assays.

### **2.2.6 Mutation and Recombination Rate Measurements**

The *TOP2* or *top2-FY,RG* allele was constitutively expressed from the *DED1* promoter (*pDED1*) in a YCp50-derived (262) centromeric plasmid with *URA3* selectable marker (263). Following transformation, plasmid presence was selected and cells subsequently maintained on SC-Ura plates. For mutation rate measurements, individual colonies containing the EV, *TOP2*, or *top2-FY,RG* plasmid were inoculated into 1 ml of SC-Ura medium and grown to saturation (3 days) on a roller drum. Appropriate dilutions were plated on SC-Ura and SC-Arg+Can plates to determine the number of viable cells and Can-R mutants, respectively, in each culture. Colonies were counted after 3 days and mutation rates were calculated using the method of the median (264); 95% confidence intervals (CIs) were determined as previously described (265). Recombination rates were similarly measured by plating saturated cultures on the appropriate selective medium.

For the etoposide experiments, the *top2-5* mutant was grown at room temperature; the WT strain was grown at either 30° or at room temperature. A given experiment was begun by inoculating a single colony into 60 ml SC-complete liquid medium. The resulting culture was then divided and either etoposide (200 µg/ml dissolved in DMSO) or an equivalent volume of DMSO was added to each half. Each of these cultures was then further separated into

individual 5-ml cultures, and these were incubated for 3 days on a roller drum. Cultures were washed two times re-suspended in 1 ml of H<sub>2</sub>O and each was sonicated prior to diluting and plating. Appropriate dilutions were plated on YPD and SC-Arg+Can as described above. Because of variation in the total number of cells in different experiments, median Can-R frequencies are reported; 95% CI for the median was determined as described above. Experiments for the *TOP2* strain were done at both room temperature and 30°, while the *top2-5* strain was grown only at room temperature.

### **2.2.7 CAN1 Mutation Spectra**

Independent Can-R colonies were obtained using a pin-plating technique that generates ~100 mini-cultures/plate. Cells were diluted in H<sub>2</sub>O and a 100-count, flat-tipped custom pinning device was used to transfer ~10<sup>3</sup> cells/pin onto SC-Ura plates. Cells were additionally spotted onto SC-Arg+Can plates to ensure that there were no pre-existing Can-R mutants. After 3 days of growth, cells were replica-plated onto SC-Arg+Can plates and were incubated for an additional 3 days. A similar technique was used to isolate independent Can-R mutants arising in the presence of 200 µg/ml etoposide (control plates contained DMSO), with growth at 30° or room temperature, as appropriate. A single Can-R colony from each spot was used for genomic DNA extraction and subsequent sequencing.

The *CAN1* locus from genomic DNAs of mutants was amplified in a 96-well format using MyTaq DNA polymerase (Bioline). Each PCR product was uniquely barcoded using primers that contained 20 nt of *CAN1*-specific sequence

conjugated to 16 nt of PacBio forward or reverse barcodes ([https://github.com/PacificBiosciences/Bioinformatics-Training/blob/master/barcoding/pacbio\\_384\\_barcodes.fasta](https://github.com/PacificBiosciences/Bioinformatics-Training/blob/master/barcoding/pacbio_384_barcodes.fasta)). The amount of each PCR product was estimated by running on agarose gels and a similar concentration of each was used to construct the pool for subsequent SMRT sequencing. Following purification of the pooled DNA (GeneJet PCR Purification Kit, ThermoFisher Scientific), SMRT libraries were constructed and sequenced by the Duke Center for Genomic and Computational Biology or LabCorp (Burlington, NC) using the PacBio RSII and Sequel systems. Circular consensus sequence (CCS) reads were sorted by barcodes and analyzed using the in-house pipeline (SmrtSeqTool) previously described (266).

### **2.2.8 *LYS2* Mutation Spectra**

pDEDTop2 or EV was introduced into the *lys2ΔA746,NR* and *lys2ΔBgl,NR* strains SJR1467 and SJR1468, respectively (248). Transformants were selected and maintained SC-Ura plates. Independent Lys<sup>+</sup> colonies were obtained using the pin-plating technique described above. After 3 days of growth, cells were replica-plated onto SC-Lys plates and incubated for an additional 5 days. A single Lys<sup>+</sup> colony from each spot was used for genomic DNA extraction. The relevant portion of the *LYS2* gene was amplified by PCR and sequenced by GeneWiz (Durham, NC) using the primer 5'-GTAACCGGTGACGATGAT.

### 2.2.9 Recombination Experiments

The EV and pDEDtop2-FY, RG plasmids were transformed into a diploid yeast strain carrying multiple heteroallelic markers (*tyr1-1/tyr1-2*, *his7-2/his7-1*, *trp5-d/trp5-c* and *lys2-1/lys2-2*). Cells containing the plasmids were selected on SC-Ura plates. Approximately  $10^4$  cells were inoculated into 1 ml of SCP-Ura liquid medium and cultures were incubated for 4 days on a roller drum.

Appropriate dilutions were plated on SCP-Ura plates to determine the number of viable cells in each culture and onto SC-Tyr, SC-His, SC-Trp, and SC-Lys plates to determine the numbers of recombinants. The median recombination frequency was reported for each heteroallelic marker because of the variation in total cell numbers in different experiments.

### 2.2.10 Statistical Methods

The proportions of mutation types in different strains were compared using a contingency Chi-Square or Fisher Exact test as appropriate ([vassarstats.net](http://vassarstats.net));  $p < 0.05$  was considered significant. Rates were calculated using method of the median (264) and 95% CIs for rates and median frequencies were determined (267). Mutation-type rates or frequencies were calculated by multiplying the total Can-R rate or frequency by the proportion of the mutation type in the corresponding spectrum. The 95% CI for each mutation type was determined by jointly considering the CI for the Can-R rate or frequency and the CI for the proportion ([Vassarstats.net](http://Vassarstats.net)). This was done using the “root of the square of the sums” (RSS) or right triangle method (268). Rates/frequencies obtained in

different strain backgrounds were considered significantly different if the respective 95% CIs did not overlap.

## **2.3 Results**

### **2.3.1 C-terminal *TOP2* Mutations Confer Hypersensitivity to Etoposide**

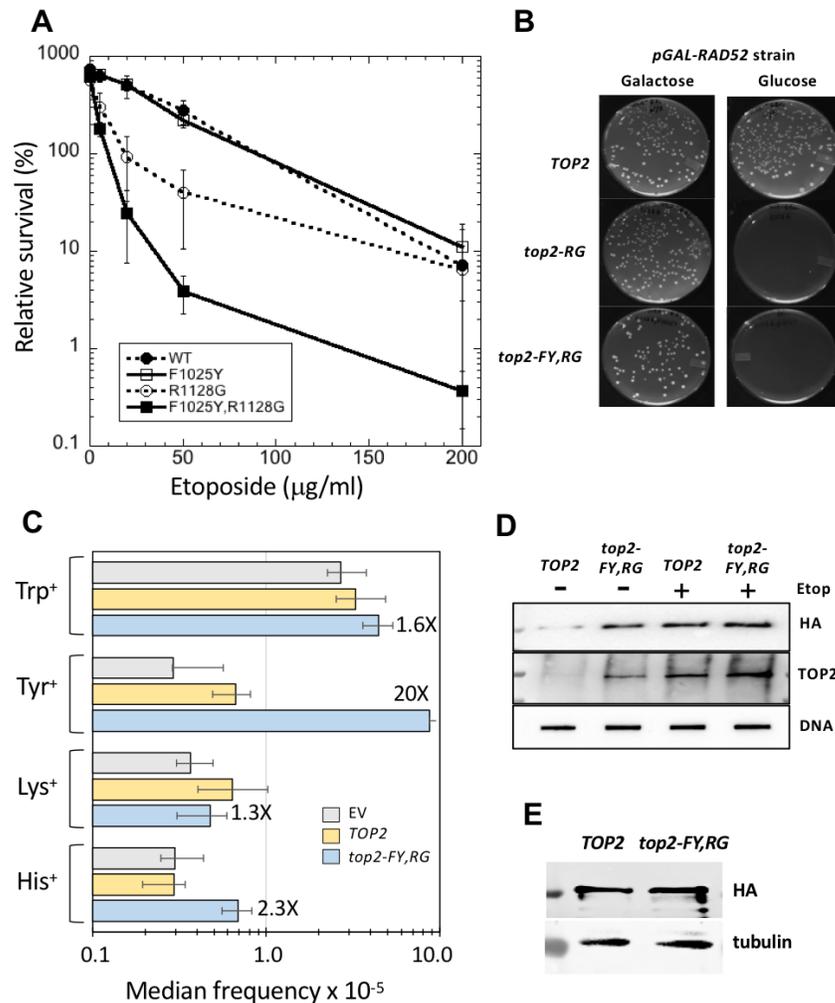
We previously described mutations in yeast *TOP2* that result in hypersensitivity to different classes of Top2-targeting agents such as etoposide and *m-AMSA* (260). To identify additional mutations conferring this phenotype, we carried out a screen following the introduction of random mutations by error-prone PCR into a DNA segment specifying amino acids 900 to 1250. The mutagenized fragment was recombined into gapped plasmid pDED1Top2 (260, 263) by co-transformation of the gapped plasmid and the mutagenized fragment into a temperature-sensitive *top2-4* strain (44). Expression of *TOP2* in strains carrying pDED1Top2 results in ~10-fold more protein than does expression from the native *TOP2* promoter (263). Transformants were selected at 34°C, which is non-permissive for the *top2-4* allele, to limit the isolation of plasmids carrying *top2* null alleles. Individual colonies were screened for sensitivity to mAMSA on solid medium and two mutants with alterations in the C-terminal part of cleavage/ligation domain of Top2 were identified. The first mutant had Arg1128 changed to Gly (R1128G) while the second (independent) mutant had the R1128G change plus an additional mutation that converted Phe1025 to Tyr (F1025Y).

The F1025Y and R1128G changes (FY and RG, respectively) were re-introduced into the pDED1Top2 plasmid individually or together by site-directed mutagenesis. Following introduction into the *top2-4* background, transformants were assessed for survival following exposure to varying concentrations of mAMSA or etoposide at 34°C. *top2-4* cells carrying the pDED1Top2 or pDED1Top2-FY plasmid were insensitive to mAMSA even at a concentration of 50 µg/ml (Appendix B). By contrast, cells containing the *top2-RG* or *top2-FY,RG* allele were sensitive to 5 µg/ml mAMSA. *top2-4* cells carrying the pDED1Top2 plasmid were modestly sensitive to etoposide (Fig. 6A) and those with the pDED1Top2-RG plasmid were more sensitive to the drug. Although over-expressing the *top2-FY* allele alone did not enhance etoposide sensitivity relative to the *TOP2* control, the double-mutant *top2-FY,RG* allele was associated with a high level of drug sensitivity. Based on drug-sensitivity profiles, only strains containing the plasmid-encoded *top2-RG* or *top2-FY,RG* allele were further analyzed.

### **2.3.2 Expression of *top2-RG* or *top2-FY,RG* is lethal in the absence of *RAD52***

Previous drug-hypersensitive *top2* mutants were viable in a *rad52*Δ background that is unable to repair DSBs (269). We were unable, however, to obtain colonies at 34°C following the transformation of either the pDED1Top2-RG or pDED1Top2-FY,RG plasmid into a *top2-4 rad52*Δ strain. To more rigorously demonstrate lethality of the *top2* alleles in the absence of recombination, we

introduced a plasmid containing a *pGAL-RAD52* fusion into a *top2-4 rad52Δ* background. Cells were thus phenotypically Rad52<sup>+</sup> or Rad52<sup>-</sup> on galactose- or glucose-containing medium, respectively. Cells transformed with pDED1Top2 grew well on both types of plates, while those carrying the mutant pDED1Top2-RG or pDED1Top2-FY, RG plasmid formed colonies only on galactose-containing medium (Fig. 6B). This effect was verified by streaking colonies obtained on galactose medium in parallel onto medium containing either glucose or galactose.



**Figure 6: Genetic Characterization of *top2* Mutants**

A) The *top2-4* host strain JN362at2-4 (263) was transformed with pDED1Top2 plasmids containing the indicated alleles. Transformants were grown to mid-log phase in SC-Ura medium at 34°C and the indicated concentration of etoposide was added. Incubation was continued for an additional 24 h before plating cells for survival. The survival plotted is relative to that at the time of etoposide addition. Error bars are  $\pm$ SEM. B) The pGAL-Rad52 plasmid (270) was transformed into a *top2-4 rad52* $\Delta$  background (JN332at2-4) and then cells grown in galactose medium were transformed with pDED1Top2 containing the indicated allele. Transformants were selected on uracil-deficient medium containing either glucose or galactose as a carbon source. C) Diploid strain CG2009 was

transformed with the relevant plasmid and grown selectively in SC/peptone-Ura medium prior to plating on drop-out media (SC/peptone) to select recombinants. The median recombination frequency is plotted, and the error bars are 95% confidence intervals. D) YMM10t2-4, a *top2-4* derivative of YMM10 (271), and was transformed with the pDED1Top2 or pDED1Top2-FY,RG plasmid containing an HA-tagged allele. Transformants were grown in the presence or absence of 200  $\mu$ g/ml etoposide prior to genomic DNA isolation using the yeast ICE protocol. After DNA recovery, quantitation, and digestion with micrococcal nuclease, DNA-associated proteins were separated by SDS-PAGE and Top2 was detected using an anti-HA or anti-TOP2 antibody. Survival data following 24 h growth in the presence of mAMSA are shown in Appendix B. E) Top2 and Top2-FY,RG levels relative to tubulin are shown.

Given the requirement for homologous recombination functions for survival, we hypothesized that expression of the mutant Top2 proteins would confer a hyper-recombination phenotype. The double-mutant plasmid, WT pDED1Top2 or the empty vector (EV) YCp50 (262) thus were introduced into a diploid yeast strain carrying multiple heteroallelic markers. Recombination frequencies were measured by selecting for tryptophan, lysine, histidine or tyrosine prototrophs. Relative to the EV control, over-expression of *TOP2* from *pDED1* did not enhance recombination frequencies between any of the heteroallelic pairs examined (Fig. 6C). Over-expression of the *top2-FY,RG* allele, however, resulted in 2.3-fold increase in His<sup>+</sup> recombinants and a 20-fold increase in Tyr<sup>+</sup> recombinants; there was no significant increase in Lys<sup>+</sup> or Trp<sup>+</sup> recombinants. The reason for the variable recombination effects of *top2-FY,RG* allele is unclear, but could reflect the relative distances between heteroalleles at the loci monitored, differing levels of gene expression, preferred sites of Top2 cleavage, or other features related to local chromosome structure.

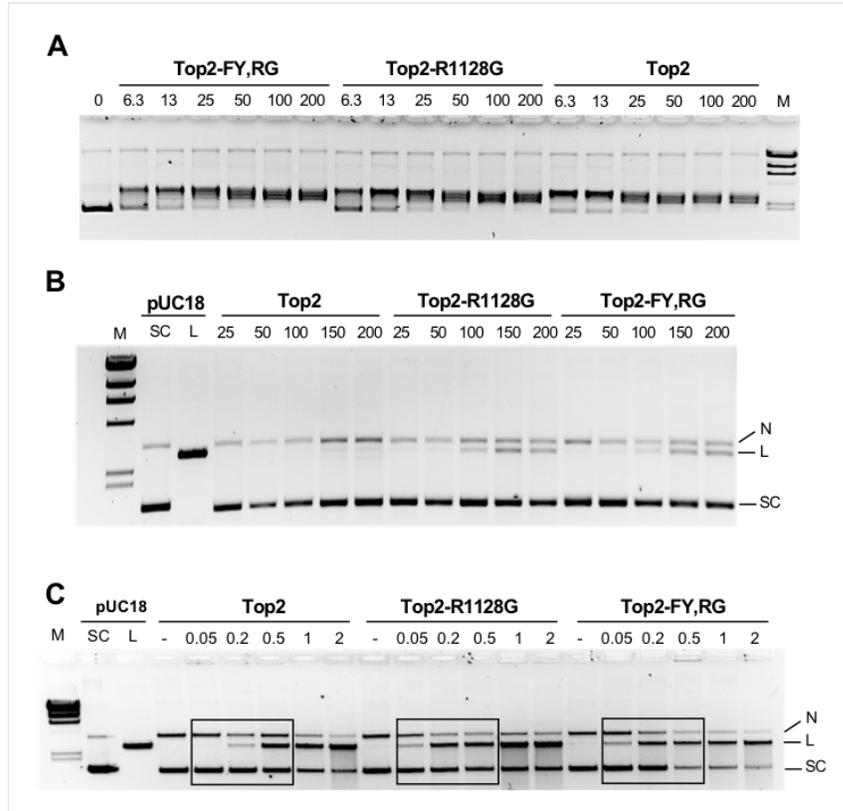
### **2.3.3 Top2-FY,RG is trapped on yeast DNA in the absence of etoposide**

Top2 poisons such as etoposide interfere with DNA re-ligation catalyzed by the enzyme and create DNA damage in the form of strand breaks with covalently attached protein at the ends. Adducts formed *in vivo* can be detected using an ICE (*in vivo* complex of enzyme) assay that immunologically detects protein covalently associated with DNA (271, 272). For the ICE assay in yeast, we constructed variants of WT and double-mutant pDED1 plasmids in which the C-terminus of the Top2 protein was tagged with HA. These variants were then expressed in a *top2-4* strain with increased etoposide sensitivity due to the absence of multiple drug-efflux pumps (271). Cells expressing WT *TOP2* showed a faint Top2 signal that was greatly enhanced when cells were treated with etoposide (Fig. 6D). By contrast, cells expressing the double-mutant allele showed a robust signal in the absence of etoposide that was further enhanced by drug treatment. Given the similar levels of the Top2 and Top2-FY,RG proteins (Fig. 6E), we conclude that Top2-FY,RG is a self-poisoning enzyme that frequently becomes trapped on DNA in the absence of etoposide.

### **2.3.4 Relaxation and cleavage activities of mutant Top2 proteins**

For biochemical characterization of Top2-RG and Top2-FY,RG, the relevant mutations were introduced into plasmid pGAL1Top2 (273). The proteins were then overexpressed in and purified from a *top1* $\Delta$  background in order to eliminate the confounding effects of Top1 activity. The relaxation activities of the

purified proteins were assessed using negatively supercoiled pUC18 DNA as substrate (261). Under our standard conditions, complete relaxation of pUC18 by the WT or Top2 mutant proteins required 25-50 ng of protein (Fig. 7A). This result indicates that the overall catalytic activity of the mutant proteins was similar to that of WT yeast Top2. We next examined the ability of the mutant Top2 proteins to cleave pUC18 DNA in the absence of Top2 targeting agents. Double-strand DNA cleavage of plasmid DNA results in the formation of linear DNA molecules. For WT Top2, a very faint linear band was seen with 150 ng of purified protein (Fig. 7B). By contrast, linear fragments were observed with Top2-RG and Top2-FY,RG at protein concentrations as low as 50 ng. We also examined the response of Top2-RG and Top2-FY,RG to etoposide (Fig. 7C) and *m-AMSA* (Appendix B). In the presence of these small-molecule inhibitors, the mutant proteins were associated with a higher level of linear DNA as the drug concentration was increased. At higher etoposide concentrations (e.g., Top2-RG at etoposide concentrations greater than 1  $\mu\text{g/ml}$ ) smearing of DNA was observed, which presumably represents plasmids cleaved at two or more separate sites. These results demonstrate that the purified Top2-RG and Top2-FY,RG proteins have an intrinsic hypersensitivity to *mAMSA* and etoposide that is in agreement with the phenotypes of Top2-RG and Top2-FY,RG *in vivo*.

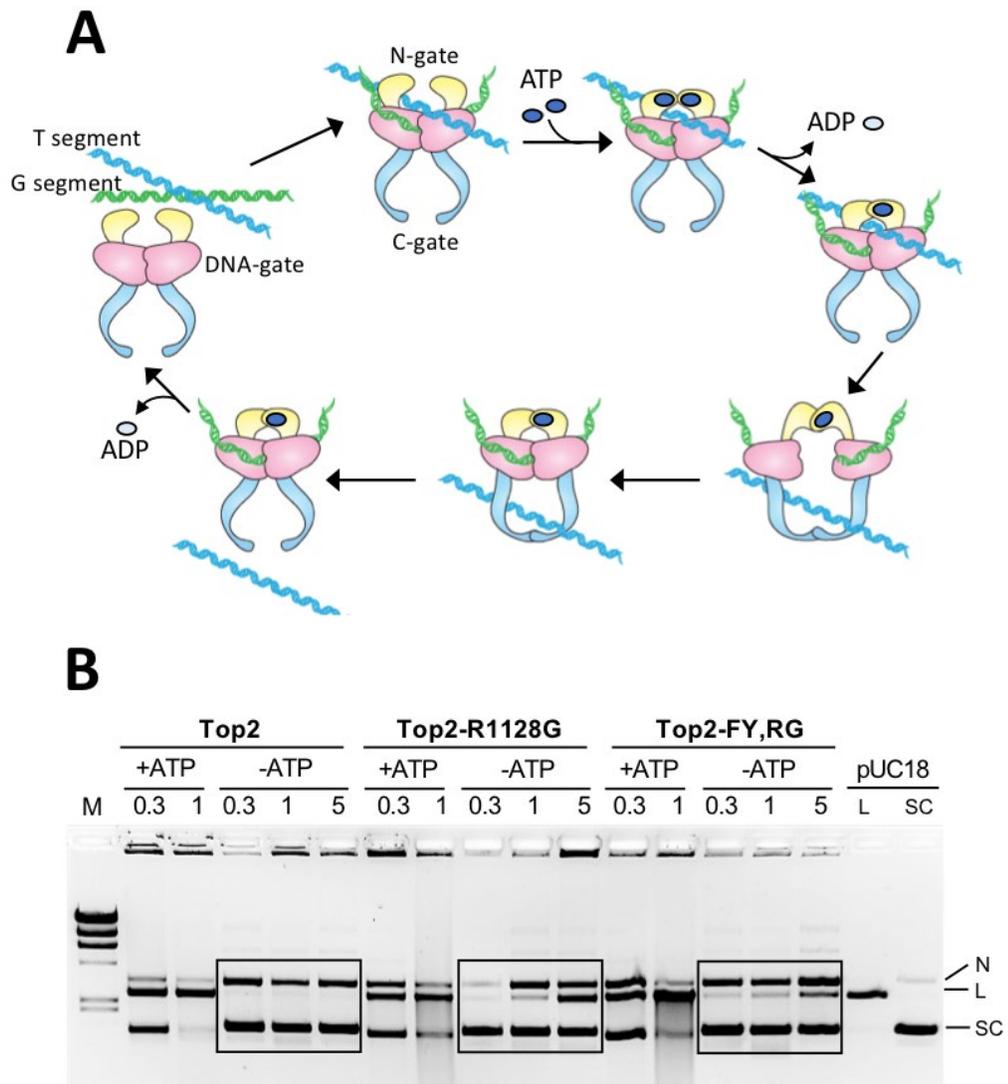


**Figure 7: Biochemical Activities of WT and Mutant Proteins.**

Top2 proteins were expressed from the *pGAL1* promoter in the *top1* $\Delta$  strain JELt1 (260) and 200 ng of negatively supercoiled pUC18 DNA was used in all reactions. A) DNA strand-passage activity of purified WT, Top2-R1128G and Top2-FY,RG proteins; the amount of protein (ng) is indicated above each lane. B) DNA cleavage activity of WT and mutant Top2 proteins; the positions of nicked (N), linear (L) and supercoiled (SC) pUC18 are indicated. DNA cleavage is quantitated in Appendix C. C) DNA cleavage activity of WT and mutant Top2 proteins in the presence of etoposide; etoposide concentration ( $\mu\text{g/ml}$ ) is indicated above each lane. Data in the presence of mAMSA are in Appendix B.

The catalytic cycle of Top2 is illustrated in Fig. 8A. Although ATP is required for supercoil relaxation and for decatenation by Top2, eukaryotic Top2 cleaves DNA, albeit at a much lower level, even in the absence of ATP (230). We were interested in determining whether the mutants that resulted in elevated

DNA cleavage required progression through the catalytic cycle, and were specifically defective in re-ligation after strand passage had occurred. To test this possibility, we examined stable cleavage of pUC18 DNA by the WT and mutant proteins in the absence of ATP (Fig. 8B). For WT Top2, linearization of pUC18 was barely detectable, although DNA nicking was seen. By contrast, linearized DNA was readily detected when pUC18 was incubated with 1  $\mu$ g Top2-RG or 0.3  $\mu$ g Top2-FY, RG. These results demonstrate that ATP is not required for elevated DNA cleavage by the mutant proteins, although it still strongly potentiates cleavage. Because ATP (or a non-hydrolyzable analog) is required for progression through the catalytic cycle and for strand passage, these results suggest that progression of the catalytic cycle is not required for elevated, drug-independent cleavage of DNA by these proteins.



**Figure 8: Progression Through the Top2 Catalytic Cycle is not Required for Elevated Cleavage.**

A) The catalytic cycle of eukaryotic Top2 is illustrated (274). Top2 cleaves DNA in the presence of a divalent cation but cannot proceed to strand passage without ATP binding (blue circles) and dimerization of the N-gate (yellow). The two ATP molecules are hydrolyzed sequentially. Strand passage through the G segment (green) is followed by re-ligation and release of the T segment (blue) through the C-gate. B) DNA cleavage activity of purified WT and mutant Top2 proteins in the presence or absence ATP; protein concentration ( $\mu\text{g}$ ) is indicated above each

lane. Positions of nicked (N), linear (L) and supercoiled (SC) pUC18 are indicated.

### **2.3.5 Top2-FY,RG is mutagenic and is associated with *de novo* duplications**

Expression of the *top2-FY,RG* allele resulted in a hyper-recombination phenotype and was lethal in a *rad52Δ* background (Fig. 6), consistent with formation of potentially toxic DSBs. *In vitro*, the mutant protein generated persistent nicks as well as DSBs (Fig. 7), leading us to examine whether its expression might be mutagenic. For this analysis, the EV, pDEDTop2 or pDED1Top2-FY,RG plasmid were introduced into a haploid *TOP2* background. The *CAN1* forward-mutation assay was used to measure mutation rates and analyze mutation types. In this assay, any mutation that disables function of the encoded protein confers resistance to canavanine, a toxic arginine analog. The canavanine-resistance (Can-R) rate was indistinguishable in cells containing either the EV or pDED1TOP2. By contrast, expression of the *top2-FY,RG* allele elevated the Can-R rate 2.7-fold (Appendix D).

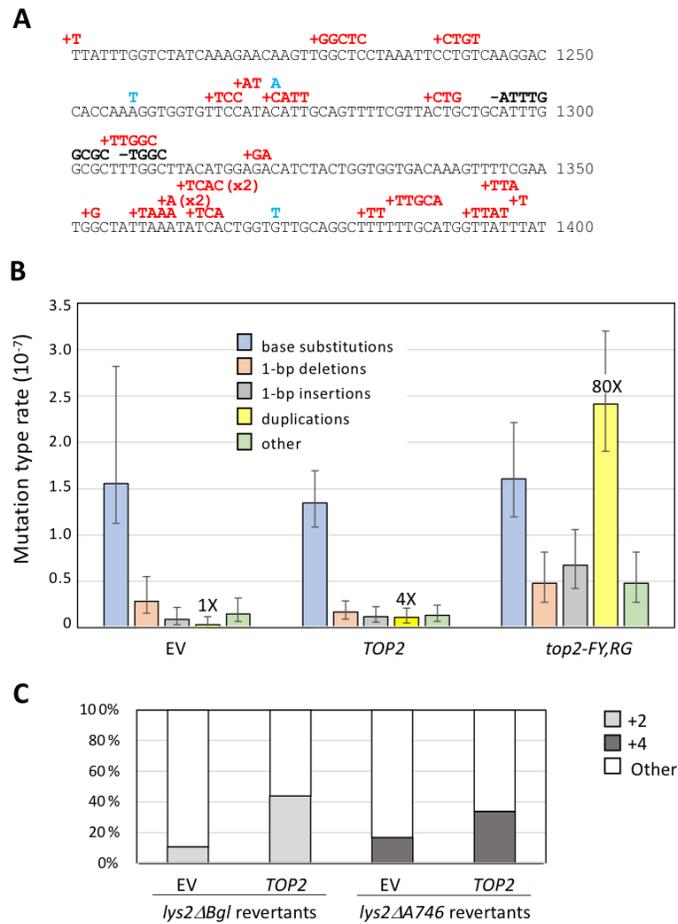
Approximately 75% of *can1* mutations detected were base substitutions when either the EV or the pDED1Top2 plasmid was present (105/142 and 111/155, respectively;  $p=0.74$  by contingency Chi-square). By contrast, only 28% of mutations were base substitutions in the strain containing the pDED1Top2-FY,RG plasmid (50/176;  $p<0.0001$ ). The proportional decrease in base substitutions was accompanied by a large increase in insertions of more than

one base pair (from 2/142 with the EV to 75/176;  $p < 0.0001$ ). Most of these insertions (49/75) corresponded to *de novo* duplications, which are defined as the creation of a repeat where one did not previously exist. In the segment of the *CAN1* ORF shown in Fig. 9A, for example, CTGT (nt 1239-1242) became CTGTCTGT in one mutant and CATT (nt 1272-1275) was duplicated to CATTCATT in another. The remainder of insertions >1-bp occurred within a pre-existing repeat; because of the genetic requirements for their formation (see below), we consider these jointly with the *de novo* duplications. The most frequent size of duplications was 4 bp, which matches the distance between Top2-generated nicks *in vitro* (275) and is relevant to the proposed mechanism of duplication formation (see Discussion).

The rates of specific mutation types (base substitutions, 1-bp deletions, 1-bp insertions, *de novo* duplications and others) in the presence of the EV, the pDED1Top2 or the pDED1Top2-FY,RG plasmid are presented in Fig. 9B. We estimate an ~80-fold increase in the rate of duplications, but no change in the base substitution rate, when the *top2-FY,RG* allele was overexpressed. It should be noted that there was also a 7.6-fold increase in the rate of +1 insertions associated with pDED1Top2-FY,RG. Although most of these 1-bp insertions were in short homopolymer runs and could reflect DNA polymerase slippage during replication, their increase suggests that many were likely generated by the same mechanism as the larger insertions.

### 2.3.6 Top2 overexpression elevates duplications in frameshift reversion assays

Although there was no increase in the overall rate of Can-R in the presence of the pDED1Top2 plasmid, there was an ~4-fold increase in duplications. Given their small number, however, neither the proportional increase in these events (2/142 for EV and 9/155 for pDED1Top2;  $p=0.09$ ) nor their corresponding rate was significantly increased relative to cells that did not overexpress Top2. To focus specifically on 2- and 4-bp duplications, we used frameshift-reversion assays where similar events were previously observed (248). The *lys2 $\Delta$ Bgl, NR* allele reverts by acquisition of a net -1 frameshift and so can detect 2-bp insertions; the *lys2 $\Delta$ A746, NR* reverts by net +1 frameshifts, which includes 4-bp insertions. Top2 over-expression was accompanied by a proportional increase in 2-bp (from 3/28 to 22/50;  $p=0.0056$ ) as well as 4-bp duplications (from 6/36 to 33/98,  $p=0.088$ ; Fig. 9C and Appendix E). When these events were jointly considered, their increase was highly significant ( $p=0.0014$ ), demonstrating an association of a specific class of mutations with Top2 overexpression.



**Figure 9: Top2-FY,RG Expression is Associated with *de novo* Duplications**

A) Partial *CAN1* mutation spectrum. Mutations are above the sequence; insertions are in red, base substitutions in blue and deletions in black font. B) Rates of specific mutation types in a *TOP2* strain containing the EV, pDED1Top2 or pDED1Top2-FY,RG. Error bars are 95% confidence intervals. C) Proportions of 2-bp and 4-bp duplications among *Lys*<sup>+</sup> revertants of the *lys2ΔBgl,NR* (SJR1467) and *lys2ΔA746,NR* (SJR1468) alleles (276). Can-R rates and associated spectra are in Appendix D. Complete *Lys*<sup>+</sup> spectra are in Appendix E.

### 2.3.7 Top2cc-dependent duplications require the NHEJ pathway

*In vitro*, the Top2-FY,RG protein generates nicks in addition to DSBs (Fig. 7), either of which potentially could initiate the *de novo* sequence duplications observed *in vivo*. Because the NHEJ pathway can introduce sequence changes at the junction of joined ends and was previously implicated in generating the *de novo* duplications in the frameshift-reversion assays described above (248), we examined the relevance of this pathway to the mutagenesis associated with stabilization of the Top2 cleavage complex (Top2cc; Fig. 10A and Appendix D). We first deleted the *DNL4* gene, which encodes the DNA ligase required for NHEJ (277), in cells containing the pDED1Top2-FY,RG plasmid. In the *dnl4* $\Delta$  background, there was a 40% reduction in the Can-R rate that was accompanied by a large proportional decrease in *de novo* duplications (from 75/176 to 7/194;  $p < 0.0001$ ) as well as 1-bp insertions (from 21/176 to 7/194;  $p = 0.005$ ). The Ku complex is also required for NHEJ in yeast, and results in *ku70* $\Delta$  background were similar to those obtained in the *dnl4* $\Delta$  background (Appendix D). In addition to the requirement for Dnl4 and Ku, most end/gap-filling that occurs during NHEJ requires DNA polymerase 4 (Pol4). Deletion of *POL4* from the strain containing the pDED1Top2-FY,RG plasmid also significantly reduced the Can-R rate and the proportion of duplications (from 75/176 to 25/126;  $p < 0.0001$ ). The effect of *POL4* loss on duplications, however, was not as dramatic as that observed in the absence of *DNL4* (25/126 and 7/194, respectively;  $p < 0.0001$ ). Loss of *POL4* had no effect on 1-bp insertions (21/176 and 12/126 in WT and *pol4* $\Delta$ , respectively;

p=0.63), which is consistent with its reduced requirement for duplications and the relatively high background of +1 events. Together, these data demonstrate that the majority of sequence duplications in this system are products of Top2-generated DSBs that are repaired by NHEJ.

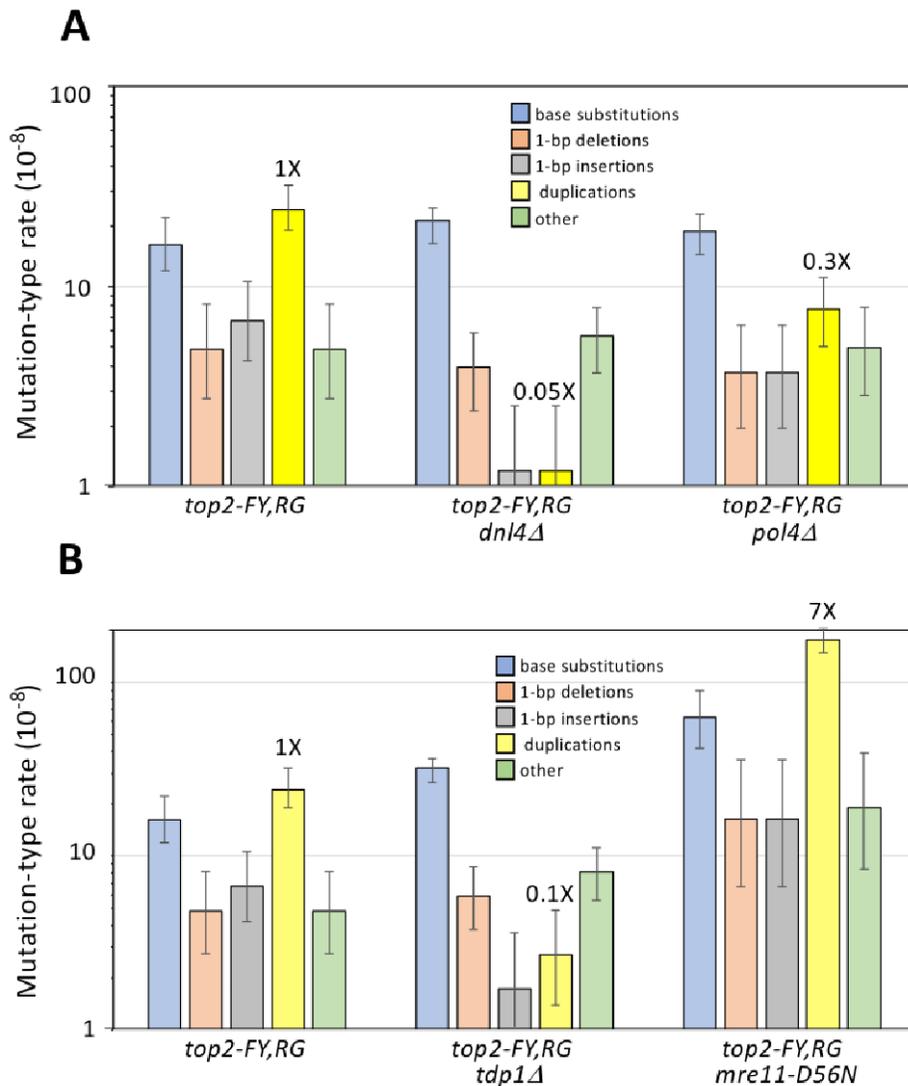
### **2.3.8 Pathways for Top2cc removal affect duplication rates**

A stabilized yeast Top2cc can potentially be removed by either (1) proteolytic degradation followed by peptide extraction through cleavage of the phosphotyrosyl bond or (2) nuclease-dependent removal of a Top2-linked oligonucleotide. In the first pathway, the Top2 peptide that remains after proteolysis is removed by tyrosyl-DNA phosphodiesterase 1 (Tdp1; (190)).

Although *TDP1* loss had no significant effect on the Can-R rate in the presence of the pDED1Top2-FY,RG plasmid, there was a large proportional reduction in duplications (75/176 to 11/205; p<0.0001) as well as 1-bp insertions (21/176 to 7/205; p=0.003) and rates of these events were significantly reduced (Fig. 10B and Appendix D).

In higher eukaryotes, a major pathway for Top2cc removal from a DNA end is by the endonuclease activity of the MRE11 component of the MRN (MRE11-RAD50-NBS1) complex, which releases a Top2-linked oligonucleotide (208). In *S. cerevisiae*, MRX (Mre11-Rad50-Xrs2) is the equivalent complex, and loss of any of the component proteins confers etoposide hypersensitivity (278). Although we were unable to obtain colonies following transformation of *mre11Δ* cells with the pDED1Top2-FY,RG plasmid, transformants were readily obtained

in a background containing the nuclease-dead *mre11-D56N* allele (279). The essential role of yeast MRX in dealing with Top2-associated damage is thus distinct from its nuclease activity. Expression of the *top2-FY,RG* allele in an *mre11-D56N* background was associated with a 7.3-fold increase in the rate of *de novo* duplications (Fig. 10B), indicating that Mre11 nuclease activity is primarily responsible for yeast Top1cc removal from DNA ends. In its absence, phosphotyrosyl peptide that remains after Top2 proteolysis is cleanly removed Tdp1 and the ends give rise to NHEJ-dependent duplications.



**Figure 10: Genetic Control of Mutations Associated with *top2-FY,RG* Expression.**

Indicated candidate genes were deleted in the *TOP2* background and the resulting strains were transformed with EV, pDED1Top2 or pDED1Top2-FY,RG. A) Duplications require Dnl4 and partially require Pol4. B) Duplications are promoted by Tdp1 and suppressed by the nuclease activity of Mre11 (*mre11-D56N*). Rates and associated spectra are in Appendix D. The fold change in duplication rate in each mutant relative to the WT background (1.0X) is indicated.

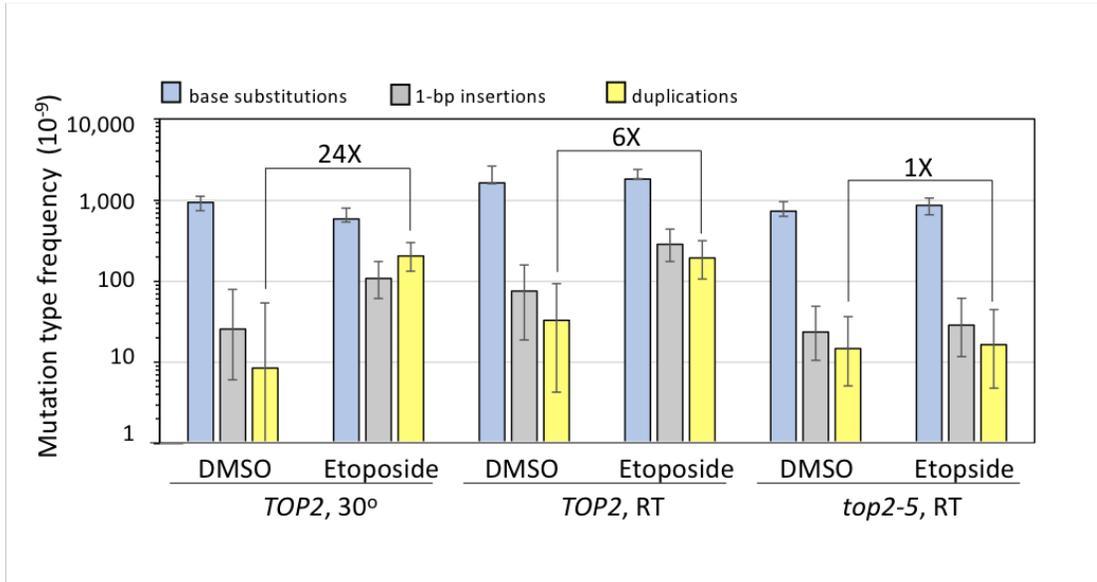
### **2.3.9 Wss1 and Rad2 are not involved in the formation of Top2-dependent duplications**

Wss1 is a yeast metalloprotease and is a functional homolog to SPRTN, a protease in higher eukaryotes. SPRTN degrades both TOP1cc and TOP2cc, while there is only evidence of Wss1 acting on Top1cc (157, 158, 170-177). To determine if Wss1 can degrade stabilized Top2cc in yeast, as well as to investigate its involvement in Top2-dependent duplications, *WSS1* was deleted from a strain containing the pDED1Top2-FY,RG plasmid. No differences between mutation rates or spectrum were observed in the WT and *wss1* strains containing the pDED1Top2-FY,RG plasmid (Appendix D).

The Rad2 nuclease is essential for nucleotide excision repair in eukaryotes. Rad2 specifically cleaves the 5' junction of the damage-containing bubble, suggesting that it could be involved in the removal of a stabilized Top2cc from short 5' overhangs (280). Additionally, previous studies have shown that a *tdp1Δ rad2Δ* double mutant has a substantial growth defect on plates containing etoposide. This sensitivity is additive with respect to that of the single mutants alone (190), suggesting that Rad2 is involved in the removal of stabilized Top2cc. To determine if Rad2 plays a role in the formation of the Top2-FY,RG duplications, *RAD2* was deleted from strains containing the pDED1Top2-FY,RG plasmid. There was no significant difference in the Can-R mutation rate or in the distribution of mutation types in the *rad2* strain compared to WT, both containing the pDED1Top2-FY,RG plasmid (Appendix D).

### 2.3.10 Mutagenic effects of etoposide

To further examine the dependence of duplications on Top2cc stabilization, we isolated Can-R mutants in a drug-sensitized *TOP2* strain grown in the presence of either DMSO, the vehicle for etoposide, or DMSO plus 200  $\mu\text{g/ml}$  etoposide. Although etoposide did not alter the median Can-R frequency (Appendix F), there were significant changes in the corresponding spectrum and in the frequencies of some mutation types (Fig. 11 and Appendix F). Relative to the DMSO-treated cultures, there was a proportional increase in duplications (from 1/162 to 36/222;  $p < 0.0001$ ) as well as 1-bp insertions (from 3/162 to 19/222;  $p < 0.0001$ ) in the presence of etoposide. Confirmation that the mutagenic effect of etoposide was mediated through Top2cc stabilization was obtained using the etoposide resistant *top2-5* allele (257). Because this allele confers temperature sensitivity, experiments were performed at room temperature (RT) rather than 30°C. Interestingly, the Can-R frequency in the WT strain was elevated ~2-fold when cells were grown at RT in the presence of DMSO or etoposide (Fig. 11 and Appendix F). As at 30°C, however, etoposide treatment of the WT strain at RT stimulated duplications (from 4/288 to 18/305;  $p = 0.007$ ) as well as 1-bp insertions (9/288 and 27/305;  $p = 0.006$ ), and their rates were not significantly different from those observed at 30°C. By contrast, etoposide stimulated neither duplications (5/344 in DMSO and 4/297 in etoposide, respectively;  $p = 1$ ) nor 1-bp insertions (8/344 and 7/297;  $p = 1$ ) in the *top2-5* background.



**Figure 11: Etoposide Stimulates Duplication Through its Interaction with Top2.**

A *TOP2 pdr1DBD-CYC8::LEU2* (drug-sensitized) strain was grown in the presence of DMSO or DMSO+etoposide at RT or at 30°C. The temperature-sensitive *top2-5 pdr1DBD-CYC8::LEU2* strain only grows at RT and is etoposide resistant. Median Can-R frequencies were determined rather than rates because of the variation in the number of viable cells in etoposide-treated cultures; error bars are 95% confidence intervals. Median frequencies and associated spectra are in Appendix F.

## 2.4 Discussion

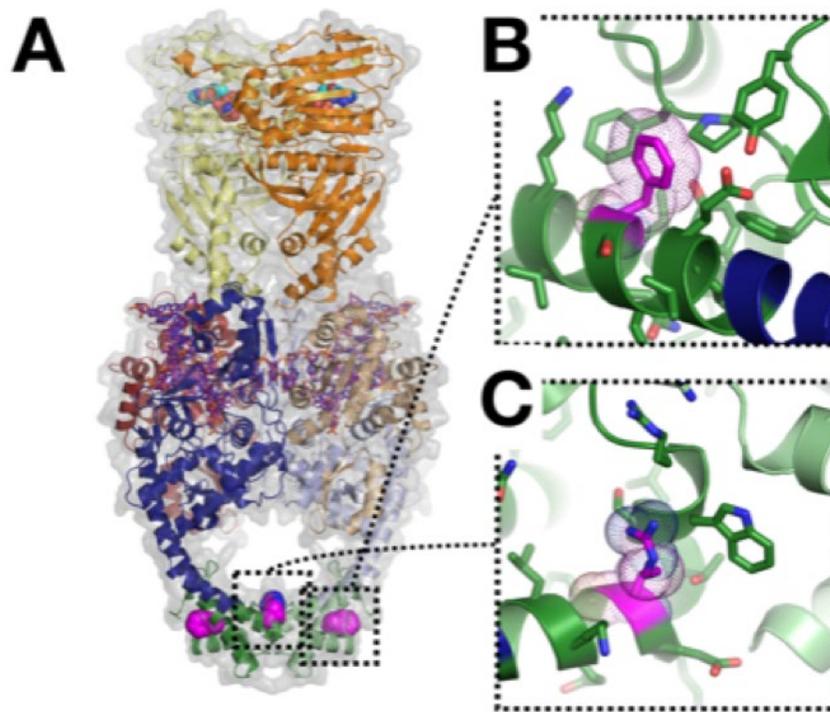
Topoisomerases carry out essential reactions that require DNA cleavage and a failure to quickly re-ligate DNA can lead to genome destabilization (243). In the current study, we identified and characterized yeast *top2* alleles (*top2-R1128G* and *top2-F1025Y,R1128G*) that produce proteins that are defective in quickly following up cleavage with re-ligation. Although the mutant proteins supported the essential function of Top2 *in vivo*, they conferred lethality in strains defective in the recombinational repair of DSBs or in the MRX complex.

Consistent with elevated DSBs, overexpression of the Top2-FY,RG protein conferred hypersensitivity to Top2 poisons, was associated with a mitotic hyper-recombination phenotype and led to elevated levels of covalent Top2-DNA complexes. This phenotype is consistent with the elevated DNA cleavage observed biochemically and is similar to that of etoposide in WT cells.

Although the Top2-FY,RG protein shows elevated cleavage, both amino acid substitutions map to the C-terminal dimer interface or 'C-gate' region (Fig. 12A) rather than near the catalytic tyrosine. The C-gate constitutes the primary dimer interface that holds Top2 subunits together in the absence of DNA or ATP (281-283). During the Top2 catalytic cycle, this interface splits apart to allow a newly transported duplex that has crossed the cleaved-DNA segment to exit the enzyme (284). The stability of the C-gate interface, together with that of the nucleotide-dimerized ATPase domains, is thought to hold the two Top2 subunits together and guard against the accidental dissociation of the dimer during its duplex cleavage and passage reaction. Such dissociation would be expected to lead to persistent DSB formation.

A close up of the interface shows that F1025 protrudes from the coiled-coil arms of the principal DNA-binding region and nestles into a small hydrophobic pocket on one side of the globular region of the C-gate (Fig. 12B). Given the location of F1025 and the protrusion of the tip of its benzyl ring so that it is solvent exposed, it seems unlikely that its replacement with tyrosine, which would project its phenolic oxygen into solution, would have an effect on C-gate stability.

By comparison, R1128 packs against a tryptophan (W1122) that directly forms part of the dimer interface (Fig. 12C). This interaction, together with the introduction of a glycine in the middle of a helical element, would be expected to locally destabilize the region. Such destabilization could impact the integrity of the C-gate or the kinetics with which the dimer interface separates and re-associates. Either behavior could, in turn, detrimentally affect the cleavage propensity of Top2 and increase the life-time of the cleaved DNA state, the possibility of subunit dissociation, or both.



**Figure 12: The F1025Y/R1128G Mutations Map to the Top2 C-gate.**

A) Cartoon representation of a full-length, nucleotide-bound *S. cerevisiae* Top2 dimer bound to a cleaved DNA segment (PDB ID 4GFH; (285)). One subunit is colored in dark hues and the other in light. Dark/light-orange – ATPase domains;

dark/light-red – TOPRIM Mg<sup>2+</sup>-binding domain; dark/light-blue – principal DNA binding region; dark/light-green – C-gate; dark-purple – cleaved DNA. AMPPNP is shown as cyan spheres; the locations of F1025 and R1128 are marked as magenta spheres. B) Close-up of the region around F1025 is highlighted, with F1025 shown in magenta. C) Close-up of the region around R1128, illustrating the packing of this residue adjacent to W1122 at the dimer interface.

The hyper-recombination phenotype associated with Top2-FY,RG expression in yeast, as well as the synthetic lethality in the absence of recombination, are consistent with stabilization of the cleavage intermediate. A mutator phenotype was not expected, however, and could reflect either the persistence of nicked intermediates (see Fig. 7) or errors associated with the repair of DSBs. The sequencing of Can-R mutants was particularly informative, revealing a shift from predominantly base substitutions to a rarely observed type of insertion: *de novo* duplication of 2 to 4 bp of sequence. The largest class of duplications was 4 bp in size, which matches the distance between the nicks Top2 makes on complementary DNA strands, and all duplications were NHEJ dependent. As illustrated on the left side of Fig. 13, the complete filling in of the 5' overhangs generated by Top2 cleavage, followed by NHEJ-mediated ligation, creates a 4-bp duplication.

A composite spectrum of insertions >1 bp that were identified in the WT and *mre11-D56N* backgrounds (Appendix G) suggests the occurrence duplication hotspots. These could reflect sites of preferred Top2-FY,RG cleavage/stabilization or sites where end filling is more frequent and NHEJ is relatively error prone. The positions of Top2 cleavage can be inferred from 4-bp

duplications, but not for smaller duplications. These cleavage sites were aligned and analyzed for possible sequence conservation, such as the weak dyad symmetry reported for human TOP2 cleavage *in vitro* (115). Although none was observed in the nick-flanking region where the Top2 monomers are expected to contact DNA, there was a strong preference for AT base pairs in the region between the nicks (Appendix H). We suggest that the AT-richness may increase the stability of the cleavage intermediate or facilitate the strand separation that precedes the filling of 3'-recessed ends. Although stronger dyad symmetry has been associated with etoposide/VP16-stabilized human TOP2 and yeast Top2 cleavage intermediates (115, 220), there were not enough duplications identified following etoposide treatment to perform a similar analysis in the current study.

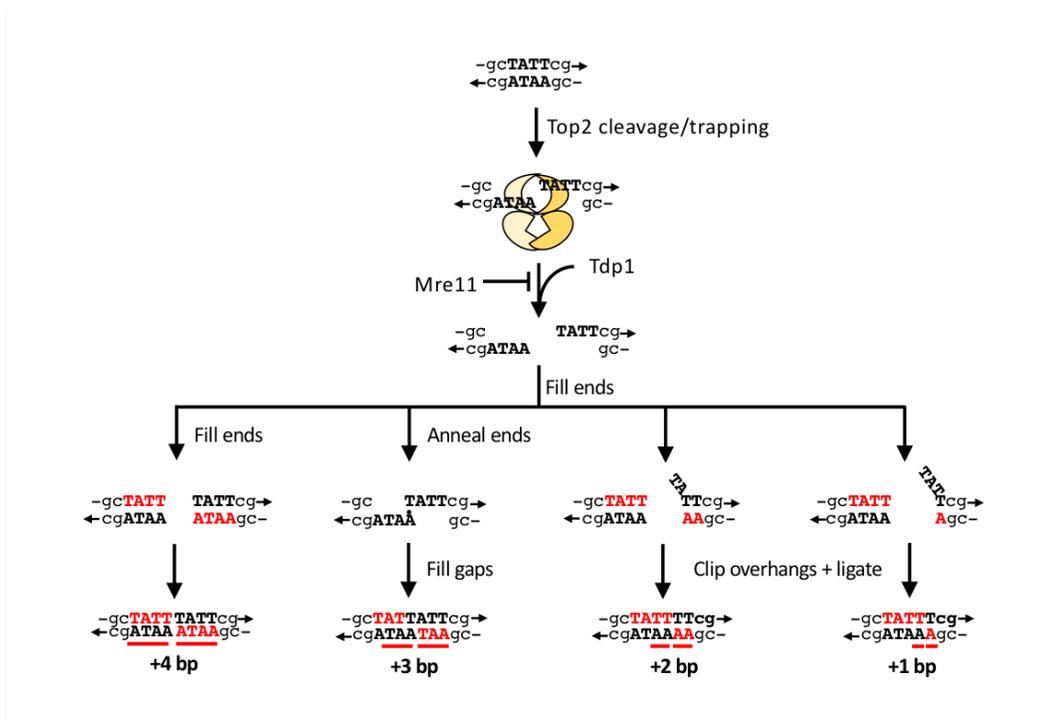


Figure 13: Models for Top2-associated Duplications.

Complementary strands are shown with arrowheads indicating 3' ends; the 4 bp flanked by Top2 nicks is in bold capital letters. Each Top1 monomer (different shades of yellow) uses an active-site tyrosine to nick one DNA strand, forming a covalent phosphotyrosyl bond with the 5' end and creating a recessed 3'-OH end. Mre11 suppresses duplications by nicking the 5'-terminal strand and thereby removing Top2 as part of an oligonucleotide. Tdp1 cleaves the phosphotyrosyl bond to create a clean 5' end. New bases added to recess 3' ends after Top2 removal are in bold red font and duplicated segments are underlined in red. There are multiple ways to generate 1- to 3-bp duplications that depend on where Top2 cleaves the duplex and if/how ends anneal before fill-in reactions.

The Top2 cleavage site depicted in Fig. 13 corresponds to position 1007, where ~10% of the 4-bp duplications occurred (Appendix G). This position was also one of the few sites where 3-bp duplications were detected. The rarity of 3-bp duplications presumably reflects the fact that, unlike the 2- or 4-bp insertions, the addition of 3 bp maintains the correct reading frame of *CAN1* and rarely disables protein function. As illustrated in Fig.13, 3-bp duplications can be readily formed at position 1007 by annealing the terminal nucleotide at each end of the TATT cleavage site and then filling the adjacent gaps. Three-bp duplications could also arise by removing a single nt from one end of the DSB, followed by end filling and ligation. As illustrated in Fig. 13, 2-bp *de novo* duplications require the addition of 4-nt to one end and 2-nt to the other (or alternatively 3 nt to each end). If a Top2 cleavage site encompasses a dinucleotide repeat, however, misaligned pairing between complementary strands would seem the more likely mechanism. Finally, 1-bp insertions can be generated by partial filling of one or both ends, or by misaligned pairing between the ends and subsequent gap filling. Based on the biochemistry of Top2, duplication sizes should not exceed the

distance between the enzyme-generated nicks, and yet 21 of 163 (~13% ) of the duplications were 5 bp. Eight of these occurred at a single position that increased a 5-bp repeat (GGGCT) from two copies to three and most likely reflected DNA polymerase slippage. The remainder were not in repetitive sequence and raise the intriguing possibility that Top2 monomers might occasionally create nicks that are 5 bp instead of 4 bp apart. Such altered spacing may be a specific feature associated with the *top2-FY,RG* allele or other mutations that destabilize the dimer interface.

Before end-filling and ligation can occur, the trapped Top2 must be removed from DNA ends. The increase in *de novo* duplications observed in the absence of Mre11 nuclease activity (*mre11-D56N* background) indicates that, as in higher eukaryotes (208), Top2 is primarily removed by the MRX-Sae2 complex (MRN-CtIP in mammals). In contrast to mammalian cells, however, the nuclease-dead Mre11 protein was compatible with viability; this difference may simply reflect a much lower load of persistent Top2 damage in the much smaller yeast genome or a more robust role of yeast Tdp1 in Top2cc removal from DNA ends. Although Mre11 nuclease activity was not required for viability in the presence of overexpressed Top2-FY,RG, it was necessary for the MRX complex to be present. This requirement does not reflect its essential role in yeast NHEJ, as NHEJ was dispensable for survival. In this context, the MRX complex may be required for mediating an appropriate checkpoint response (286).

*In vitro*, Mre11 nicks 15-40 nt from a 5'-blocked end (287), which precludes the end filling required for *de novo* duplications. 5'-end cleavage also facilitates more extensive end resection and commits repair to HR (288), although microhomology-mediated end joining (MMEJ) is an alternative outcome if HR is not possible (289). It should be noted that the mutation assay done here was in haploid cells, where the only recombination option was the identical sister chromatid. The dramatic increase in *de novo* duplications in the absence of Mre11 nucleolytic activity indicates that MRX is the primary remover of trapped Top2, as it is in mammalian cells. Loss of Tdp1 had the reverse effect and duplications were virtually eliminated. Although Tdp1 was originally identified as a phosphodiesterase that removes 3'-linked peptides from DNA (183), it also has been implicated in removal of 5'-linked peptides and its loss confers etoposide sensitivity (190). It should be noted that yeast does not have a protein analogous to the TDP2 protein of mammals, which also is important for TOP2 removal (189) and suppresses chromosome rearrangements by creating DSB termini that are substrates for NHEJ (224).

Even though Mre11 and Tdp1 are the major proteins involved in the removal of stabilized Top2cc in yeast, cells have evolved multiple pathways to prevent associated cytotoxic events. The role of Wss1, a yeast metalloprotease, and Rad2, a nuclease involved in nucleotide excision repair, were investigated as alternative pathways for Top2cc removal and contributors to the formation of Top2-FY, RG duplications (170, 190, 280, 290). Loss of either Wss1 or Rad2 in

the presence of the pDED1Top2-FY,RG plasmid had no effect on the Can-R mutation rate or spectrum, indicating that neither affects the formation of the duplication events. Alternative pathways to Wss1 involved in degradation of the Top2cc are the 26S proteasome and Ddi1. The 26S proteasome is a large protease complex that degrades ubiquitinated Top2cc (162), while Ddi1 is a novel aspartate protease recently identified in yeast that can remove polyubiquitinated Top2cc (159). Redundancy of these pathways with Wss1 may explain why an effect on duplication rate was not observed in the *wss1* $\Delta$  single mutant. Future studies in which Wss1 activity is lost in addition to either the 26S proteasome using protease inhibitors or a *wss1* $\Delta$  *ddi1* $\Delta$  double mutant may reveal a role of Wss1 in Top2cc removal.

Previous studies have implicated Rad2 in the removal of stabilized Top2cc since loss of *TDP1* or *RAD2* conferred etoposide sensitivity. There was an additive effect in the double mutant (190), suggesting that Rad2-dependent processing of Top2cc is independent of Tdp1. Although Rad2 has been implicated in the removal of Top2cc, the persistence of duplications in *rad2* $\Delta$  mutants suggest that Rad2 does not affect the formation of these events. Since Rad2 is able to remove 5' tails through endonucleolytic cleavage, it may act redundantly to the MRX pathway in the removal of the stabilized Top2cc (291). Determining the effects of a *rad2* $\Delta$  *mre11-D56N* double mutant on the duplication rate would determine if these pathways are redundant. If so, we would expect a further increase in the duplication rate relative to the *mre11-D56N* single mutant.

The broader implications of the Top2-dependent mutation signature described here are two-fold and derive from the highly conserved biochemistry of Top2 and subsequent repair mechanisms. First, the stabilization of the covalent-cleavage intermediate by chemotherapeutic drugs, or the presence of an appropriate mutant TOP2 protein, is expected to have a similar consequence in mammalian cells. Insertions are much more likely to disrupt gene function than are base substitutions, with even a very low level having detrimental consequences. This may contribute to secondary malignancies that arise following the clinical use of TOP2 inhibitors and the corresponding genomes would be expected to have a distinctive mutation signature. Mutations in TOP2 or protein overproduction could also be potential drivers of tumorigenesis. In this regard, we note that the yeast Top2 signature identified here matches insertion-deletion signature 17 (ID17) in human cancers, which is comprised primarily of 4-bp *de novo* duplications (292, 293). A second implication of Top2-associated mutagenesis is that it provides a mechanism for the birth of repetitive sequences, as demonstrated here in yeast strains that overproduce the WT protein. In addition to its essential role during genome duplication, Top2 activity may thus be an important contributor to genome evolution.

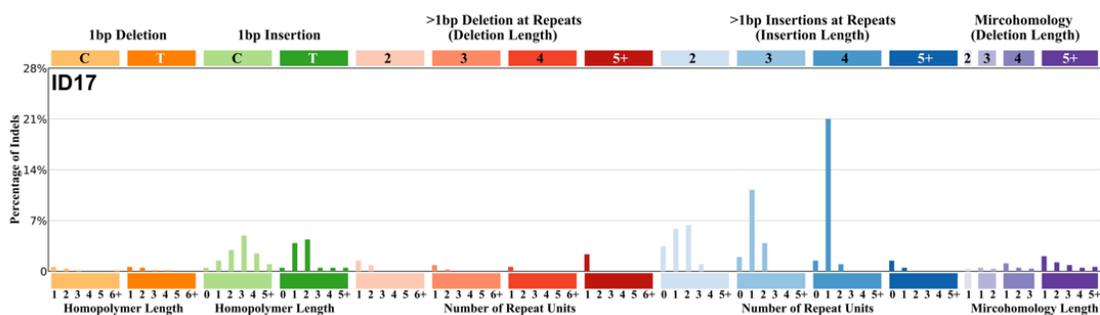
# Chapter 3. Characterization of Yeast Top2 Mutant Equivalent to a Somatic TOP2A Mutation Identified in Human Gastric Cancers and Cholangiocarcinomas

## 3.1 Introduction

Multiple processes cause the somatic mutations found in cancer genomes, but the drivers of these mutations remain unclear. One approach to identifying drivers is through the computational analysis of somatic mutations found in cancer genomes and the characterization of mutation signatures. The Pan-Cancer Analysis of Whole Genomes Consortium of the International Cancer Genome Consortium (PCAWG-ICGC) and The Cancer Genome Atlas (TCGA) have led extensive efforts to accomplish these goals. On the Catalogue of Somatic Mutations in Cancer (COSMIC) website, 90 single-base substitutions (SBS) signatures are listed. Some of these signatures have been associated with exogenous and endogenous exposures to DNA-damaging agents, as well as with defects in DNA-maintenance processes (294-299). For example, germline and somatic mutations that inhibit the tumor suppressor function of *BRCA1* and *BRCA2* are associated with the SBS Signature 3 and can drive tumorigenesis in breast cancers (299).

Recently, 18 small insertion and deletion (ID) signatures were identified (292, 300). ID17 is characterized as 2 to 4 bp insertions in a non-repetitive sequence to create a duplication (Fig. 14). All tumors that contain the ID17 mutation signature are heterozygous for a somatic mutation in topoisomerase 2A (TOP2A), TOP2A p.K743N (293). Furthermore, any tumor with this TOP2A

mutation also contains the ID17 signature. TOP2A p.K743N has been observed in four gastric cancers and three cholangiocarcinomas, none of which are hypermutated. Importantly, the ID17 mutation signature matches the 2 to 5 bp duplication signature associated with the expression of the yeast *top2-F1025Y,R1128G* (*top2-FY,RG*) allele described in Chapter 2 (228). These duplications result from the stabilization of a Top2 intermediate during its catalytic cycle and error-prone repair by NHEJ.



**Figure 14: ID17 Mutation Signature from COSMIC.**

The mutation profile is divided into types of insertion(s) or deletion(s) events at the top of the graph. For each insertion(s) or deletion(s) type, the X-axis is divided into either the homopolymer length or the number of repeat units. The height of the mutation profile bar represents the proportion of one ID mutation type among all the other ID mutation types in the signature. ID17 is classified as 2- to 4-bp insertions at a non-repetitive sequence.

TOP2 is an enzyme that resolves topological structures that arise during cellular processes such as transcription, replication, and chromatin remodeling (35). Human cells have two isoforms: TOP2A and TOP2B. TOP2A is expressed in proliferating cells and is essential for mitotic functions, while TOP2B is ubiquitously expressed and plays a role in gene expression (35). Yeast only have one form of Top2, and it is an essential gene. During the TOP2 catalytic cycle,

the enzyme creates a double-strand break (DSB) to allow for passage of an intact DNA duplex and resolution of topological structures. When TOP2 forms the DSB, each subunit of the homodimer creates a transient phosphotyrosyl bond with the 5' end of the complementary nicked strands, and this intermediate is called a TOP2 cleavage complex (Top2cc). Once the topological structure is resolved, TOP2 re-ligates the DNA to restore genome integrity. However, the creation of a DSB poses a threat to genome stability. Failure of TOP2 to complete its catalytic cycle can result in persistence of the DSB and activate a DNA damage response that leads to cell cycle arrest, senescence or apoptosis (107-109). This has made TOP2 an effective target for commonly used antibiotics and anti-cancer drugs (2, 232).

The association between the TOP2A p.K743N mutant and ID17 is the first report of topoisomerase-associated mutagenesis in human cancers (293). To confirm and explore the mechanism by which TOP2A p.K743N creates ID17, a yeast equivalent mutant was created, Top2-K720N. Expression of the *top2-K720N* allele was associated with the formation of duplications similar to those observed with the Top2-FY,RG mutant and ID17. Additionally, these duplication events were dependent on Top2cc removal by Tdp1 and the non-homologous end-joining (NHEJ) pathway for DSB repair. Although the mechanism for duplication formation is the same in strains containing the Top2-FY,RG and Top2-K720N plasmids, some differences in the genetic requirements and genetic consequences were observed. These differences may reflect the protein domain

that contains the Top2 mutation. While the Top2-FY, RG mutations lie in the C-gate domain of the protein, the Top2-K720N mutation occurs in the DNA binding domain. These results provide compelling evidence that the ID17 mutation signature is associated with TOP2 p.K743N expression and may be a driver of tumorigenesis. These findings additionally have implications for the treatment of tumors with ID17, as these may be more susceptible to topoisomerase poisons.

## **3.2 Materials and Methods**

### **3.2.1 Strains and growth conditions**

YPD (1% yeast extract, 2% Bacto-peptone, 2% dextrose, 250 mg/liter adenine; 2% agar for plates) was used for non-selective growth. Synthetic complete (SC) medium contained 0.15% yeast nitrogen base, 0.5% ammonium sulfate, and 2% dextrose (2% agar added for plates) and was supplemented with all amino acids plus adenine and uracil. Drop-out plates missing one amino acid or base (e.g., SC-Ura medium contained no uracil) were used for selective growth. Canavanine-resistant (Can-R) mutants were selected on SC-Arg plates containing 60µg/ml L-canavanine sulfate. All growth occurred at 30°C.

### **3.2.2 Strain constructions**

A list of all yeast strains is provided in Appendix I. Haploid strains used for mutation analyses were *RAD5* derivatives of *W303* [*ade2-1 his3-11,15 ura3-1 leu2-3,112 trp1-1 can1-100 rad5-G535R*] (249). *DNL4*, *TDP1*, and *TOP1* were deleted by one-step allele replacement using PCR fragments amplified from a plasmid containing a selectable drug resistance marker. The *natMX4* cassette

was amplified from pAG52 (250) and the *loxP-hph-loxP* from pSR955 (254). The nuclease-dead *mre11-D56N* allele was introduced using two-step allele replacement following transformation with the *SphI*-digested pSM444 (253).

All non-integrating plasmids were derived from a YCp50 centromeric plasmid with a *URA3* selectable marker (262, 263). The various *top2* alleles were constitutively expressed from the *DED1* promoter (*pDED1*) (262), and the empty vector and pDED1Top2-FY,RG plasmids were previously used in (228). The *top2-K720N* mutation was introduced into the pDED1Top2 plasmid using site-directed mutagenesis (Gene Wiz, Durham, NC).

### **3.2.3 Mutation rate measurements**

Following transformation, the presence of the plasmid was selected and cells were subsequently maintained on SC-Ura plates. Mutation rates were measured by inoculating single colonies containing the empty vector (EV), *top2-FY,RG*, or *top2-K720N* plasmid into 1 ml of SC-Ura medium and grown to saturation (three days) on a roller drum. Appropriate dilutions were plated onto SC-Ura and SC-Arg+Can plates to determine the number of viable cells and Can-R mutants, respectively, in each culture. Colonies were counted after three days of growth, and mutation rates were calculated using the method of the median (264); 95% confidence intervals (CIs) were determined as previously described (265).

### 3.2.4 *CAN1* mutation spectrum

Independent Can-R colonies were obtained using a pin-plating technique that generates ~100 mini-cultures/plate. Single colonies were inoculated into 1ml of SC-Ura and were incubated overnight on a roller drum. The cells were diluted in water, and a 100-count, flat-tipped custom pinning device was used to transfer ~10<sup>3</sup> cells/pin onto SC-Ura plates. Additionally, cells were spotted onto SC-Arg+Can plates to ensure that there were no pre-existing Can-R mutants. After 3 days of growth, cells were replica plated onto SC-Arg+Can plates and were incubated for an additional 3 days. Single Can-R colonies were inoculated in 150ul of YPD medium, and the cultures were incubated overnight prior to genomic DNA extraction.

The *CAN1* locus from the genomic DNAs of mutants was amplified in a 96-well format using MyTaq DNA polymerase (Bioline). Each PCR product was uniquely barcoded using primers containing 20 nucleotides (nt) of *CAN1*-specific sequence conjugated to 16 nt of PacBio forward and reverse barcodes ([https://github.com/PacificBiosciences/Bioinformatics-Training/blob/master/barcoding/pacbio\\_384\\_barcodes.fasta](https://github.com/PacificBiosciences/Bioinformatics-Training/blob/master/barcoding/pacbio_384_barcodes.fasta)). The amount of each PCR product was estimated on agarose gels, and a similar concentration of each was used to construct the pool for subsequent SMRT sequencing. Following purification of the pooled DNA (GeneJet PCR Purification Kit, ThermoFisher Scientific), SMRT libraries were constructed and sequenced by the Duke Center for Genomic and Computational Biology using the PacBio

Sequel system. Circular consensus sequence (CCS) reads were sorted by barcodes and analyzed using the in-house pipeline SmrtSeqTool; (266), except that Deep Variant was used for variance calling (301). A variant was only considered when the variant allele frequency was  $\geq 60\%$ , the QUAL score was  $\geq 20$ , and the total coverage was  $\geq 10$ .

A portion of Can-R mutants were sequenced using Sanger sequencing. The *CAN1* locus from the genomic DNAs of mutants was amplified in a 96-well format using MyTaq DNA polymerase (Bioline). *CAN1* was subsequently sequenced by GeneWiz (Durham, NC) using three primers: 5'-TTATGAGGGTGAGAATGCGA, 5'-CAGTGGAACCTTTGTACGTCCAA. 5'-TTCTCACAAAGATTCCTTTCTC.

### 3.2.5 Generation of heat maps

MATLAB was used to create the duplication heat maps. The positions of the duplications that occurred in strains expressing *top2-FY, RG* or *top2-K720N* in the WT and *mre11-D56N* backgrounds were imported from Excel files into MATLAB vector objects. Using the position vectors as indices for another matrix that was the same length as *CAN1*, these positions were tabulated. This new matrix contained the number of times a duplication event occurred at each location in *CAN1*. The matrix was then divided by the total number of mutations, so that the values corresponded to a proportion of total events. A MATLAB colormap was created with hotspots appearing red, and locations where

mutations were not observed appearing black. The MATLAB image function was used to visualize the matrices created.

### **3.2.6 RT-qPCR**

Single colonies were inoculated into 5ml of SC-Ura and were incubated for 18h on a roller drum. 500ul of the overnight culture were inoculated into 5ml of fresh SC-Ura and cells were incubated an additional 3.5h to obtain a mid-log phase population (OD<sub>600</sub> 0.3-0.8). Cells were treated with zymolyase to degrade the cell wall, and the RNA was extracted using the TRIzol Reagent protocol (Invitrogen). Contaminating DNA was removed using the “Rigorous DNase treatment” protocol from the DNA-free Kit (Ambion by Life Technologies). RNA concentrations were determined using the Qubit RNA BR Assay Kit (Thermo Fisher Scientific). 2 $\mu$ g of total RNA were used to to synthesize cDNA using the High-Capacity RNA-to-cDNA Kit (Applied Biosystems). PowerTrack SYBR Green Master Mix was used for the RT-qPCR reaction (Applied Biosystems) using *TOP2*-specific primers 5'-AAAGATCGTCGAGAGCTGCG-3' (forward) and 5'-AAGAATGGCGCTTTCTCTGGAT-3' (reverse). RNA from the housekeeping gene *ALG9* served as an internal normalization standard and was amplified using primers 5'-CACGGATAGTGGCTTTGGTGAACAATTAC-3' (forward) and 5'-TATGATTATCTGGCAGCAGGAAAGAACTTGGG-3' (reverse).

### **3.2.7 Statistical methods**

The proportions of mutation types in different strains were compared using a contingency Chi-Square test as appropriate (vassarstats.net); p<0.05 was

considered significant. Rates were calculated using the method of the median (264), and 95% CIs for rates were determined (267). Mutation type rates were calculated by multiplying the total Can-R rate by the proportion of the mutation type in the corresponding spectra. The 95% CIs for each mutation type were determined by jointly considering the CI for the Can-R rate and the CI for the proportion (vassarstats.net). This was done using the “root of the square of the sums” (RSS) or right triangle method (268). Rates obtained in different strain backgrounds were considered significantly different if the respective 95% CIs did not overlap. The lower and upper bounds of the error bars for the qRT-PCR expression were calculated using the equation  $2^{-(\Delta\Delta C_T \pm \text{Standard deviation of the } \Delta\Delta C_T)}$ .

Binomial sampling was used to calculate the number of times a duplication is expected to occur at a single location within *CAN1*. First, the possible Top2 cleavage sites were determined using all the 4-bp duplication positions observed in strains expressing *top2-FY,RG* or *top2-K720N* in the WT and *mre11-D56N* backgrounds; 70 sites were identified. The BINOM.DIST function in Excel was used to calculate the probability that a duplication would occur at a single location 0, 1, 2, etc. times in N trials. For the BINOM.DIST function, a probability of 1/70 was used. The number of trials was N = 67 for pDED1Top2-FY,RG and N = 87 for pDED1Top2-K720N (Appendix J), corresponding to the total number of 4-bp duplications observed in each strain. For example, in the *top2-FY,RG* strain, the probability that a duplication occurred zero times at a given location was

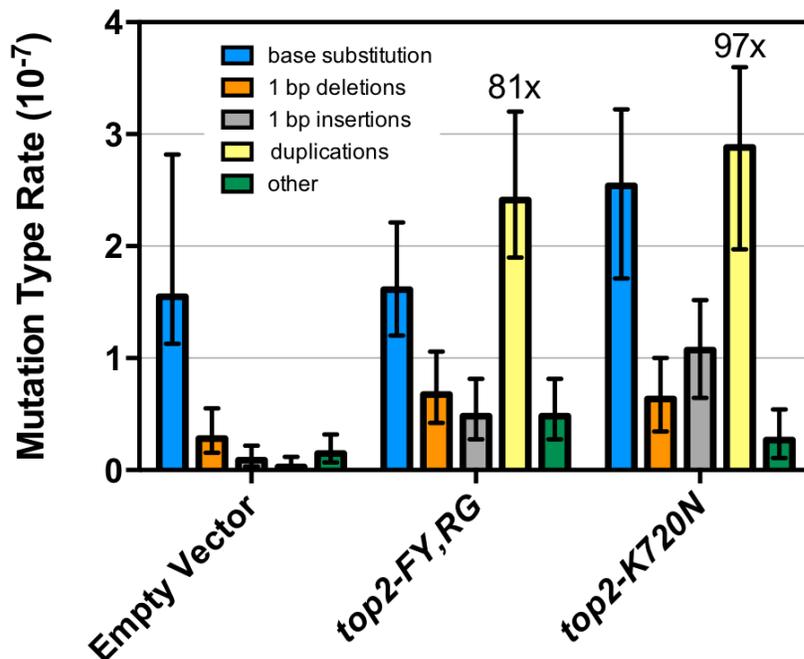
approximately 38%. This probability was then multiplied by the total number of possible locations, 70, to find the expected number of sites with zero duplications (38% x 70 = 27 locations expected to have zero duplications). These expected numbers were compared to the observed duplication distributions using a Chi-square Goodness of Fit test (vassarstats.net);  $p < 0.05$  was considered significant. This method was adapted from (302).

### 3.3 Results

#### 3.3.1 Top2-K720N is associated with *de novo* duplications

Human TOP2A-K743N is associated with the indel mutational signature ID17, which is characterized by duplications of 2- to 4-bp in non-repetitive sequences (292, 293). To investigate whether this *TOP2A* mutant can generate *de novo* duplications *in vivo*, the yeast equivalent of the *TOP2-K743N* mutation, *top2-K720N*, was introduced into the pDED1TOP2 plasmid previously used to characterize the *top2-FY,RG* allele (228). The pDED1Top2-K720N plasmid was introduced into a haploid *TOP2* background and the *CAN1* forward mutation assay was used to measure mutation rates and analyze mutation types. In this assay, any mutation within *CAN1* that results in a non-functional protein confers resistance to the drug canavanine, a toxic arginine analog. The Can-R mutation rate increased ~4-fold in cells containing pDED1Top2-K720N compared to the Can-R mutation rate previously determined in cells expressing the EV plasmid (Appendix K)(228). There was no difference in the Can-R mutation rate in strains expressing the *top2-FY,RG* and *top2-K720N* alleles (Appendix K)(228).

Approximately 42% of the *can1* mutations in strains expressing the *top2-FY,RG* allele were duplications (228). Similarly, 35% of the *can1* mutations associated with expression of *top2-K720N* allele were duplications (75/176 *top2-FY,RG* and 76/221 for *top2-K720N*;  $p = 0.3$  by contingency Chi-square). The proportion of duplications in the pDED1Top2-K720N-containing strain was significantly higher than in the EV strain (76/221 and 2/142, respectively;  $p < 0.0001$ ). When the proportions were converted into mutation type rates, there was a 97-fold increase in the duplication rate in cells containing the *top2-K720N* plasmid compared to the EV plasmid (Fig. 15). The most common duplication size was 4 bp, which matches the distance between the Top2-generated nicks *in vitro* (Appendix L) (275). A complete mutation spectrum of 2 to 5 bp duplications observed in cells containing the pDED1Top2-K720N plasmid is in Appendix M.



**Figure 15: Top2-FY,RG and Top2-K720N are associated with *de novo* duplications.**

Rate of specific mutation types in a *TOP2* strain containing the EV, pDED1Top1-FY,RG, or pDED1Top2-K720N plasmid. Error bars are 95% CIs.

The rate and proportion of the 1-bp insertions also increased in strains expressing *top2-K720N* compared to the EV strain (31/221 and 6/142, respectively;  $p = 0.005$ ) (Fig. 15). Additionally, there was a slight but significant increase in the 1-bp insertions in the *top2-K720N* strain compared to *top2-FY,RG* strain (Fig. 15). However, there was no difference in the proportion of these events (31/221 for *top2-K720N* and 21/176 for *top2-FY,RG*;  $p = 0.6$ ). Although all of the 1-bp insertions in the *top2-FY,RG* and *top2-K720N* strains occurred within short homopolymer runs, their dependence on *top2-FY,RG* or *top2-K720N*

expression indicates that they were likely generated through the same mechanism as the duplications rather than by replication slippage.

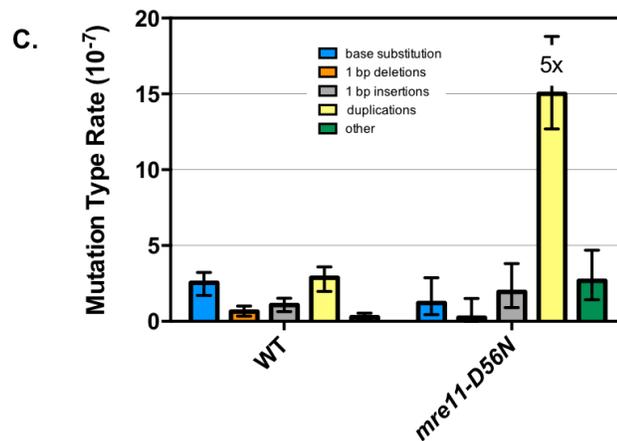
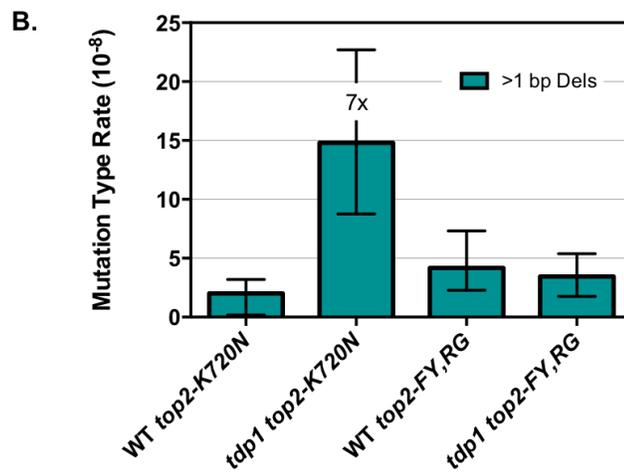
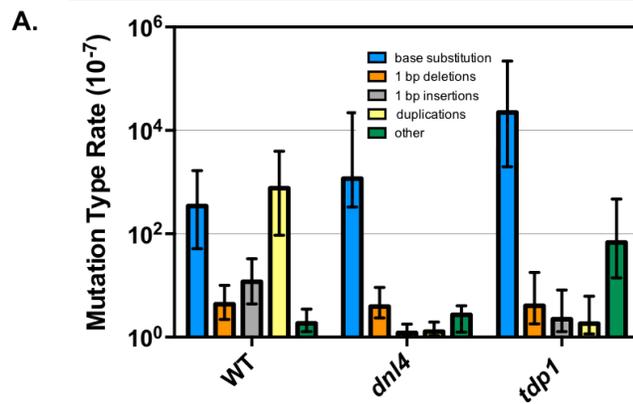
### **3.3.2 Top2-K720N duplications have similar genetic requirements as Top2-FY,RG duplications**

The duplications associated with *top2-FY,RG* expression are dependent on Top2cc removal by Tdp1 and DSB repair by NHEJ. In order to determine if the Top2-K720N duplications have the same genetic requirements, *DNL4*, the ligase required for NHEJ, and *TDP1* were deleted from the *TOP2* background. Although loss of *DNL4* or *TDP1* did not affect the overall Can-R mutation rate (Appendix K), there was a significant proportional reduction in duplications (86/221 WT to 4/159 *dnl4* $\Delta$  or 3/85 *tdp1* $\Delta$ ;  $p < 0.0001$  for both *dnl4* $\Delta$  and *tdp1* $\Delta$ ) and a corresponding decrease in the duplication rate (Fig. 16A). Additionally, there was a reduction in the proportion and rate of 1-bp insertions in the absence of Dnl4 (32/221 WT to 3/159 *dnl4* $\Delta$ ;  $p < 0.0001$ )(Fig. 16A). In the absence of Tdp1, there was a significant reduction in the 1-bp insertions (32/221 WT to 4/85 *tdp1* $\Delta$ ;  $p = 0.03$ ), but the rate of these events was not different relative to WT (Fig. 16A).

Loss of Tdp1 in *top2-K720N*-expressing cells also resulted in a significant increase in the rate of “other” mutation types in the spectra. “Other” mutation types include >1 bp deletions, complex events, and insertions that do not create duplications. There was no significant change in the proportion or rate of the complex events or insertions (complex events: 1/221 WT to 3/85 *tdp1* $\Delta$ ;  $p = 0.12$ ;

insertions: 1/221 WT to 1/85 *tdp1Δ*;  $p = 0.9$ ). However, there was a proportional increase in the >1 bp deletions compared to WT (6/222 WT to 17/85 *tdp1Δ*;  $p < 0.0001$ ) (Appendix N) that corresponded to a 7-fold increase in the rate of these events (Fig. 16B). The >1 bp deletions ranged from 2 to 16 bp in length (Appendix N) and 5/17 were short deletions within repeats, which are generally diagnostic of Top1 activity. In strains expressing *top2-FY,RG*, there was no difference in the proportion or rate of >1 bp deletions when the WT and *tdp1Δ* strains were compared (12/176 WT to 27/205 *tdp1Δ*;  $p = 0.06$ ) (Figure 16B; Appendix N).

Mre11-dependent removal of Top2-FY,RG prevents the formation of the duplications associated with this mutant protein (228). Like cells expressing *top2-FY,RG*, transformants were not viable when the pDED1Top2-K720N plasmid was introduced into *mre11Δ* cells (data not shown). However, transformants were obtained in strains containing the nuclease-dead version of *MRE11*, *mre11-D56N*. Loss of Mre11 nuclease activity in cells containing the pDED1Top2-K720N plasmid increased the overall Can-R mutation rate 2.9-fold compared to WT *top2-K720N* strains (Appendix K). This was associated with a proportional increase in duplications (86/221 WT to 62/87 *mre11-D56N*;  $p < 0.0001$ ), and an ~5-fold increase in duplication rate (Fig. 16C). Combined with the *tdp1Δ* and *dnl4Δ* results, the duplications associated with the expression of *top2-K720N* have similar genetic requirements as the *top2FY,RG* stains.

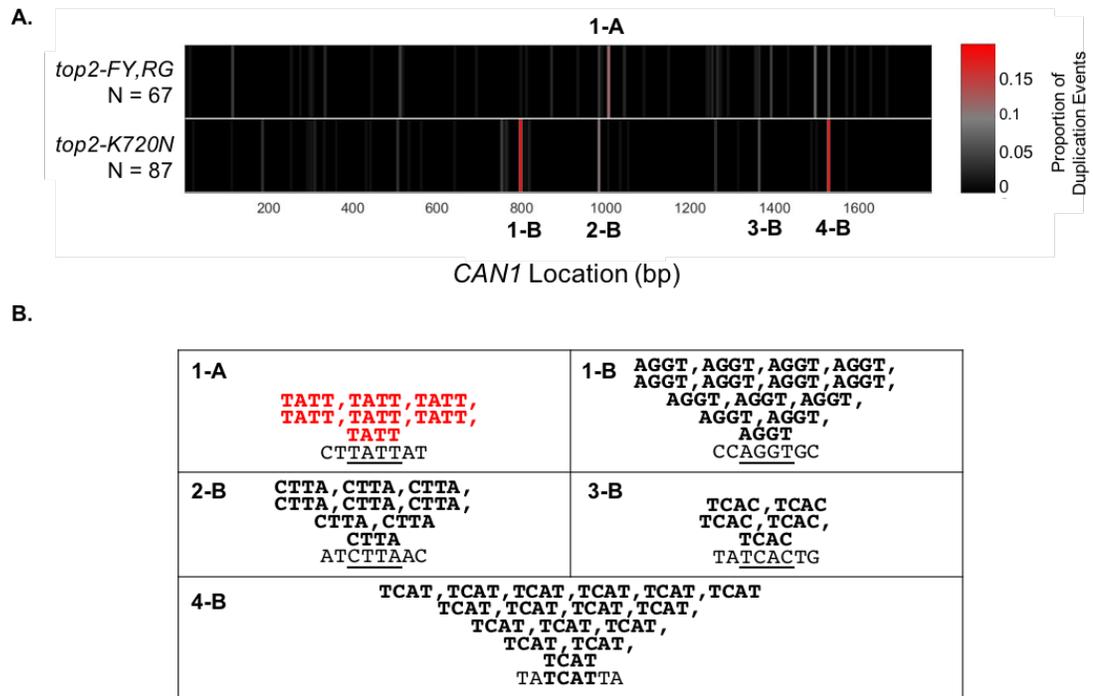


**Figure 16: Genetic Requirements for Top2-K720N Associated Duplications.**

A) Mutation-type rates for WT, *dnl4* $\Delta$ , and *tdp1* $\Delta$  strains containing the pDED1TOP2-K720N plasmid. B) >1 bp deletion rates in WT and *tdp1* $\Delta$  strains containing either the pDED1Top2-K720N or pDED1Top2-FY,RG plasmid. C) Mutation-type rates for the WT and *mre11-D56N* strains containing pDED1Top2-K720N. Error bars are 95% CIs.

### **3.3.3 Distribution of the Top2-K720N duplications differs from Top2-FY,RG duplications**

We previously suggested the occurrence of duplication hotspots in *CAN1* in strains expressing *top2-FY,RG* in the WT and *mre11-D56N* backgrounds (228). Duplication hotspots were also identified in strains expressing *top2-K720N* in the WT and *mre11-D56N* backgrounds. However, the hotspot locations are different in the pDED1Top2-FY,RG- and pDED1Top2-K720N-containing strains (Fig.17A). A composite of the 2- to 5-bp duplications found in the WT and *mre11-D56N* backgrounds containing the pDED1Top2-FY,RG or pDED1Top2-K720N plasmid are in Appendix M.



**Figure 17: Top2-FY,RG and Top2-K720N Duplication Distribution and Hotspots.**

A) Heat maps representing 4-bp duplications in WT and *mre11-D56N* strains containing the pDED1Top2-FY,RG or pDED1Top2-K720N plasmid. The heat maps span the entire *CAN1* gene (1.7 kb) and values correspond to a proportion of total events. Positions where no duplications were observed appear black. As more duplications were observed at a single location, the color changes from black to red. Hotspots determined using binomial sampling are labeled. B) The sequence of each hotspot is displayed at the bottom of the box (underlined), and the duplications are centered above. The labels correspond to the hotspots indicated in A. The hotspots observed in *top2-FY,RG* are in red, and the hotspots observed in *top2-K720N* are black. A complete composite of the 4-bp duplications from pDED1Top2-FY,RG- and pDED1Top2-K720N-containing strains in the WT and *mre11-D56N* backgrounds is in Appendix O.

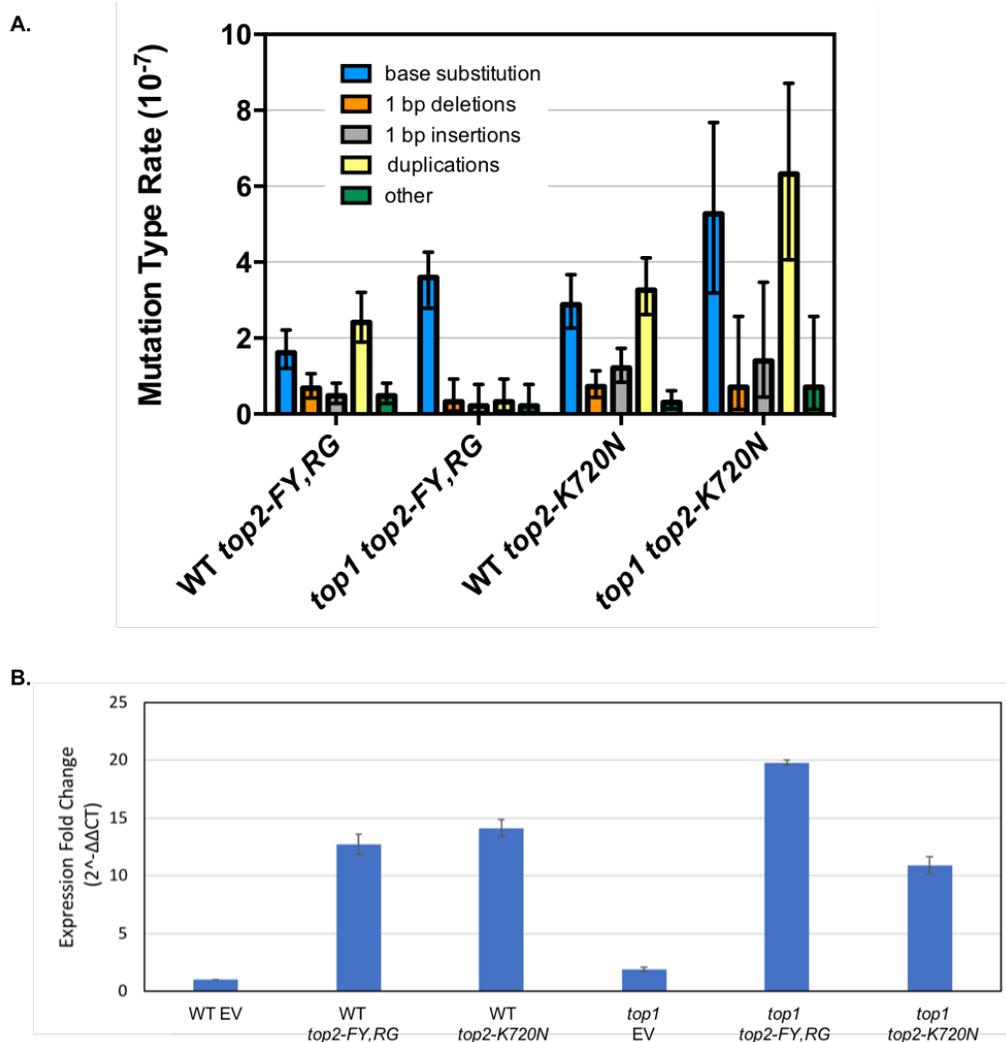
To statistically identify duplication hotspots within *CAN1*, we used binomial sampling (302). The possible Top2 cleavage locations were determined by identifying the positions in *can1* where 4-bp duplications occurred in strains

expressing *top2-FY,RG* or *top2-K720N* in the WT and *mre11-D56N* backgrounds. The analysis was restricted to 4-bp duplications because the Top2 cleavage locations are known only for these events and 70 locations were identified (Appendix O). If mutations are randomly distributed, then the number of times that a duplication occurs at single position in *CAN1* 0, 1, 2, etc. times should follow a binomial distribution. The expected number of sites at which a duplication occurred 0, 1, 2, etc. times was compared to the observed number using a Chi-Square Goodness of Fit Test. For both the *top2-FY,RG* and *top2-K720N* strains, the distribution of duplications in *CAN1* was not randomly distributed (pDED1Top2-FY,RG:  $p = 0.03$ ; pDED1Top2-K720N:  $p < 0.0001$ ) (Appendix J). The binomial distribution was also used to determine how many duplications would have to occur at a single location to be considered a hotspot. A p-value less than 0.01 was considered significant (Appendix J). Therefore, a duplication needs to occur  $\geq 5$  times at a single location in the pDED1Top2-FY,RG- and pDED1Top2-K720N-containing strains to be considered a hotspot.

There was one duplication hotspot in the *top2-FY,RG* strains. This hotspot was located at position 1007 (**TATT** to **TATTTATT**), where approximately 10% of the 4-bp duplications occurred (7/67) (Fig. 17). In the pDED1Top2-K720N strains, there were four hotspots: nt 798, 984, 1364, and 1529 (Fig. 17). Sixteen percent (14/87) of the 4-bp duplications occurred at position 798; 10% (9/87) at position 984; 6% (5/87) at position 1364; and 18% (16/87) of the events at position 1529 (Fig 17).

### **3.3.4 Top1 is required for Top1-FY,RG associated duplications but not Top2-K720N duplications**

Topoisomerase 1 (Top1) is a type IB topoisomerase that resolves positive and negative supercoils during cellular processes such as replication and transcription. Loss of Top1 did not affect the mutation rate in strains expressing *top2-FY,RG* (Appendix K). However, there was a proportional reduction in the duplications (75/176 WT to 3/44 *top1Δ*;  $p < 0.0001$ ), as well as a significant decrease in the duplication rate (Fig 18A). In the *top2-K720N* strain, however, there was no significant difference in either the proportion or rate of duplications (86/221 WT to 18/41 *top1Δ*;  $p = 0.7$ ) (Appendix K).



**Figure 18: Effects of Top1 Loss on Top2-FY,RG and Top2-K720N Associated Duplications.**

A) Mutation-type rates for WT and *top1*Δ strains containing the pDED1Top2-FY,RG or pDED1Top2-K720N plasmid. The error bars are 95% CIs. B) Relative *TOP2* transcript level in WT and *top1*Δ strains containing the EV, pDED1Top2-FY,RG, or pDED1Top2-K720N plasmid. The results were normalized to WT EV. The upper and lower bounds of the error bars were calculated using the standard deviation of the  $\Delta\Delta C_T$  value.

Previously, loss of Top1 was shown to decrease plasmid expression in yeast (303). If loss of Top1 results in low *top2-FY,RG* expression, then a

decrease in the duplication rate would be expected. RT-qPCR was used to measure *top2-FY,RG* and *top2-K720N* transcript levels in the WT and *top1Δ* backgrounds. These plasmids were maintained in a *TOP2* background, and previous experiments showed an ~10-fold increase in overall Top2 protein levels when the plasmids were present (228). Therefore, if *top2-FY,RG* and *top2-K720N* are being expressed in the absence of Top1, then an ~10-fold increase in the transcript relative to the EV control is expected. In the WT cells, there was a 12- and 14-fold increase in *TOP2*-specific RNA when *top2-FY,RG* or *top2-K720N* was expressed, respectively (Fig. 18B). Similarly, there was a 20- and 11-fold increase in *TOP2* RNA in the absence of *TOP1* in *top2-FY,RG* and *top2-K720N* expressing strains (Fig. 18B). Loss of Top1 thus did not affect the expression level of *top2-FY,RG* or *top2-K720N*.

### 3.4 Discussion

In this study, we characterized a novel yeast *top2* allele (*top2-K720N*) that is the equivalent of human TOP2 p.K743N. This TOP2 mutant was identified in gastric cancers and a cholangiocarcinoma. Additionally, it is associated with the ID17 mutation signature, characterized as 2- to 4-bp duplications. Like the *top2-FY,RG* allele previously studied in yeast, the expression of *top2-K720N* resulted in a significant increase in the duplications ranging from 2- to 5-bp in length (228). These events also were dependent on Top2cc removal by Tdp1 and DSB repair by NHEJ. Removal of the Top2cc by the MRX prevented the formation of the Top2-K720N associated duplications. This phenotype is consistent with

Top2-K720N stabilization on the DNA during its catalytic cycle. Furthermore, this study provides biological evidence that the etiology of ID17 is TOP2A p.K743N.

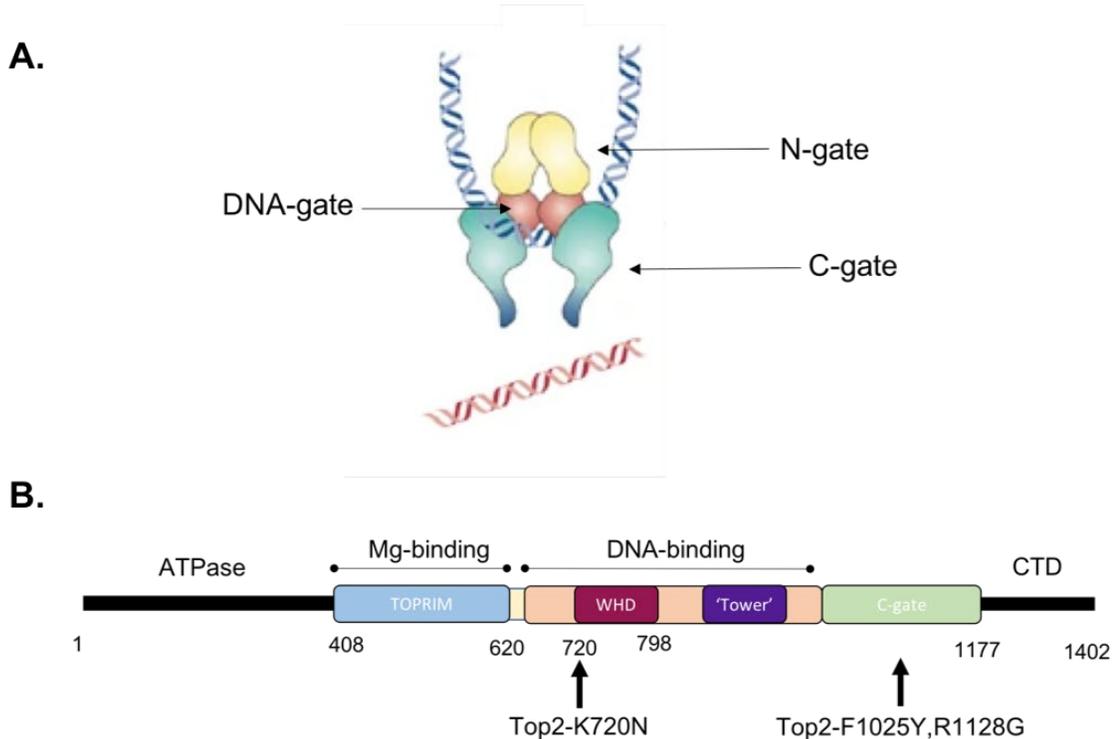
### **3.4.1 Using binomial sampling to define duplication hotspots**

As with the *top2-FY,RG* allele, potential hotspots for duplications were observed within *CAN1* when *top2-K720N* was expressed. In order to better define duplication hotspots, we adapted a binomial sampling method (302). Previous use of this method assumed that insertions and deletions were equally likely to occur at each position in a reporter gene (302). Using this criterion, the probability for a Top2-dependent duplication to occur within *CAN1* was 1/1776. This criterion makes two assumptions. The first is that Top2 can cleave at each position within *CAN1*, and the second assumption is the NHEJ is equally efficient at generating duplications at all 5' overhangs (277). To avoid these assumptions, we defined Top2 cleavage sites as positions within *CAN1* in where 4-bp duplications were observed in pDED1Top2-FY,RG- and pDED1Top2-K720N-containing strains in the WT and *mre11-D56N* backgrounds. The mutations were restricted to 4-bp duplications because the Top2 cleavage positions are unambiguous only for these events, and the sequence of the overhangs determines whether 2- or 3-bp deletions are alternatives. A composite spectrum of the 4-bp duplications identified in the pDED1Top2-FY,RG- and pDED1Top2-K720N-containing strains in the WT and *mre11-D56N* backgrounds can be found in Appendix O. We identified 70 positions where 4-bp duplications occurred, so the probability was adjusted to 1/70 rather than 1/1773.

Using this probability and the number of 4-bp duplications observed in each strain (67 for *top2-FY,RG*; 87 for *top2-K720N*), we were able to calculate an expected value for the number of positions a 4-bp duplication would occur 0, 1, 2, etc. times (Appendix J). These expected values were compared to the observed frequency of 4-bp duplications using a Chi-square Goodness of Fit Test. The observed values for both *top2-FY,RG* and *top2-K720N* were significantly different from the expected values ( $p = 0.03$  *top2-FY,RG*;  $p < 0.001$  *top2-K720N*), indicating that the 4-bp duplications are not randomly distributed across *CAN1* (Appendix J). Furthermore, we also determined the number of times a 4-bp duplication has to occur at a single position for it to be considered a hotspot; a p-value of  $<0.01$  was chosen (Appendix J). Using this threshold, a position where a 4-bp duplication occurred  $\geq 5$  times was considered a hotspot (Appendix J). By this definition, there is one hotspot in the *top2-FY,RG* strains and four hotspots in the *top2-K720N* strains (Fig. 17 and Appendix O). Although both strains create duplications, these hotspots occur at different positions within *CAN1* (Fig. 17A and Appendix O).

The amino acid changes found in Top2-FY,RG occur within the C-terminal dimer interface of the protein, known as the C-gate (228) (Fig. 19). This domain is critical for protein stability and is involved in opening the C-gate to allow for the newly transported DNA duplex (T-segment) to exit the enzyme (281-284). By contrast, the amino acid change in Top2-K720N is located in the winged helix-domain (WHD) within the DNA-binding domain of the protein (Fig. 19) (304). This

domain contains the active site tyrosine and is responsible for DNA cleavage and separation of the gate- or G-segment (Fig. 19). Additionally, the WHD contains a third-dimer interface called the DNA-gate located in the interior of Top2. The DNA-gate helps coordinate the passage of the T-segment through the cleaved G-segment (304). The differences observed in the *CAN1* mutation spectra may reflect sequence specificity in DNA recognition and binding by Top2-K720N. Alternatively, both Top2 mutants contain amino acid changes involved in the accurate passage of the T-segment and stability of the homodimer, which may influence where the Top2cc becomes stabilized on the DNA. Dissociation of the dimer during the catalytic cycle can lead to persistent DSBs in the genome. Further biochemical and structural studies are needed to better understand the consequences of the amino acid changes in Top2 on DNA binding, cleavage and re-ligation.



**Figure 19: Structure of Yeast Top2.**

A) Cartoon schematic of the Top2 cleavage complex. Two DNA duplexes are involved in the Top2 catalytic cycle. The gate- or G-segment (blue) is cleaved by Top2, while the transfer- or T-segment remains intact and is passed through the DSB of the G-segment. The Top2 homodimer has three dimerization interfaces, called the N-gate, DNA-gate, and G-gate. Adapted from (1). B) Schematic of the Top2 domains. The ATPase domain uses ATP binding and hydrolysis to promote T-segment capture, stimulate G-segment cleavage, and coordinate subsequent opening and closing of the gates (285). The TOPRIM domain assists in the formation of the Top2cc by coordinating magnesium ions essential for DNA cleavage (304). The DNA binding domain contains the winged-helix domain (WHD), the 'Tower' domain, and the active site tyrosine. The WHD contains the N-gate and is important for the passage of the T-segment through the DSB. The 'Tower' domain directly interacts with the DNA, flanking the DSB; the precise function of this domain is unclear. The C-gate forms the carboxyl-terminal dimerization interface, which allows the T-segment to exit the enzyme (304). The C-terminal domain (CTD) is where the main differences in the mammalian TOP2 isoforms, TOP2A and TOP2B, are observed. This domain helps modulate enzyme activity (305). The CTD also contains residues specific for modifications such as SUMOylation and ubiquitination (305). The positions of the amino acid

change(s) seen in the Top2-K720N and Top2-FY,RG mutants are indicated by black arrows.

### 3.4.2 Top1 involvement in Top2-FY,RG associated duplications

Top1 is a type IB topoisomerase that nicks DNA and forms a 3' phosphotyrosyl link to resolve topological structures by allowing one DNA end to rotate around the intact strand (9). Top1 is associated with a distinctive mutation signature: 2- to 5-bp deletions within short tandem repeats (306, 307). Loss of Top1 in strains expressing *top2-FY,RG* resulted in a significant decrease in the duplication rate, indicating that Top1 is needed to observe these events (Fig. 18A). By contrast, the duplication rate in strains expressing *top2-K720N* was not affected by the loss of Top1. Top1 and Top2 actively resolve positive and negative supercoils that arise during replication and transcription and can compensate for one another in these functions. Interactions between the Top1 and Top2 proteins have not been observed, however, and the role Top1 plays in the formation of Top2-FY-RG associated duplications is unclear. Even more puzzling is why Top1 would affect Top2-FY,RG associated duplications, but not Top2-K720N associated duplications. Loss of Top1 can negatively affect plasmid expression in yeast, so one hypothesis was that the pDED1Top2-FY,RG plasmid was not being expressed in the *top1Δ* background, resulting in the loss of duplications (303). RT-qPCR was used to measure pDED1Top2-FY,RG and pDED1Top2-K720N transcript levels in the WT and *top1Δ* backgrounds. At least a 10-fold increase was observed in the WT and *top1Δ* strains containing either

pDED1Top2-FY,RG or pDED1Top2-K720N relative to empty vector (Fig. 18B), indicating that loss of Top1 does not affect Top2 expression levels. Since transcript levels do not always equate to protein levels, our next step is to measure Top2 protein levels in pDED1Top2-FY,RG- and pDED1Top2-K720N-containing strains in the WT and *top1* $\Delta$  backgrounds. Additionally, we will assess whether the catalytic activity of Top1 is required for the *CAN1* forward mutation assay.

#### **3.4.4 TOP2 p.K7432N is associated with 6- to 8-bp deletions not observed in Top2-K720N**

In addition to the 2- to 4-bp duplications observed in tumors containing TOP2 p.K743N, there was an increase in small deletions  $\geq 5$ bp that resembled the ID8 mutation signature (293). ID8 occurs in a majority of tumors lacking TOP2A p.K743N and across most cancer types. However, there was a correlation between ID17 mutations and  $\geq 5$ bp deletions, specifically enriched for 6- to 8-bp deletions, in TOP2 p.K743N carriers relative to other tumors. Additionally, these 6- to 8-bp deletions correlated with transcription activity. This led to the proposal of a novel mutation signature ID\_TOP2A that incorporated ID17 and 6- to 8-bp deletions (293). No evidence of ID8 or 6- to 8-bp deletions was observed in strains expressing *top2-K720N*. This may reflect a difference in repair of the stabilized Top2cc between yeast and humans.

### **3.4.5 Conclusion**

Top2-K720N expression in yeast is associated with the formation of duplications. This provides compelling evidence that the TOP2A p.K743N mutation identified in gastric cancer and cholangiocarcinoma is the cause the ID17 mutation signature. This finding has implications for future cancer treatment plans. For example, patients that present with ID17 and TOP2A p.K743N may be more susceptible to Top2 poisons. Additionally, the discovery of multiple Top2 mutations (Top2-FY, RG and Top2-K720N) that contain amino acid changes in different locations, but have similar mutation signatures, suggests that other Top2 variants may also be associated with duplications. As more cancer genomes are sequenced, more Top2 variants may be discovered.

## 4. Conclusions

Top2 is critical for maintaining genome stability by resolving topological structures that arise during processes such as replication, transcription, and repair (9). This enzyme removes supercoils and disentangles DNA through the creation of DSBs, after which Top2 efficiently re-ligates the DNA in an error-free manner. There are endogenous and exogenous factors, however, that impede re-ligation (9). Stabilization of the Top2cc on the DNA poses a risk to genomic integrity and can lead to cell cycle arrest, senescence or apoptosis (107-109). Top2cc trapping has also been exploited in the design of antibiotic and chemotherapeutic drugs used in the clinic (2, 232). One common chemotherapeutic drug that prevents Top2-dependent re-ligation and stabilizes the TOP2cc is etoposide. This thesis characterized mutant forms of the *top2* allele to better understand the genetic consequences of stabilizing the Top2cc on DNA.

The focus of chapter 2 was a novel *top2* allele (*top2-FY,RG*) that confers hypersensitivity to etoposide and encodes a protein that becomes trapped on DNA even in the absence of the drug. Through the use of a forward mutation assay, I found that *top2-FY,RG* expression is associated with the formation of *de novo* duplications. These events vary from 2- to 5-bp in length, with the most common size being 4 bp. The duplications are dependent on the removal of the trapped Top2cc by Tdp1 and DSB repair by NHEJ. Removal of the Top2cc by the MRX complex precludes the formation of duplications by creating a 3' tail and

presumably committing the DNA for repair by homologous recombination. A summary of the proposed mechanism for the formation of the duplications can be found in Fig. 13.

The Top2cc is a bulky protein that requires degradation in order for Tdp1 to access the phosphotyrosyl bond (162). The protein or protein complex that is responsible for degrading the Top2cc prior to formation of the duplication is unclear. There are three candidates: Wss1, the 26S proteasome, and Ddi1 (Fig. 20). In Chapter 2, the role of the metalloprotease Wss1 in the formation of the Top2-FY,RG associated duplications was investigated by deleting *WSS1*. Loss of Wss1 did not affect the *CAN1* mutation or duplication rate, indicating that it is not a significant contributor to these events. The roles of the 26S proteasome and Ddi1 have not been investigated. The 26S proteasome is a large complex that degrades intracellular proteins that have been polyubiquitinated (162). Ddi1 is a recently identified yeast aspartate protease that degrades polyubiquitinated Top2cc and other DNA-protein crosslinks (159, 180, 181). The roles of the 26S proteasome and Ddi1, as well as further investigation of Wss1 in the formation of Top2-dependent duplications can be assessed using a combination of deletion mutants and a proteasome inhibitor (MG132) (159). Wss1 and Ddi1 are non-essential genes that can be deleted from strains. The 26S proteasome has 33 components, six of which are non-essential but downregulate proteasome activity (308). A list of the non-essential proteins in the 26S proteasome is summarized in Table 4. Reduction in the duplication rate in the presence of MG132 or in one

of the deletion backgrounds would reflect disruption of the major degradation pathway for Top2cc. It also is possible that functionally redundant pathways may be responsible for the removal of stabilized Top2cc. If that is the case, then the loss of duplications would require simultaneous disruption of distinct degradation pathways. For instance, if Wss1 has a redundant role in Top2cc degradation, we may see a loss of duplication events in the *wss1Δ ddi1Δ* double mutant, deletion of *WSS1* in combination with a component of the 26S proteasome, or when the *wss1Δ* strain is treated with MG132. Additionally, the loss of Ddi1 in combination MG132 can also be investigated, as well as the loss of all three potential pathways.

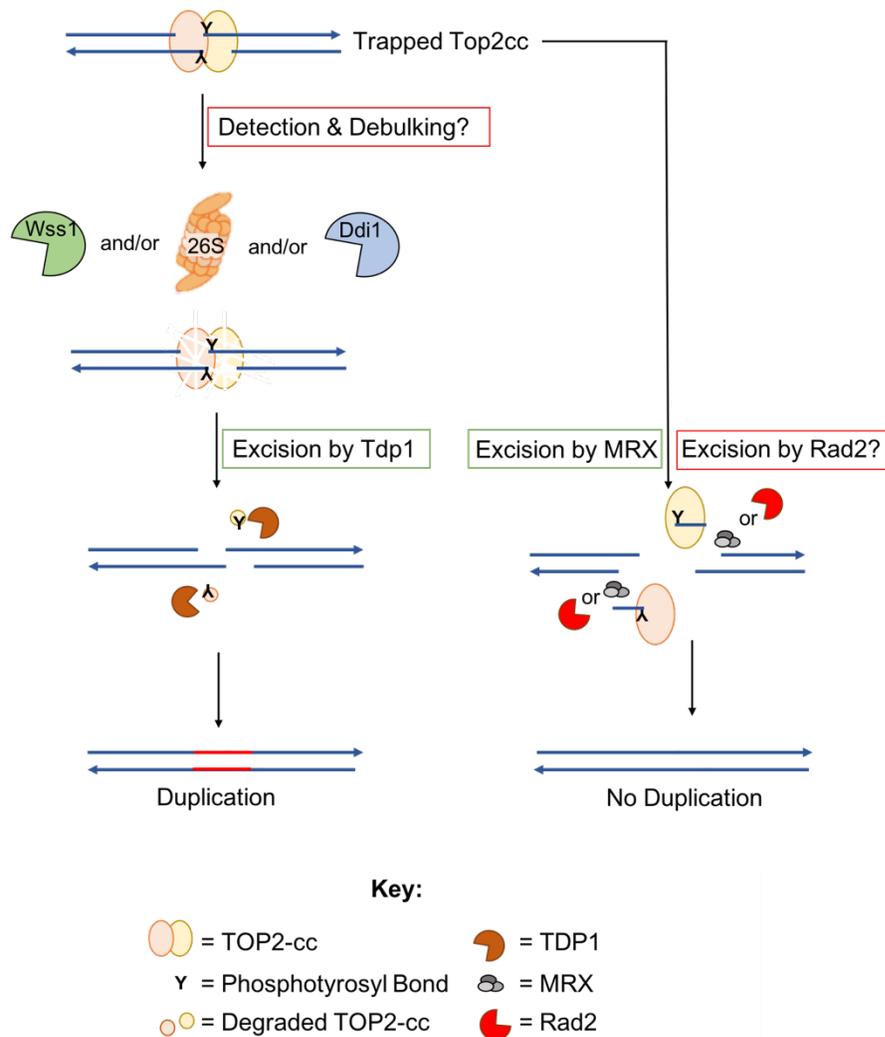
**Table 4: Non-essential Components of the 26S Proteasome.**

<b>Protein</b>	<b>Function</b>
Sem1	Regulatory subunit of the 26S proteasome
Pre9	Alpha 3 subunit of the 20S proteasome
Rpn4	Transcription factor that stimulates expression of proteasome genes
Rpn13	Subunit of the 19S regulatory particle; acts as a ubiquitin receptor
Ubp6	Ubiquitin-specific protease located in the base of the 26S proteasome
Rpn10	Subunit of the 19S regulatory particle; proteasome polyubiquitin receptor

It is important to better understand which pathway is responsible for the degradation of the Top2cc prior to cleavage of the phosphotyrosyl bond by Tdp1 because a combination of proteasome inhibitors and TOP2 poisons have been suggested for clinical use (309). Inhibition of the proteasome increases the half-life of the TOP2cc and enhances the likelihood that irreparable damage will be created in cancer cells (156). Furthermore, the use of protease inhibitors can restore TOP2 levels in drug-resistant cells (156). However, recent studies suggest that this might not be as simple as previously thought. Inhibition of the proteasome in primary splenic mouse B-cells, for example, prevented the proteolytic processing of the stabilized TOP2cc and protected cells from etoposide-induced DNA damage and cell death (167). It is possible that the combination of proteasome inhibitors with Top2 poisons may be advantageous in some cancers but ineffective in others (309).

The prominent protein complex involved in the removal of stabilized Top2cc in higher eukaryotes is MRX/N (208). As demonstrated in Chapter 2, loss of Mre11 nuclease activity results in a significant increase in the rate of duplications, indicating that removal of the Top2cc by MRX precludes the formation of duplications (Fig. 10B; Appendix D). Additionally, MRX-dependent removal does not require the degradation of the Top2cc (160, 162). An alternative protein that may work redundantly with MRX is Rad2, a nuclease involved in nucleotide excision repair that can remove 5' flaps. Previous studies showed additive sensitivity to etoposide in *rad2Δ tdp1Δ* double mutant (190),

suggesting that Rad2-dependent processing of Top2cc is independent of Tdp1. Deletion of *RAD2* in strains expressing *top2-FY,RG* did not affect the Can-R mutation or duplication rate (Appendix D), suggesting that Rad2 is not a major contributor to these events. However, Rad2 may act redundantly to MRX in the removal of Top2cc (Fig. 20). To assess this hypothesis, the duplication rate can be measured in a *mre11-D56N rad2Δ* double mutant. If Rad2 is redundant to MRX, then an even greater increase in the duplication rate would be expected in the double mutant relative to loss of Mre11 nuclease activity alone.



**Figure 20: Potential Mechanisms for Top2cc Degradation and Removal.**

There are two potential pathways for stabilized Top2cc removal. In the first pathway, the Top2cc needs to be degraded prior to removal by Tdp1 and formation of the duplication. There are three candidates for degradation: Wss1, the 26S proteasome, and Ddi1. An alternative pathway involves the MRX complex, in which the Top2cc is released after endonucleolytic cleavage by Mre11 as part of a short oligo. Rad2 is a nuclease that can also remove 5' flaps, suggesting that Rad2 may also be able to remove stabilized Top2cc, but may be redundant to MRX.

Chapter 3 focused on the characterization of *top2-K720N*, which is the yeast equivalent of the human TOP2p.K743N allele identified in several gastric cancers and cholangiocarcinomas (293). TOP2p.K743N is associated with the ID17 mutation signature in COSMIC, characterized as 2 to 4 bp *de novo* duplications, and resembles the mutation signature associated with *top2-FY,RG* expression in yeast (292). Expression of *top2-K720N* was similarly associated with the formation of duplications that range from 2 to 5 bp in length, suggesting that Top2-K720N is stabilized on DNA. As with *top2-FY,RG*, duplications observed in strains expressing *top2-K720N* were dependent on Top2cc removal by Tdp1 and DSB repair by NHEJ. Furthermore, loss of Mre11 nuclease activity resulted in a significant increase in the duplication rate, indicating that removal of the Top2cc by MRX prevents these events. Although the genetic requirements for duplication formation are the same in the *top2-FY,RG* and *top2-K720N* strains, two major differences were observed. First, detection of Top2-FY,RG associated duplications were dependent on Top1 activity, while the Top2-K720N events were not. Further genetic and biochemical studies are necessary to determine the role of Top1 in the Top2-FY,RG duplications. Secondly, the duplication hotspots observed in the *top2-FY,RG* and *top2-K720N* strains occurred at different positions within *CAN1* (Fig. 17).

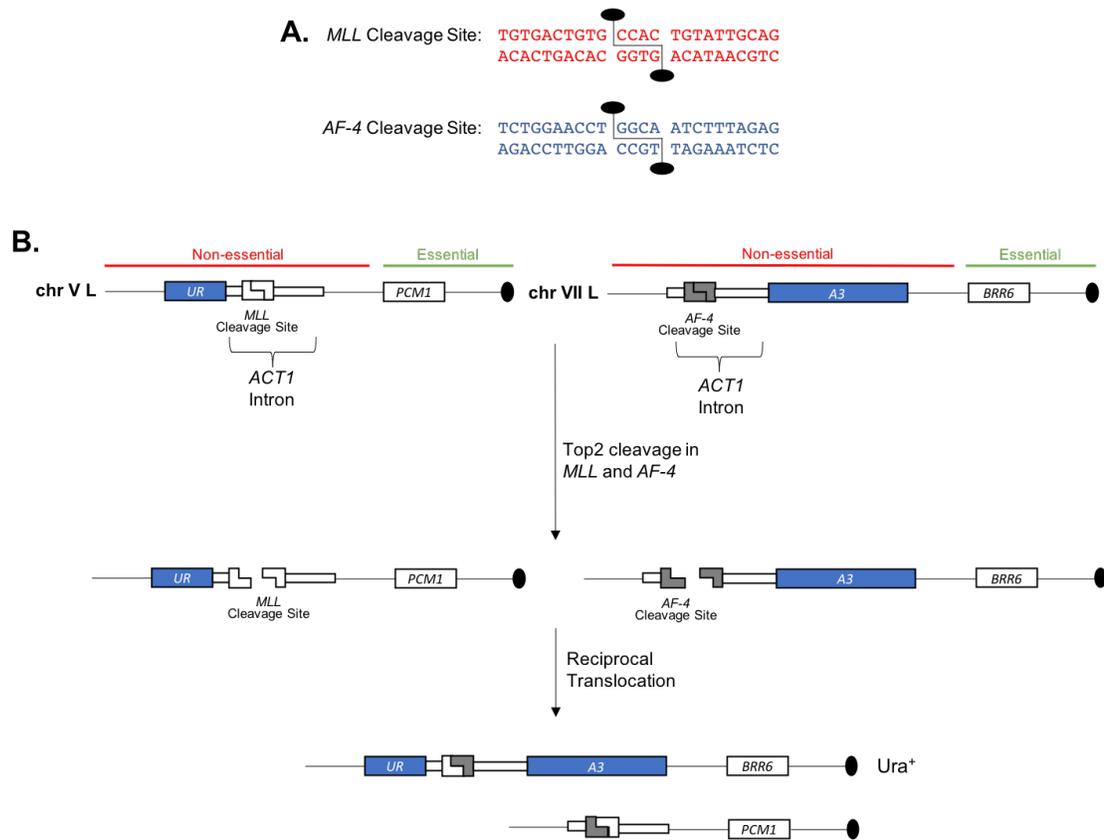
The *CAN1* forward mutation assay was initially used to characterize the *top2-FY,RG* and *top2-K720N* strains because any mutation in *CAN1* that disrupts the function of the protein results in resistance to the drug canavanine, allowing a

relatively unbiased examination of mutations caused by either *top2* allele. However, to further investigate the hotspot differences seen in *top2-FY,RG* and *top2-K720N*, the *lys2* frameshift reversion assay can be used to look at specific duplication events. The duplication hotspots from *CAN1* can be transplanted out-of-frame into the reversion window of *lys2* to create a +1-frameshift allele. If a 4-bp duplication occurs at the transplanted hotspot, then *LYS2* will be in-frame, and the cells will be able to grow on plates lacking lysine. Other +1, +4, and -2 mutations will also give rise to Lys<sup>+</sup> colonies, but their contribution should be reduced in the presence of a hotspot. Furthermore, the reversion window can accommodate the insertion of at least 1kb of sequence (310), making it possible to simultaneously introduce multiple hotspots from *CAN1* into *LYS2*. This approach is summarized in Fig. 21. In strains containing the EV, few, if any, duplications are expected at the hotspots, and any mutations that occur in this strain will be considered background. For the strains containing the pDED1Top2-FY,RG plasmid, duplications are expected to occur at *top2-FY,RG* hotspots identified in *CAN1* but not at the *top2-K720N* hotspots if the hotspots are indeed allele-specific. Similar results are expected for the pDED1Top2-K720N-containing strains with regard to the *top2-K720N* hotspots. The *lys2* reversion assay will also be useful for other genetic tests, such as investigating the role of the proteasome and proteases, as sequencing the reversion window of *lys2* is substantially easier than sequencing *CAN1*.



therapy-induced or primary infant leukemia display translocations within an 8.3 kb breakpoint cluster in the *MLL* gene and are in close proximity to TOP2 cleavage sites mapped *in vitro* (126). The presence of the TOP2 cleavage sites in conjunction with the nonrandom association of etoposide treatment with these events suggests that DSBs induced by TOP2 may initiate *MLL* translocations. An alternative hypothesis for the formation of these translocations involves cleavage by apoptotic nucleases within the breakpoint cluster region of *MLL* (128-130). To further investigate the role of Top2 in the formation of the *MLL* reciprocal translocation *in vivo*, a phenotypic assay in *S. cerevisiae* could be utilized. In this assay, the Top2 cleavage site identified in *MLL* will be inserted into the *ACT1* intron that is fused to a 3' truncated *URA3* gene on the left arm of chromosome V (Fig. 22). The Top2 cleavage site from *AF-4*, one of the partner genes involved in the *MLL* translocation, will be inserted into the *ACT1* intron fused to the 5' truncated *URA3* gene on the left arm of chromosome VII. Both truncated forms of the *URA3* gene are positioned telomere proximal to the last essential gene on the chromosome arm, *PCM1* on chromosome V and *BRR6* on chromosome VII, which also will allow non-reciprocal translocations to be detected. Top2 cleavage at both the *MLL* and *AF-4* sequences will create DSBs with 4-bp 5' overhangs. If a translocation occurs between the *MLL* and *AF-4* sequences, the *URA3* gene will be reconstituted and loss of *ACT1* by splicing will cause the cells containing the translocation to be Ura<sup>+</sup> (Fig. 22). If the reciprocal translocations are caused by Top2 cleavage, then an increase in these events is expected in cells treated

with etoposide or expressing one of the mutant *top2* alleles relative to the vehicle control or empty vector plasmid expression.



**Figure 22: *MLL* Translocation Assay.**

A) Sequences of the TOP2 cleavage sites in the *MLL* and *AF-4* genes identified *in vitro*. B) The *MLL* cleavage site will be introduced into the 5' half of the *ACT1* intron fused to the 3' truncated *URA3* gene (*UR*) located on the left (L) arm of chromosome V (chr V). *PCM1* is the most telomeric essential gene on chr V, and the remainder of the L arm is non-essential. The 5' truncated *URA3* gene (*A3*) will be inserted on the L arm of chr VII and will be fused to the 3' half of the *ACT1* intron containing the *AF-4* cleavage site. *BRR6* is the most telomeric essential gene on chr VII. If Top2 cleaves at the *MLL* and *AF-4* cleavage sites, 4-bp 5' overhangs will be produced. A translocation between the 3' truncated *URA3* containing half the *MLL* site, and the 5' truncated *URA3* containing half the *AF-4* site will reconstitute the full length *URA3* gene. Splicing of the *ACT1* intron will confer a *Ura*<sup>+</sup> phenotype. This figure is not drawn to scale.

In summary, I have characterized two *top2* mutants (*top2-FY,RG* and *top2-K720N*), both of which are associated with the formation of duplications. Top2 is highly conserved from yeast to humans, and the 2 to 5 bp duplications associated with *top2-FY,RG* and *top2-K720N* closely resemble the ID17 mutation signature found in some human cancers. This finding suggests that Top2 may be a driver of tumorigenesis, especially in gastric cancers and cholangiocarcinoma. Additionally, this work has potential implications for future treatment plans. For example, patients who present with ID17 may be more susceptible to TOP2 poisons. Finally, from an evolutionary standpoint, the Top2-associated duplications may provide a mechanism for the birth of repetitive sequences and may be a critical contributor to genome evolution.

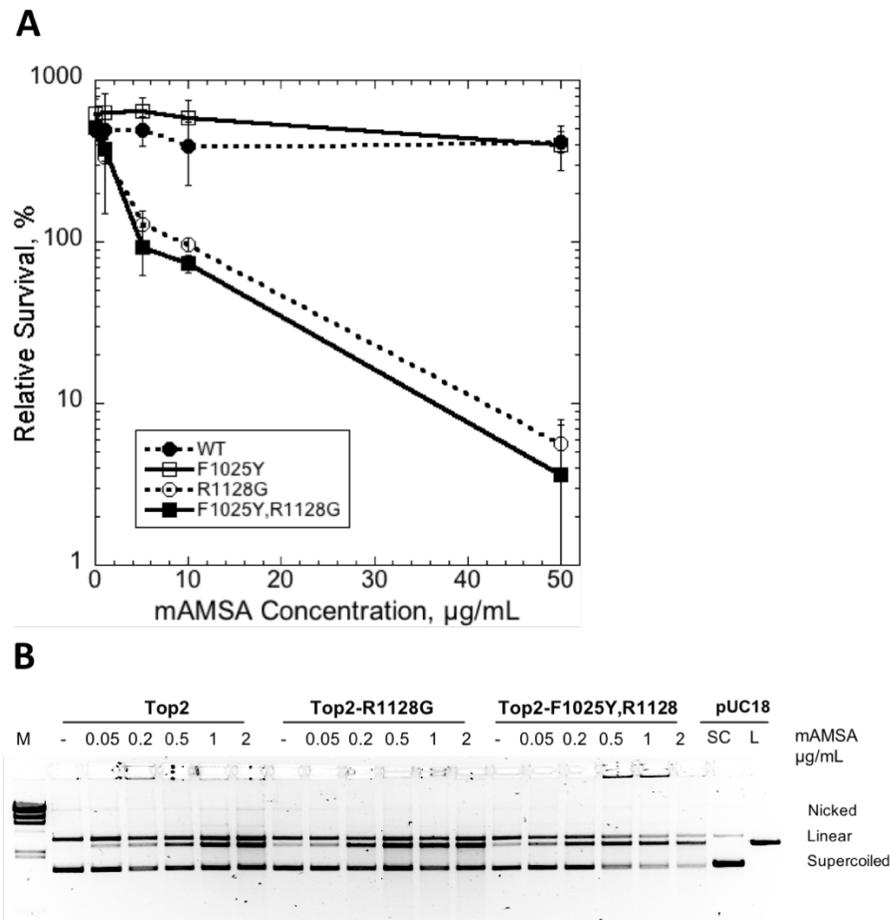
## Appendix A.

**Table 5: List of Yeast Strains from Chapter 2**

Strain	Genotype	Comments/Reference
JELt1	<i>MAT trp1 leu2 ura3-52 pbr1-1122 pep4-3 his3::pGAL1GAL4 top1::LEU2</i>	(234)
JN362at2-4	<i>MATa ura3-52 leu2 trp1 his7 ade1-2 top2-4</i>	(237)
JN394at2-4	<i>MATa ura3-52 leu2 trp1 his7 ade1-2 top2-4 rad52::LEU2</i>	(237)
JN332at2-4	<i>MATa ura3-52 leu2 trp1 his7 ade1-2 top2-4 rad52::TRP1</i>	Nitiss lab collection
YMM10	<i>ura3-52 his3Δ200 leu2Δ1 trp1Δ63 lys2-801<sub>am</sub> ade2-101<sub>oc</sub> pdr18Δ::hisG-URA3-his3; pdr12Δ::hisG; snq2Δ::hisG pdr5::TRP1 pdr10Δ::hisG; pdr15Δ::loxP-kanMX-loxP yor1Δ::HIS3 bat1Δ::HIS3 ycfΔ::HIS3</i>	Provided by Dr. Karl Kuchler
YMM10t2-4	<i>ura3-52 his3Δ200 leu2Δ1 trp1Δ63 lys2-801<sub>am</sub> ade2-101<sub>oc</sub> pdr18Δ::hisG-URA3-his3; pdr12Δ::hisG; snq2Δ::hisG pdr5::TRP1 pdr10Δ::hisG; pdr15Δ::loxP-kanMX-loxP yor1Δ::HIS3 bat1Δ::HIS3 ycfΔ::HIS3 top2-4</i>	Derivative of YMM10 converted to <i>top2-4</i> using <i>pCH1042</i> (237)
CG2009	<i>MATa/MATα lys2-1/lys2-2 tyr1-1/tyr1-2 his7-2/his7-1 leu2Δ/leu2Δ ura3Δ/ura3-1 trp5-d/trp5-c met13-d/met13-c ade5/ADE5 ade2/ade2</i>	Provided by C. Giroux
HK2211 (=SJR3099)	<i>MATa can1Δ::his3Δ3',0 leu2-3,112 his3Δ ura3-1 trp1</i>	Provided by H.L. Klein
SJR3425	<i>MATα can1Δ::his3',0 leu2-3,112 his3Δ ura3-1 trp1</i>	Mating-type switch of SJR3099
MC42-2d (=SJR3659)	<i>Mata CAN1 leu2-3,112 ura3-1 his3-11 ade2-1 trp1-1</i>	(244)
SJR4960	<i>MATa can1Δ::his3Δ3',0 leu2-3,112 his3Δ ura3-1 trp1 leu2Δ::loxP-TRP1-loxP</i>	Transformation of SJR3099
SJR4973	<i>MATa can1Δ::his3Δ3',0 leu2-3,112 his3Δ ura3-1 trp1 leu2Δ::loxP-TRP1-loxP pdr1DBD-CYC8::LEU2</i>	Transformation of SJR4960
SJR5042	<i>MATa CAN1 leu2-3,112 ura3-1 his3-11 ade2-1 trp1-1 hxt13Δ::loxP-TRP-loxP</i>	Transformation of SJR3659
SJR5045	<i>MATα can1Δ::his3',0 leu2-3,112 his3Δ ura3-1 trp1 tdp1Δ::loxP-hph-loxP</i>	Transformation of SJR3425
SJR5046	<i>MATα can1Δ::his3',0 leu2-3,112 his3Δ ura3-1 trp1 leu2Δ::loxP-TRP1-loxP</i>	Transformation of SJR3425

SJR5049	<i>Mat<math>\alpha</math> CAN1 leu2-3,112 ura3-1 his3-11 ade2-1 trp1 hxt13<math>\Delta</math>::loxP-TRP1-loxP tdp1<math>\Delta</math>::loxP-hph-loxP</i>	Spore from SJR5042 x SJR5045
SJR5051	<i>Mat<math>\alpha</math> pdr1DBD-CYC8::LEU2 leu2<math>\Delta</math>::loxP-TRP1-loxP ura3-1 can1<math>\Delta</math>::his3<math>\Delta</math>3',0 his3<math>\Delta</math> trp1</i>	Spore from SJR4973 x SJR5046
SJR5065	<i>Mat<math>\alpha</math> pdr1DBD-CYC8::LEU2 leu2<math>\Delta</math>::loxP ura3-1 can1<math>\Delta</math>::his3<math>\Delta</math>3',0 his3<math>\Delta</math> trp1</i>	Cre recombinase used to remove TRP1 from SJR5051
SJR5074	<i>Mat a CAN1 leu2<math>\Delta</math>::loxP ura3-1 his3-11 ade2-1 hxt13<math>\Delta</math>::loxP-TRP-loxP pdr1DBD-CYC8::LEU2</i>	Spore from SJR5065 x SJR5042
SJR5089	<i>Mat<math>\alpha</math> CAN1 leu2-3,112 ura3-1 his3-11 ade2-1 trp1-1 hxt13<math>\Delta</math>::loxP-TRP1-loxP dnl4<math>\Delta</math>::natMX4</i>	Transformation of SJR5042
SJR5090	<i>Mat<math>\alpha</math> CAN1 leu2-3,112 ura3-1 his3-11 ade2-1 trp1-1 hxt13<math>\Delta</math>::loxP-TRP1-loxP pol4<math>\Delta</math>::loxP-hph-loxP</i>	Transformation of SJR5042
SJR5095	<i>Mat<math>\alpha</math> CAN1 leu2-3,112 ura3-1 his3-11 ade2-1 trp1-1 hxt13<math>\Delta</math>::loxP-TRP1-loxP mre11-D56N</i>	Two step allele replacement using SJR5042
SJR5110	<i>Mat<math>\alpha</math> CAN1 leu2<math>\Delta</math>::loxP ura3-1 his3-11 ade2-1 hxt13<math>\Delta</math>::loxP-TRP1-loxP pdr1DBD-CYC8::LEU2 top2-5</i>	Twp step allele replacement using SJR5074
SJR5143	<i>Mat<math>\alpha</math> CAN1 leu2-3,112 ura3-1 his3-11 ade2-1 trp1-1 hxt13<math>\Delta</math>::loxP-TRP1-loxP yku70<math>\Delta</math>::natMX4</i>	Transformation of SJR5042
SJR1467	<i>Mat<math>\alpha</math> ade2-101<sub>oc</sub> his3<math>\Delta</math>200 ura3<math>\Delta</math>Nco lys2<math>\Delta</math>A746,NR</i>	(241)
SJR1468	<i>Mat<math>\alpha</math> ade2-101<sub>oc</sub> his3<math>\Delta</math>200 ura3<math>\Delta</math>Nco lys2<math>\Delta</math>Bgl,NR</i>	(241)
SJR5123	<i>MATa CAN1 leu2-3,112 ura3-1 his3-11 ade2-1 trp1-1 hxt13<math>\Delta</math>::loxP-TRP-loxP wss1<math>\Delta</math>-loxP-G418-loxP</i>	Transformation of SJR5042
SJR5101	<i>MATa CAN1 leu2-3,112 ura3-1 his3-11 ade2-1 trp1-1 hxt13<math>\Delta</math>::loxP-TRP-loxP rad2<math>\Delta</math>-loxP-G418-loxP</i>	Transformation of SJR5042

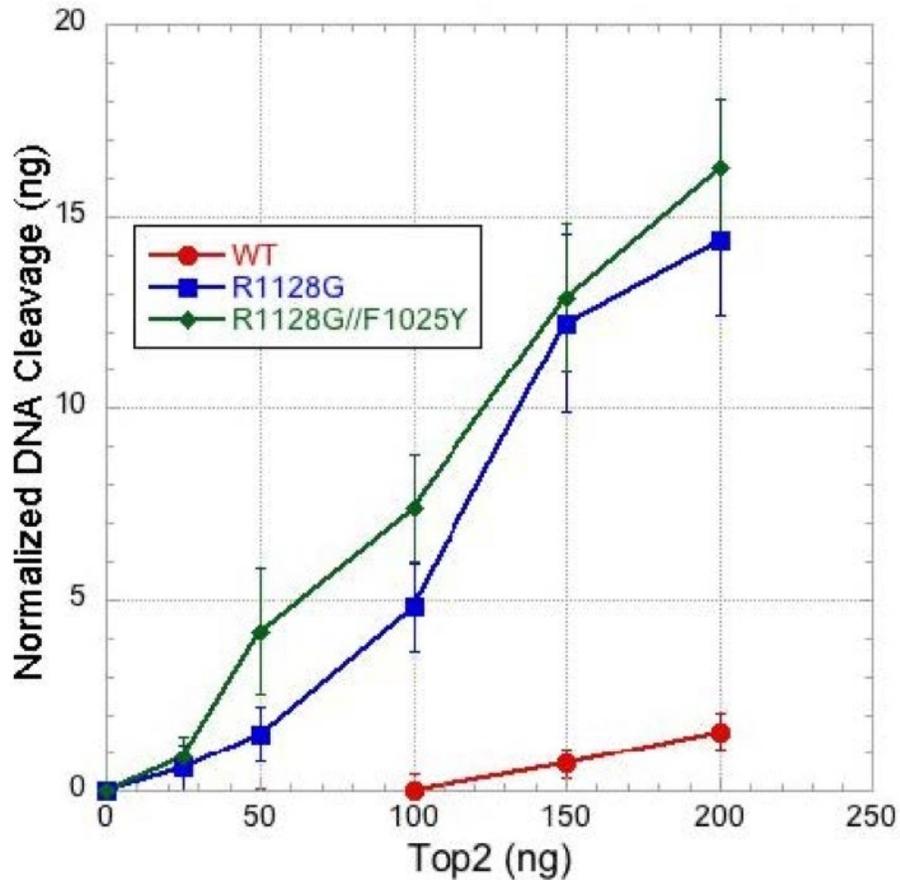
## Appendix B.



**Figure 23: Effects of *m*-AMSA on Viability and Top2 Cleavage Activity.**

A) Cell viability was determined by a yeast colony formation assay in JN362 top2-4 cells carrying either pDED1Top2-F1025Y, pDED1Top2-R1128G, pDED1Top2-F1025Y,R1128G or pDED1Top2 (WT) plasmid after 24-hour exposure to mAMSA at 34°C. Survival is relative to that at the time of mAMSA addition (100%). Yeast cells carrying the top2-R1128G allele were sensitive to mAMSA at concentrations as low as 5 µg/ml. B) Plasmid cleavage assay with purified Top2 proteins. DNA cleavage of 200 ng of negatively supercoiled plasmid DNA (pUC18) was in the presence of mAMSA as indicated above each lane. Each enzyme produced DNA cleavage (linear and nicked products) in an mAMSA dose-response. At equal ng of purified protein, the mutant proteins demonstrated elevated levels of DNA cleavage in the absence of mAMSA as well as a greater response to mAMSA.

## Appendix C.



**Figure 24: Quantitation of Top2 cleavage activity.**

DNA cleavage reactions were carried out using 200ng of pUC18 DNA and were analyzed by agarose gel electrophoresis. DNA bands that co-migrated with linear DNA in the plasmid cleavage assay shown in Fig 2B were quantitated using the GelDoc-It imaging system (UVP). Results shown are the mean  $\pm$ SD of three independent determinations of DNA cleavage for wild-type Top2, Top2-R1128G and Top2-FY,RG proteins.

## Appendix D.

**Table 6: Can-R Median Mutation Rates and Associated Mutation Types.**

Strain	Can-R rate x e-7 (95% CI)	# of cultures	# of mutants with mutation type					Total sequenced
			Base substitution	1-bp deletion	1-bp insertion	Duplications	Other	
WT + EV	2.10 (1.57-3.80)	18	105	19	6	2	10	142
WT + <i>TOP2</i>	1.88 (1.58-2.33)	28	111	14	10	9	11	155
WT + <i>top2-FY,RG</i>	5.66 (4.97-7.20)	34	50	15	21	75	15	176
<i>dnl4Δ</i> + <i>top2-FY,RG</i>	3.32 (2.67-3.73)	18	124	23	7	7	33	194
<i>pol4Δ</i> + <i>top2-FY,RG</i>	3.88 (3.34-4.34)	32	61	12	12	25	16	126
<i>ku70Δ</i> + <i>top2-FY,RG</i>	7.84 (6.47-9.58)	24	53	7	6	3	7	76
<i>tdp1Δ</i> + <i>top2-FY,RG</i>	5.04 (4.39-5.50)	17	130	24	7	11	33	205
<i>mre11-D56N</i> + EV	3.28 (3.12-3.56)	36	114	16	5	5	8	148
<i>mre11-D56N</i> + <i>top2-FY,RG</i>	29.1 (28.1-30.8)	36	38	12	9	87	13	159
<i>wss1Δ</i> + <i>top2-FY,RG</i>	6.99 (6.21-8.29)	24	58	10	17	60	12	157
<i>rad2Δ</i> + <i>top2-FY,RG</i>	7.92 (6.96-9.40)	31	67	15	22	64	15	183



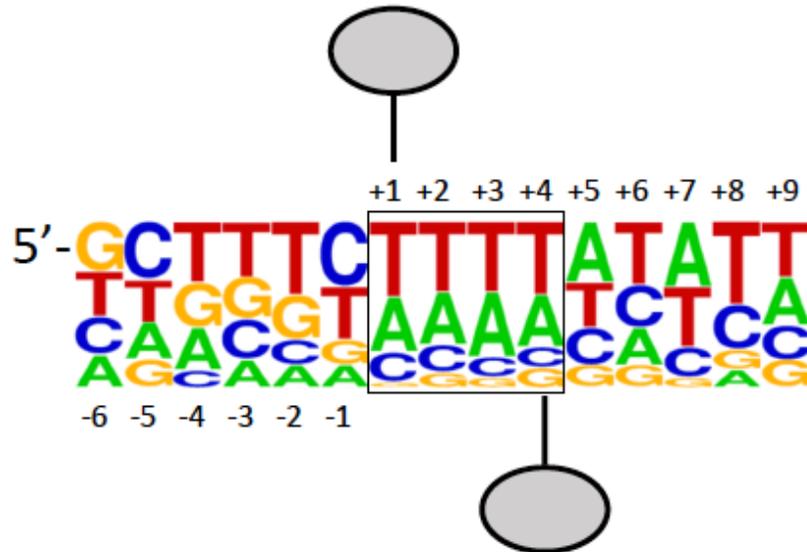
## Appendix F.

**Table 7: Can-R Median Frequencies and Associated Mutation Types in the Presence of Etoposide.**

Strain	Can-R median frequency x 10 <sup>-6</sup> (95% CI)	# of cultures	# of mutants with mutation type					Total sequenced
			Base substitution	1-bp deletion	1-bp insertion	Duplications	Other	
WT + DMSO, 30°	1.38 (1.14-1.60)	44	110	12	3	1	36	162
WT + etoposide, 30°	1.25 (0.92-1.66)	35	104	30	19	36	33	222
WT + DMSO, RT	2.39 (1.08-3.88)	33	194	16	9	4	65	288
WT + etoposide, RT	3.25 (2.52-4.18)	35	172	42	27	18	46	305
<i>top2-5</i> + DMSO, RT	1.02 (0.87-1.33)	48	246	39	8	5	46	344
<i>top2-5</i> + etoposide, RT	1.22 (0.99-1.50)	47	208	29	7	4	49	297



## Appendix H.



**Figure 27: Sequence logo of 4-bp Duplications in WT and *mre11-D56N* strains containing pDED1Top2-FY,RG.**

The logo was obtained by aligning the duplicated sequences (boxed) in 67 4-bp duplications. The lollipops indicate Top2 subunits, which are above/below the nucleotide to which each forms a covalent phosphotyrosyl link. By convention, the site of linkage on the top strand is labeled +1. The logo was constructed as described (22, 23) using website <https://weblogo.berkeley.edu/>.

## Appendix I.

**Table 8: List of Yeast Strains from Chapter 3.**

<b>Strain</b>	<b>Genotype</b>	<b>Comments/Reference</b>
MC42-2d (=SJR3659)	<i>Mat a CAN1 leu2-3,112 ura3-1 his3-11 ade2-1 trp1-1</i>	(244)
SJR5042	<i>MATa CAN1 leu2-3,112 ura3-1 his3-11 ade2-1 trp1-1 hxt13Δ::loxP-TRP-loxP</i>	(228)
SJR5049	<i>Matα CAN1 leu2-3,112 ura3-1 his3-11 ade2-1 trp1 hxt13Δ::loxP-TRP1-loxP tdp1Δ::loxP-hph-loxP</i>	(228)
SJR5089	<i>Mata CAN1 leu2-3,112 ura3-1 his3-11 ade2-1 trp1-1 hxt13Δ::loxP-TRP1-loxP dnl4Δ::natMX4</i>	(228)
SJR5095	<i>Mata CAN1 leu2-3,112 ura3-1 his3-11 ade2-1 trp1-1 hxt13Δ::loxP-TRP1-loxP mre11-D56N</i>	(228)
SJR5186	<i>MATa CAN1 leu2-3,112 ura3-1 his3-11 ade2-1 trp1-1 hxt13Δ::loxP-hph-loxP</i>	Transformation of SJR5186

## Appendix J.

A.

FYRG: B(67, 1/70)			
# of Times an Event Occurs at One Location	Probability	Expected	Observed
0	3.81E-01	27	27
1	3.70E-01	26	32
2	1.77E-01	12	4
3	5.56E-02	4	4
≥4	1.57E-02	1	3

p = 0.03

K720N: B(87, 1/70)			
# of Times an Event Occurs at One Location	Probability	Expected	Observed
0	2.86E-01	20	34
1	3.61E-01	25	25
2	2.25E-01	16	4
3	9.23E-02	6	2
≥4	3.64E-02	3	5

p < 0.001

B.

FYRG: B(67, 1/70)			
# of Times an Event Occurs at One Location	Probability	Expected	Observed
0	3.81E-01	27	27
1	3.70E-01	26	32
2	1.77E-01	12	4
3	5.56E-02	4	4
4	1.29E-02	1	2
5	2.35E-03	0	0
6	3.53E-04	0	0
7	4.45E-05	0	1

K720N: B(87, 1/70)			
# of Times an Event Occurs at One Location	Probability	Expected	Observed
0	2.86E-01	20	34
1	3.61E-01	25	25
2	2.25E-01	16	4
3	9.23E-02	6	2
4	2.81E-02	2	1
5	6.76E-03	0	1
6	1.34E-03	0	0
7	2.24E-04	0	0
8	3.25E-05	0	0
9	4.14E-06	0	1
10	4.68E-07	0	0
11	4.75E-08	0	0
12	4.36E-09	0	0
13	3.64E-10	0	0
14	2.79E-11	0	1
15	1.97E-12	0	0
16	1.28E-13	0	1

**Figure 28: Using Binomial Sampling to Identify Duplication Hotspots in *can1*.**

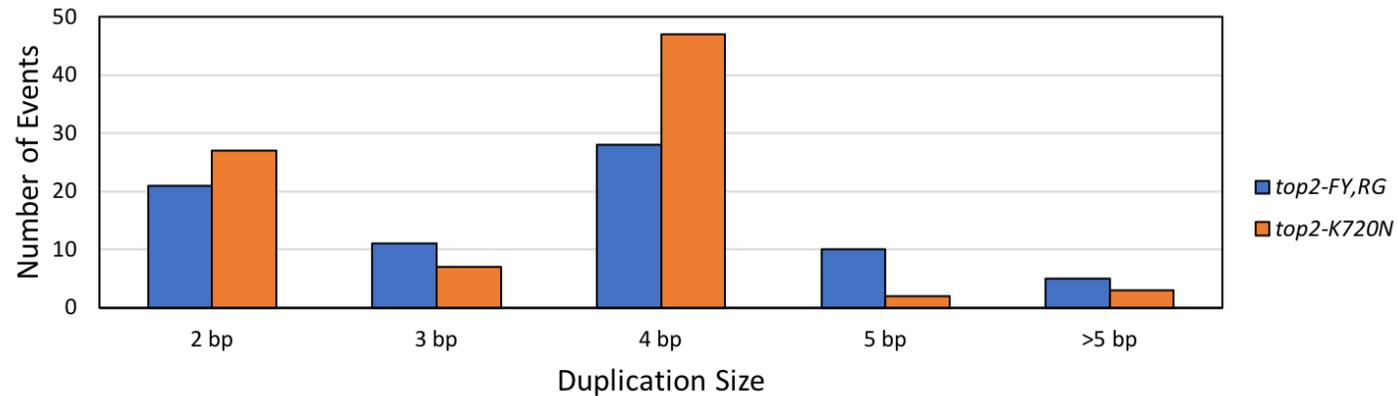
In *CAN1*, there are 70 locations in which a 4-bp duplication was observed at least one time in pDED1Top2-FY,RG- or pDED1Top2-K720N-containing strains in the WT or *mre11-D56N* backgrounds. Therefore, the probability used for the binomial sampling calculations is 1/70. There were 67 4-bp duplications observed in strains expressing *top2-FY,RG* and 87 4-bp duplications observed in strains expressing *top2-K720N*. The binomial sampling probability was calculated using the BINOM.DIST function in Excel. A) To determine if duplications are randomly distributed in *can1*, the expected number of times a duplication occurs at one location was compared to the observed using a Chi-square Goodness of Fit test. The duplications in both Top2-FY,RG and Top2-K720N were not randomly distributed; p = 0.03 in pDED1Top2-FY,RG-containing strains p < 0.001 in pDED1Top2-K720N-containing strains. B) Duplication hotspots were defined as positions where p < 0.01 using binomial sampling. Using this criteria, positions with ≥5 duplications, as indicated by the red highlight, are hotspots in both *top2-FY,RG* and *top2-K720N* strains.

## Appendix K.

**Table 9: Can-R Rates and Associated Mutation Types.**

Strain	Can-R rate x e-7 (95% CI)	# of cultures	# of mutants with mutation type					Total sequenced
			Base substitution	1-bp deletion	1-bp insertion	Duplications	Other	
WT + EV*	2.10 (1.57 – 3.80)	18	105	19	6	2	10	142
WT + <i>top2-FY,RG</i> *	5.66 (4.97 – 7.20)	34	50	15	21	75	15	176
WT + <i>top2-K720N</i>	7.40 (5.40 – 8.72)	24	76	19	32	86	8	221
<i>dnl4Δ</i> + <i>top2-K720N</i>	4.8 (3.67 – 6.01)	18	114	22	3	4	16	159
<i>tdp1Δ</i> + <i>top2-K720N</i>	7.40 (6.31 – 8.45)	36	50	7	4	3	21	85
<i>mre11-D56N</i> + <i>top2-K720N</i>	21.1 (20.2 – 25.6)	36	5	1	8	62	11	87
<i>top1Δ</i> + <i>top2-FY,RG</i>	4.66 (4.17– 5.22)	36	34	3	2	3	2	44
<i>top1Δ</i> + <i>top2-K720N</i>	14.4 (13.0 – 15.5)	35	15	2	4	18	2	41

## Appendix L.



135

**Figure 29: Distribution of Duplication Sizes in the *top2-FY,RG* and *top2-K720N* Strains in the WT Background.**

A majority of the duplications observed in the *top2-FY,RG* and *top2-K720N* WT strains were 2-5 bp in length. There was no significant difference in the distribution of duplication sizes between *top2-FY,RG* and *top2-K720N* ( $p = 0.1$ ; Chi-square contingency test).



## Appendix N.

**Table 10: Can-R Rates and “Other” Mutation Types Associated with *top2-FY,RG* and *top2-K720N***

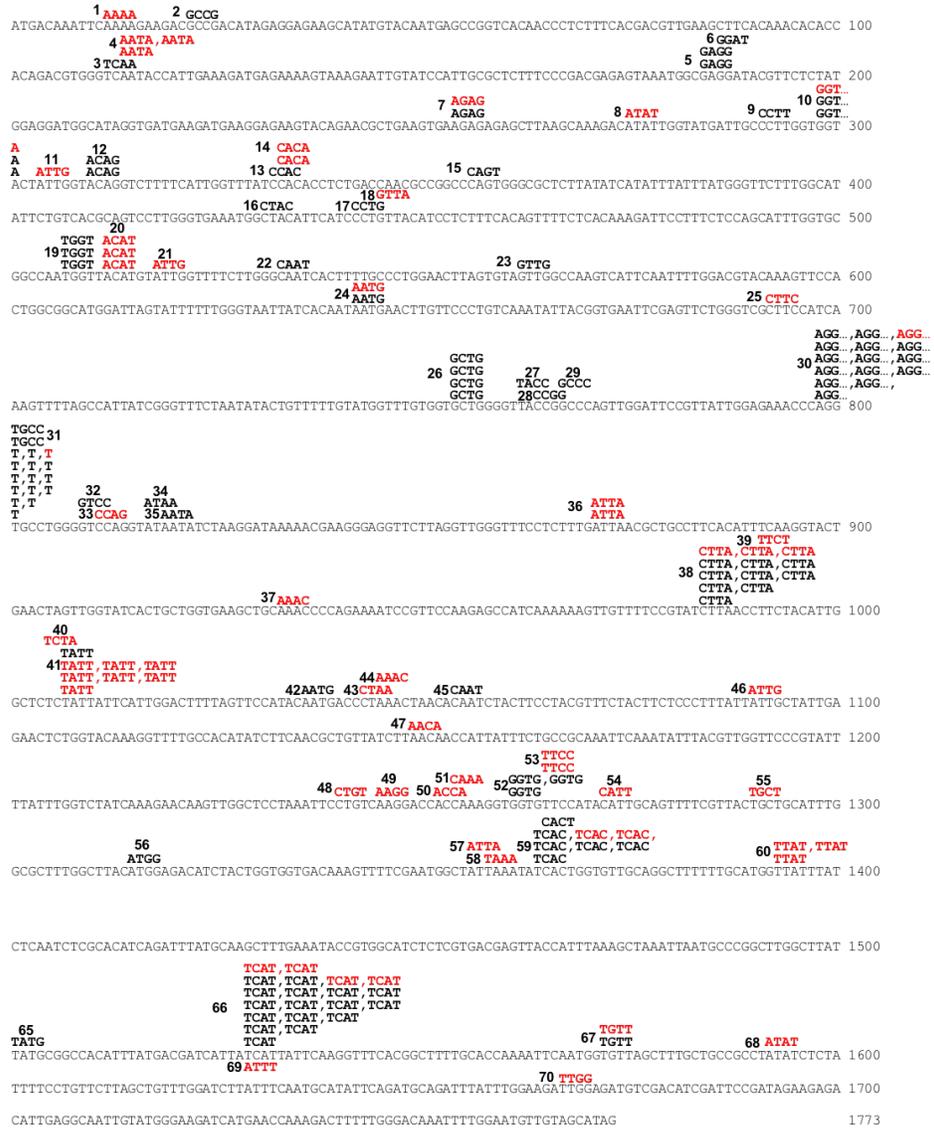
Strain	Can-R rate x e-7 (95% CI)	# of cultures	“Other” Mutation Types			Total sequenced
			> 1 bp deletions	Complex Events	Insertions	
WT + <i>top2-FY,RG*</i>	5.66 (4.97 – 7.20)	34	13	2	0	176
<i>tdp1Δ</i> + <i>top2-FY,RG*</i>	5.04 (4.39 – 5.50)	17	27	6	1	205
WT + <i>top2-K720N</i>	8.37 (7.49 – 9.99)	12	6	1	1	221
<i>tdp1Δ</i> + <i>top2-K720N</i>	7.40 (6.31 – 8.45)	36	17	3	1	85

**Table 11: >1 bp Deletions Observed in the *tdp1Δ* Strains Expressing *top2-K720N***

CAN1 Location	Sequence	Size of Deletion
223	GATGAAGG	2
273	ACATATG	2
291	CCTTGGTGG	2
732	CTGTTTITG	2
1124	CACATATCTT	2
1124	CACATATCTT	2
291	CCTTGGTGGTAC	3
1034	ACAATGACCC	4
978	CCGTATCTAA	5
103	AGACGTGGGTC	7
1035	CAATGACCCTAAA	7
1572	TGTAGCTTTGCTGC	8
1250	CCACCAAAGGTGGT	10
589	TACAAAGTTCCACTG	11
1247	GGACCACCAAAGGTGGT	11
1493	TGGCTTATTATGCGGCC	11
1313	ACATGGAGACATCTACTGGTGG	16

The sequence deleted is in red.

# Appendix O.



**Figure 31: Composite of 4-bp Duplications in *top2-FY, RG* and *top2-K720N* strains**

The 4-bp duplications are positioned above the *CAN1* sequence (black). The duplications in red were observed in the pDED1Top2-FY, RG-containing strains, while the duplications in black were observed in the pDED1Top2-K720N-containing strains. The numbers adjacent to the duplications indicate their position number in the binomial distribution calculations.

## References

1. J. L. Nitiss, DNA topoisomerase II and its growing repertoire of biological functions. *Nature reviews. Cancer* **9**, 327-337 (2009).
2. Y. Pommier, Drugging topoisomerases: lessons and challenges. *ACS Chem Biol* **8**, 82-95 (2013).
3. Y. C. Tse, K. Kirkegaard, J. C. Wang, Covalent bonds between protein and DNA. Formation of phosphotyrosine linkage between certain DNA topoisomerases and DNA. *J Biol Chem* **255**, 5560-5565 (1980).
4. P. O. Brown, N. R. Cozzarelli, Catenation and knotting of duplex DNA by type 1 topoisomerases: a mechanistic parallel with type 2 topoisomerases. *Proceedings of the National Academy of Sciences of the United States of America* **78**, 843-847 (1981).
5. K. Kirkegaard, J. C. Wang, Bacterial DNA topoisomerase I can relax positively supercoiled DNA containing a single-stranded loop. *Journal of molecular biology* **185**, 625-637 (1985).
6. J. J. Champoux, A first view of the structure of a type IA topoisomerase with bound DNA. *Trends Pharmacol Sci* **23**, 199-201 (2002).
7. N. M. Baker, R. Rajan, A. Mondragón, Structural studies of type I topoisomerases. *Nucleic acids research* **37**, 693-701 (2009).
8. P. Forterre, S. Gribaldo, D. Gadelle, M. C. Serre, Origin and evolution of DNA topoisomerases. *Biochimie* **89**, 427-446 (2007).
9. S. M. Vos, E. M. Tretter, B. H. Schmidt, J. M. Berger, All tangled up: how cells direct, manage and exploit topoisomerase function. *Nature reviews. Molecular cell biology* **12**, 827-841 (2011).

10. D. A. Koster, V. Croquette, C. Dekker, S. Shuman, N. H. Dekker, Friction and torque govern the relaxation of DNA supercoils by eukaryotic topoisomerase IB. *Nature* **434**, 671-674 (2005).
11. M. A. Bjornsti, S. H. Kaufmann, Topoisomerases and cancer chemotherapy: recent advances and unanswered questions. *F1000Res* **8** (2019).
12. B. Taneja, B. Schnurr, A. Slesarev, J. F. Marko, A. Mondragón, Topoisomerase V relaxes supercoiled DNA by a constrained swiveling mechanism. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 14670-14675 (2007).
13. A. I. Slesarev *et al.*, DNA topoisomerase V is a relative of eukaryotic topoisomerase I from a hyperthermophilic prokaryote. *Nature* **364**, 735-737 (1993).
14. P. Forterre, DNA topoisomerase V: a new fold of mysterious origin. *Trends Biotechnol* **24**, 245-247 (2006).
15. L. F. Liu, C. C. Liu, B. M. Alberts, Type II DNA topoisomerases: enzymes that can unknot a topologically knotted DNA molecule via a reversible double-strand break. *Cell* **19**, 697-707 (1980).
16. P. O. Brown, N. R. Cozzarelli, A sign inversion mechanism for enzymatic supercoiling of DNA. *Science (New York, N.Y.)* **206**, 1081-1083 (1979).
17. K. Mizuuchi, L. M. Fisher, M. H. O'Dea, M. Gellert, DNA gyrase action involves the introduction of transient double-strand breaks into DNA. *Proceedings of the National Academy of Sciences of the United States of America* **77**, 1847-1851 (1980).
18. K. Mizuuchi, M. Gellert, R. A. Weisberg, H. A. Nash, Catenation and supercoiling in the products of bacteriophage lambda integrative recombination in vitro. *Journal of molecular biology* **141**, 485-494 (1980).

19. T. Hsieh, D. Brutlag, ATP-dependent DNA topoisomerase from *D. melanogaster* reversibly catenates duplex DNA rings. *Cell* **21**, 115-125 (1980).
20. K. B. Tan *et al.*, Topoisomerase II alpha and topoisomerase II beta genes: characterization and mapping to human chromosomes 17 and 3, respectively. *Cancer Res* **52**, 231-234 (1992).
21. P. Grue *et al.*, Essential mitotic functions of DNA topoisomerase IIalpha are not adopted by topoisomerase IIbeta in human H69 cells. *J Biol Chem* **273**, 33660-33666 (1998).
22. M. M. Heck, W. N. Hittelman, W. C. Earnshaw, Differential expression of DNA topoisomerases I and II during the eukaryotic cell cycle. *Proceedings of the National Academy of Sciences of the United States of America* **85**, 1086-1090 (1988).
23. R. D. Woessner, M. R. Mattern, C. K. Mirabelli, R. K. Johnson, F. H. Drake, Proliferation- and cell cycle-dependent differences in expression of the 170 kilodalton and 180 kilodalton forms of topoisomerase II in NIH-3T3 cells. *Cell Growth Differ* **2**, 209-214 (1991).
24. Y. L. Lyu *et al.*, Role of topoisomerase IIbeta in the expression of developmentally regulated genes. *Mol Cell Biol* **26**, 7929-7941 (2006).
25. B. G. Ju *et al.*, A topoisomerase IIbeta-mediated dsDNA break required for regulated transcription. *Science (New York, N.Y.)* **312**, 1798-1802 (2006).
26. R. Madabhushi *et al.*, Activity-Induced DNA Breaks Govern the Expression of Neuronal Early-Response Genes. *Cell* **161**, 1592-1605 (2015).
27. A. Bergerat, D. Gabelle, P. Forterre, Purification of a DNA topoisomerase II from the hyperthermophilic archaeon *Sulfolobus shibatae*. A thermostable enzyme with both bacterial and eucaryal features. *J Biol Chem* **269**, 27663-27669 (1994).

28. A. Bergerat *et al.*, An atypical topoisomerase II from Archaea with implications for meiotic recombination. *Nature* **386**, 414-417 (1997).
29. S. B. Malik, M. A. Ramesh, A. M. Hulstrand, J. M. Logsdon, Jr., Protist homologs of the meiotic Spo11 gene and topoisomerase VI reveal an evolutionary history of gene duplication and lineage-specific loss. *Mol Biol Evol* **24**, 2827-2841 (2007).
30. A. J. Schoeffler, J. M. Berger, DNA topoisomerases: harnessing and constraining energy to govern chromosome topology. *Q Rev Biophys* **41**, 41-101 (2008).
31. L. F. Liu, T. C. Rowe, L. Yang, K. M. Tewey, G. L. Chen, Cleavage of DNA by mammalian DNA topoisomerase II. *J Biol Chem* **258**, 15365-15370 (1983).
32. E. L. Zechiedrich, K. Christiansen, A. H. Andersen, O. Westergaard, N. Osheroff, Double-stranded DNA cleavage/religation reaction of eukaryotic topoisomerase II: evidence for a nicked DNA intermediate. *Biochemistry* **28**, 6229-6236 (1989).
33. F. Mueller-Planitz, D. Herschlag, DNA topoisomerase II selects DNA cleavage sites based on reactivity rather than binding affinity. *Nucleic acids research* **35**, 3764-3773 (2007).
34. J. E. Deweese, A. B. Burgin, N. Osheroff, Human topoisomerase IIalpha uses a two-metal-ion mechanism for DNA cleavage. *Nucleic acids research* **36**, 4883-4893 (2008).
35. Y. Pommier, Y. Sun, S. N. Huang, J. L. Nitiss, Roles of eukaryotic topoisomerases in transcription, replication and genomic stability. *Nature reviews. Molecular cell biology* **17**, 703-721 (2016).
36. J. C. Wang, Cellular roles of DNA topoisomerases: a molecular perspective. *Nature reviews. Molecular cell biology* **3**, 430-440 (2002).

37. L. Postow, C. D. Hardy, J. Arsuaga, N. R. Cozzarelli, Topological domain structure of the Escherichia coli chromosome. *Genes & development* **18**, 1766-1779 (2004).
38. L. Postow, N. J. Crisona, B. J. Peter, C. D. Hardy, N. R. Cozzarelli, Topological challenges to DNA replication: conformations at the fork. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 8219-8226 (2001).
39. J. B. Schwartzman, A. Stasiak, A topological view of the replicon. *EMBO reports* **5**, 256-261 (2004).
40. I. Lucas, T. Germe, M. Chevrier-Miller, O. Hyrien, Topoisomerase II can unlink replicating DNA by precatenane removal. *The EMBO journal* **20**, 6509-6519 (2001).
41. K. Jeppsson *et al.*, The chromosomal association of the Smc5/6 complex depends on cohesion and predicts the level of sister chromatid entanglement. *PLoS genetics* **10**, e1004680 (2014).
42. S. DiNardo, K. Voelkel, R. Sternglanz, DNA topoisomerase II mutant of *Saccharomyces cerevisiae*: topoisomerase II is required for segregation of daughter molecules at the termination of DNA replication. *Proceedings of the National Academy of Sciences of the United States of America* **81**, 2616-2620 (1984).
43. T. Goto, J. C. Wang, Yeast DNA topoisomerase II is encoded by a single-copy, essential gene. *Cell* **36**, 1073-1080 (1984).
44. C. Holm, T. Goto, J. C. Wang, D. Botstein, DNA topoisomerase II is required at the time of mitosis in yeast. *Cell* **41**, 553-563 (1985).
45. C. Holm, T. Stearns, D. Botstein, DNA topoisomerase II must act at mitosis to prevent nondisjunction and chromosome breakage. *Mol Cell Biol* **9**, 159-168 (1989).

46. D. L. Bauer, R. Marie, K. H. Rasmussen, A. Kristensen, K. U. Mir, DNA catenation maintains structure of human metaphase chromosomes. *Nucleic acids research* **40**, 11428-11434 (2012).
47. C. F. Nielsen *et al.*, PICH promotes sister chromatid disjunction and co-operates with topoisomerase II in mitosis. *Nature communications* **6**, 8962 (2015).
48. Z. Daniloski, K. K. Bisht, B. McStay, S. Smith, Resolution of human ribosomal DNA occurs in anaphase, dependent on tankyrase 1, condensin II, and topoisomerase II $\alpha$ . *Genes & development* **33**, 276-281 (2019).
49. M. S. d'Alcontres, J. A. Palacios, D. Mejias, M. A. Blasco, Topolla prevents telomere fragility and formation of ultra thin DNA bridges during mitosis through TRF1-dependent binding to telomeres. *Cell Cycle* **13**, 1463-1481 (2014).
50. N. Takahashi *et al.*, The MCM-binding protein ETG1 aids sister chromatid cohesion required for postreplicative homologous recombination repair. *PLoS genetics* **6**, e1000817 (2010).
51. T. Hirano, T. J. Mitchison, Topoisomerase II does not play a scaffolding role in the organization of mitotic chromosomes assembled in *Xenopus* egg extracts. *The Journal of cell biology* **120**, 601-612 (1993).
52. J. Baxter, L. Aragón, A model for chromosome condensation based on the interplay between condensin and topoisomerase II. *Trends Genet* **28**, 110-117 (2012).
53. K. Kimura, T. Hirano, ATP-dependent positive supercoiling of DNA by 13S condensin: a biochemical implication for chromosome condensation. *Cell* **90**, 625-634 (1997).
54. K. Kimura, V. V. Rybenkov, N. J. Crisona, T. Hirano, N. R. Cozzarelli, 13S condensin actively reconfigures DNA by introducing global positive writhe: implications for chromosome condensation. *Cell* **98**, 239-248 (1999).

55. J. B. Rattner, M. J. Hendzel, C. S. Furbee, M. T. Muller, D. P. Bazett-Jones, Topoisomerase II alpha is associated with the mammalian centromere in a cell cycle- and species-specific manner and is required for proper centromere/kinetochore structure. *The Journal of cell biology* **134**, 1097-1107 (1996).
56. A. B. Lane, J. F. Giménez-Abián, D. J. Clarke, A novel chromatin tether domain controls topoisomerase II $\alpha$  dynamics and mitotic chromosome formation. *The Journal of cell biology* **203**, 471-486 (2013).
57. C. L. Andersen, A. Wandall, E. Kjeldsen, C. Mielke, J. Koch, Active, but not inactive, human centromeres display topoisomerase II activity in vivo. *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* **10**, 305-312 (2002).
58. L. H. Wang, B. Mayer, O. Stemmann, E. A. Nigg, Centromere DNA decatenation depends on cohesin removal and is required for mammalian cell division. *J Cell Sci* **123**, 806-813 (2010).
59. A. M. Farcas, P. Uluocak, W. Helmhart, K. Nasmyth, Cohesin's concatenation of sister DNAs maintains their intertwining. *Molecular cell* **44**, 97-107 (2011).
60. M. Fernández-Casañas, K. L. Chan, The Unresolved Problem of DNA Bridging. *Genes (Basel)* **9** (2018).
61. C. Thrash, A. T. Bankier, B. G. Barrell, R. Sternglanz, Cloning, characterization, and sequence of the yeast DNA topoisomerase I gene. *Proceedings of the National Academy of Sciences of the United States of America* **82**, 4374-4378 (1985).
62. T. Goto, J. C. Wang, Cloning of yeast TOP1, the gene encoding DNA topoisomerase I, and construction of mutants defective in both DNA topoisomerase I and DNA topoisomerase II. *Proceedings of the National Academy of Sciences of the United States of America* **82**, 7178-7182 (1985).

63. S. J. Brill, S. DiNardo, K. Voelkel-Meiman, R. Sternglanz, Need for DNA topoisomerase activity as a swivel for DNA replication for transcription of ribosomal RNA. *Nature* **326**, 414-416 (1987).
64. R. A. Kim, J. C. Wang, Function of DNA topoisomerases as replication swivels in *Saccharomyces cerevisiae*. *Journal of molecular biology* **208**, 257-267 (1989).
65. A. V. Mitkova, E. E. Biswas-Fiss, S. B. Biswas, Modulation of DNA synthesis in *Saccharomyces cerevisiae* nuclear extract by DNA polymerases and the origin recognition complex. *J Biol Chem* **280**, 6285-6292 (2005).
66. V. Gaggioli, B. Le Viet, T. Germe, O. Hyrien, DNA topoisomerase II $\alpha$  controls replication origin cluster licensing and firing time in *Xenopus* egg extracts. *Nucleic acids research* **41**, 7313-7331 (2013).
67. R. E. Gonzalez *et al.*, Effects of conditional depletion of topoisomerase II on cell cycle progression in mammalian cells. *Cell Cycle* **10**, 3505-3514 (2011).
68. R. Bermejo *et al.*, Top1- and Top2-mediated topological transitions at replication forks ensure fork progression and stability and prevent DNA damage checkpoint activation. *Genes & development* **21**, 1921-1936 (2007).
69. G. Abdurashidova *et al.*, Functional interactions of DNA topoisomerases with a human replication origin. *The EMBO journal* **26**, 998-1009 (2007).
70. A. Falaschi, Binding of DNA topoisomerases I and II to replication origins. *Methods Mol Biol* **582**, 131-143 (2009).
71. H. G. Hu, M. Baack, R. Knippers, Proteins of the origin recognition complex (ORC) and DNA topoisomerases on mammalian chromatin. *BMC Mol Biol* **10**, 36 (2009).

72. J. Baxter, J. F. Diffley, Topoisomerase II inactivation prevents the completion of DNA replication in budding yeast. *Molecular cell* **30**, 790-802 (2008).
73. H. J. Edenberg, J. A. Huberman, Eukaryotic chromosome replication. *Annual review of genetics* **9**, 245-284 (1975).
74. S. C. Fields-Berry, M. L. DePamphilis, Sequences that promote formation of catenated intertwinings during termination of DNA replication. *Nucleic acids research* **17**, 3261-3273 (1989).
75. D. Fachinetti *et al.*, Replication termination at eukaryotic chromosomes is mediated by Top2 and occurs at genomic loci containing pausing elements. *Molecular cell* **39**, 595-605 (2010).
76. J. M. Dewar, M. Budzowska, J. C. Walter, The mechanism of DNA replication termination in vertebrates. *Nature* **525**, 345-350 (2015).
77. T. D. Deegan, J. Baxter, M. Ortiz Bazán, J. T. P. Yeeles, K. P. M. Labib, Pif1-Family Helicases Support Fork Convergence during DNA Replication Termination in Eukaryotes. *Molecular cell* **74**, 231-244.e239 (2019).
78. D. R. Heintzman, L. V. Campos, J. A. W. Byl, N. Osheroff, J. M. Dewar, Topoisomerase II Is Crucial for Fork Convergence during Vertebrate Replication Termination. *Cell reports* **29**, 422-436.e425 (2019).
79. E. Gilson, V. Géli, How telomeres are replicated. *Nature reviews. Molecular cell biology* **8**, 825-838 (2007).
80. T. Zhang *et al.*, Looping-out mechanism for resolution of replicative stress at telomeres. *EMBO Rep* **18**, 1412-1428 (2017).
81. L. F. Liu, J. C. Wang, Supercoiling of the DNA template during transcription. *Proceedings of the National Academy of Sciences of the United States of America* **84**, 7024-7027 (1987).

82. S. J. Brill, R. Sternglanz, Transcription-dependent DNA supercoiling in yeast DNA topoisomerase mutants. *Cell* **54**, 403-411 (1988).
83. F. Kouzine, S. Sanford, Z. Elisha-Feil, D. Levens, The functional response of upstream DNA to dynamic supercoiling in vivo. *Nature structural & molecular biology* **15**, 146-154 (2008).
84. A. S. Sperling, K. S. Jeong, T. Kitada, M. Grunstein, Topoisomerase II binds nucleosome-free DNA and acts redundantly with topoisomerase I to enhance recruitment of RNA Pol II in budding yeast. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 12693-12698 (2011).
85. R. Bermejo *et al.*, Genome-organizing factors Top2 and Hmo1 prevent chromosome fragility at sites of S phase transcription. *Cell* **138**, 870-884 (2009).
86. M. Durand-Dubief, J. Persson, U. Norman, E. Hartsuiker, K. Ekwall, Topoisomerase I regulates open chromatin and controls gene expression in vivo. *The EMBO journal* **29**, 2126-2134 (2010).
87. L. Lotito *et al.*, Global transcription regulation by DNA topoisomerase I in exponentially growing *Saccharomyces cerevisiae* cells: activation of telomere-proximal genes by TOP1 deletion. *Journal of molecular biology* **377**, 311-322 (2008).
88. J. M. Pedersen *et al.*, DNA Topoisomerases maintain promoters in a state competent for transcriptional activation in *Saccharomyces cerevisiae*. *PLoS genetics* **8**, e1003128 (2012).
89. C. Nikolaou *et al.*, Topoisomerase II regulates yeast genes with singular chromatin architectures. *Nucleic acids research* **41**, 9243-9256 (2013).
90. M. Roedgaard, J. Fredsoe, J. M. Pedersen, L. Bjergbaek, A. H. Andersen, DNA Topoisomerases Are Required for Preinitiation Complex Assembly during GAL Gene Activation. *PLoS one* **10**, e0132739 (2015).

91. A. Merino, K. R. Madden, W. S. Lane, J. J. Champoux, D. Reinberg, DNA topoisomerase I is involved in both repression and activation of transcription. *Nature* **365**, 227-232 (1993).
92. N. Mondal, J. D. Parvin, DNA topoisomerase IIalpha is required for RNA polymerase II transcription on chromatin templates. *Nature* **413**, 435-438 (2001).
93. K. L. Zobeck, M. S. Buckley, W. R. Zipfel, J. T. Lis, Recruitment timing and dynamics of transcription factors at the Hsp70 loci in living cells. *Molecular cell* **40**, 965-975 (2010).
94. L. Baranello *et al.*, RNA Polymerase II Regulates Topoisomerase 1 Activity to Favor Efficient Transcription. *Cell* **165**, 357-371 (2016).
95. S. Singh *et al.*, Pausing sites of RNA polymerase II on actively transcribed genes are enriched in DNA double-stranded breaks. *J Biol Chem* **295**, 3990-4000 (2020).
96. H. Bunch *et al.*, Transcriptional elongation requires DNA break-induced signalling. *Nature communications* **6**, 10191 (2015).
97. T. Germe, O. Hyrien, Topoisomerase II-DNA complexes trapped by ICRF-193 perturb chromatin structure. *EMBO reports* **6**, 729-735 (2005).
98. N. Mondal *et al.*, Elongation by RNA polymerase II on chromatin templates requires topoisomerase activity. *Nucleic acids research* **31**, 5016-5024 (2003).
99. R. S. Joshi, B. Piña, J. Roca, Topoisomerase II is required for the production of long Pol II gene transcripts in yeast. *Nucleic acids research* **40**, 7907-7915 (2012).
100. A. Bancaud *et al.*, Nucleosome chiral transition under positive torsional stress in single chromatin fibers. *Molecular cell* **27**, 135-147 (2007).

101. C. Lavelle, DNA torsional stress propagates through chromatin fiber and participates in transcriptional regulation. *Nature structural & molecular biology* **15**, 123-125 (2008).
102. C. Lavelle, J. M. Victor, J. Zlatanova, Chromatin fiber dynamics under tension and torsion. *International journal of molecular sciences* **11**, 1557-1579 (2010).
103. J. Salceda, X. Fernandez, J. Roca, Topoisomerase II, not topoisomerase I, is the proficient relaxase of nucleosomal DNA. *The EMBO journal* **25**, 2575-2583 (2006).
104. S. L. French *et al.*, Distinguishing the roles of Topoisomerases I and II in relief of transcription-induced torsional stress in yeast rRNA genes. *Mol Cell Biol* **31**, 482-494 (2011).
105. F. Kouzine *et al.*, Transcription-dependent dynamic supercoiling is a short-range genomic force. *Nature structural & molecular biology* **20**, 396-403 (2013).
106. I. F. King *et al.*, Topoisomerases facilitate transcription of long genes linked to autism. *Nature* **501**, 58-62 (2013).
107. F. d'Adda di Fagagna, Living on a break: cellular senescence as a DNA-damage response. *Nature reviews. Cancer* **8**, 512-522 (2008).
108. W. P. Roos, B. Kaina, DNA damage-induced cell death: from specific DNA lesions to the DNA damage response and apoptosis. *Cancer Lett* **332**, 237-248 (2013).
109. W. P. Roos, B. Kaina, DNA damage-induced cell death by apoptosis. *Trends Mol Med* **12**, 440-450 (2006).
110. A. H. Corbett, N. Osheroff, When good enzymes go bad: conversion of topoisomerase II to a cellular toxin by antineoplastic drugs. *Chem Res Toxicol* **6**, 585-597 (1993).

111. A. Y. Chen, L. F. Liu, DNA topoisomerases: essential enzymes and lethal targets. *Annu Rev Pharmacol Toxicol* **34**, 191-218 (1994).
112. R. D. Anderson, N. A. Berger, International Commission for Protection Against Environmental Mutagens and Carcinogens. Mutagenicity and carcinogenicity of topoisomerase-interactive agents. *Mutation research* **309**, 109-142 (1994).
113. L. R. Ferguson, B. C. Baguley, Topoisomerase II enzymes and mutagenicity. *Environ Mol Mutagen* **24**, 245-261 (1994).
114. S. J. Froelich-Ammon, N. Osheroff, Topoisomerase poisons: harnessing the dark side of enzyme mechanism. *J Biol Chem* **270**, 21429-21432 (1995).
115. G. Capranico, K. W. Kohn, Y. Pommier, Local sequence requirements for DNA cleavage by mammalian topoisomerase II in the presence of doxorubicin. *Nucleic acids research* **18**, 6611-6619 (1990).
116. Y. Pommier, G. Capranico, A. Orr, K. W. Kohn, Local base sequence preferences for DNA cleavage by mammalian topoisomerase II in the presence of amsacrine or teniposide. *Nucleic acids research* **19**, 5973-5980 (1991).
117. R. L. Low, S. Orton, D. B. Friedman, A truncated form of DNA topoisomerase II $\beta$  associates with the mtDNA genome in mammalian mitochondria. *Eur J Biochem* **270**, 4173-4186 (2003).
118. S. Zhang *et al.*, Identification of the molecular basis of doxorubicin-induced cardiotoxicity. *Nat Med* **18**, 1639-1642 (2012).
119. A. M. Wilstermann *et al.*, Topoisomerase II - drug interaction domains: identification of substituents on etoposide that interact with the enzyme. *Biochemistry* **46**, 8217-8225 (2007).

120. C. C. Wu *et al.*, Structural basis of type II topoisomerase inhibition by the anticancer drug etoposide. *Science (New York, N.Y.)* **333**, 459-462 (2011).
121. B. H. Long, S. T. Musial, M. G. Brattain, Single- and double-strand DNA breakage and repair in human lung adenocarcinoma cells exposed to etoposide and teniposide. *Cancer Res* **45**, 3106-3112 (1985).
122. D. Kerrigan, Y. Pommier, K. W. Kohn, Protein-linked DNA strand breaks produced by etoposide and teniposide in mouse L1210 and human VA-13 and HT-29 cell lines: relationship to cytotoxicity. *NCI Monogr*, 117-121 (1987).
123. Y. Pommier, A. Orr, K. W. Kohn, J. F. Riou, Differential effects of amsacrine and epipodophyllotoxins on topoisomerase II cleavage in the human c-myc protooncogene. *Cancer Res* **52**, 3125-3130 (1992).
124. C. A. Felix, Secondary leukemias induced by topoisomerase-targeted drugs. *Biochim Biophys Acta* **1400**, 233-255 (1998).
125. C. A. Felix, Leukemias related to treatment with DNA topoisomerase II inhibitors. *Med Pediatr Oncol* **36**, 525-535 (2001).
126. B. D. Lovett *et al.*, Near-precise interchromosomal recombination and functional DNA topoisomerase II cleavage sites at MLL and AF-4 genomic breakpoints in treatment-related acute lymphoblastic leukemia with t(4;11) translocation. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 9802-9807 (2001).
127. I. G. Cowell *et al.*, Model for MLL translocations in therapy-related leukemia involving topoisomerase II $\beta$ -mediated DNA strand breaks and gene proximity. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 8989-8994 (2012).
128. C. J. Betti, M. J. Villalobos, M. O. Diaz, A. T. Vaughan, Apoptotic triggers initiate translocations within the MLL gene involving the nonhomologous end joining repair system. *Cancer Res* **61**, 4550-4555 (2001).

129. C. J. Betti, M. J. Villalobos, M. O. Diaz, A. T. Vaughan, Apoptotic stimuli initiate MLL-AF9 translocations that are transcribed in cells capable of division. *Cancer Res* **63**, 1377-1381 (2003).
130. A. T. Vaughan *et al.*, Surviving apoptosis: a possible mechanism of benzene-induced leukemia. *Chem Biol Interact* **153-154**, 179-185 (2005).
131. J. M. Barret *et al.*, F14512, a potent antitumor agent targeting topoisomerase II vectored into cancer cells via the polyamine transport system. *Cancer Res* **68**, 9845-9853 (2008).
132. A. C. Gentry *et al.*, Interactions between the etoposide derivative F14512 and human type II topoisomerases: implications for the C4 spermine moiety in promoting enzyme-mediated DNA cleavage. *Biochemistry* **50**, 3240-3249 (2011).
133. O. Bombarde *et al.*, The DNA-Binding Polyamine Moiety in the Vectorized DNA Topoisomerase II Inhibitor F14512 Alters Reparability of the Consequent Enzyme-Linked DNA Double-Strand Breaks. *Mol Cancer Ther* **16**, 2166-2177 (2017).
134. A. Kruczynski *et al.*, Preclinical activity of F14512, designed to target tumors expressing an active polyamine transport system. *Invest New Drugs* **29**, 9-21 (2011).
135. A. Kruczynski *et al.*, F14512, a polyamine-vectorized anti-cancer drug, currently in clinical trials exhibits a marked preclinical anti-leukemic activity. *Leukemia* **27**, 2139-2148 (2013).
136. D. Tierny *et al.*, Phase I Clinical Pharmacology Study of F14512, a New Polyamine-Vectorized Anticancer Drug, in Naturally Occurring Canine Lymphoma. *Clinical cancer research : an official journal of the American Association for Cancer Research* **21**, 5314-5323 (2015).
137. M. A. Smith *et al.*, Report of the Cancer Therapy Evaluation Program monitoring plan for secondary acute myeloid leukemia following treatment with epipodophyllotoxins. *J Natl Cancer Inst* **85**, 554-558 (1993).

138. M. A. Smith, L. Rubinstein, R. S. Ungerleider, Therapy-related acute myeloid leukemia following treatment with epipodophyllotoxins: estimating the risks. *Med Pediatr Oncol* **23**, 86-98 (1994).
139. C. H. Pui, Acute leukemias with the t(4;11)(q21;q23). *Leuk Lymphoma* **7**, 173-179 (1992).
140. J. A. Ross, J. D. Potter, L. L. Robison, Infant leukemia, topoisomerase II inhibitors, and the MLL gene. *J Natl Cancer Inst* **86**, 1678-1680 (1994).
141. C. H. Pui *et al.*, Acute myeloid leukemia in children treated with epipodophyllotoxins for acute lymphoblastic leukemia. *N Engl J Med* **325**, 1682-1687 (1991).
142. P. S. Kingma, A. H. Corbett, P. C. Burcham, L. J. Marnett, N. Osheroff, Abasic sites stimulate double-stranded DNA cleavage mediated by topoisomerase II. DNA lesions as endogenous topoisomerase II poisons. *J Biol Chem* **270**, 21441-21444 (1995).
143. P. S. Kingma, N. Osheroff, Apurinic sites are position-specific topoisomerase II poisons. *J Biol Chem* **272**, 1148-1155 (1997).
144. P. S. Kingma, N. Osheroff, Spontaneous DNA damage stimulates topoisomerase II-mediated DNA cleavage. *J Biol Chem* **272**, 7488-7493 (1997).
145. M. Sabourin, N. Osheroff, Sensitivity of human type II topoisomerases to DNA damage: stimulation of enzyme-mediated DNA cleavage by abasic, oxidized and alkylated lesions. *Nucleic acids research* **28**, 1947-1954 (2000).
146. R. Vélez-Cruz *et al.*, Exocyclic DNA lesions stimulate DNA cleavage mediated by human topoisomerase II alpha in vitro and in cultured cells. *Biochemistry* **44**, 3972-3981 (2005).
147. M. Bigioni *et al.*, Position-specific effects of base mismatch on mammalian topoisomerase II DNA cleaving activity. *Biochemistry* **35**, 153-159 (1996).

148. S. D. Cline, W. R. Jones, M. P. Stone, N. Osheroff, DNA abasic lesions in a different light: solution structure of an endogenous topoisomerase II poison. *Biochemistry* **38**, 15500-15507 (1999).
149. N. Maizels, L. Davis, Initiation of homologous recombination at DNA nicks. *Nucleic acids research* **46**, 6962-6973 (2018).
150. J. E. Dewese, N. Osheroff, The DNA cleavage reaction of topoisomerase II: wolf in sheep's clothing. *Nucleic acids research* **37**, 738-748 (2009).
151. S. A. Nick McElhinny *et al.*, Abundant ribonucleotide incorporation into DNA by yeast replicative polymerases. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 4949-4954 (2010).
152. B. Hiller *et al.*, Mammalian RNase H2 removes ribonucleotides from DNA to maintain genome integrity. *The Journal of experimental medicine* **209**, 1419-1426 (2012).
153. M. A. Reijns *et al.*, Enzymatic removal of ribonucleotides from DNA is essential for mammalian genome integrity and development. *Cell* **149**, 1008-1022 (2012).
154. Y. Wang, B. R. Knudsen, L. Bjergbaek, O. Westergaard, A. H. Andersen, Stimulated activity of human topoisomerases IIalpha and IIbeta on RNA-containing substrates. *J Biol Chem* **274**, 22839-22846 (1999).
155. Y. Wang, A. Thyssen, O. Westergaard, A. H. Andersen, Position-specific effect of ribonucleotides on the cleavage activity of human topoisomerase II. *Nucleic acids research* **28**, 4815-4821 (2000).
156. Y. Ogiso, A. Tomida, S. Lei, S. Omura, T. Tsuruo, Proteasome inhibition circumvents solid tumor resistance to topoisomerase II-directed drugs. *Cancer Res* **60**, 2429-2434 (2000).
157. J. Lopez-Mosqueda *et al.*, SPRTN is a mammalian DNA-binding metalloprotease that resolves DNA-protein crosslinks. *eLife* **5** (2016).

158. B. Vaz *et al.*, Metalloprotease SPRTN/DVC1 Orchestrates Replication-Coupled DNA-Protein Crosslink Repair. *Molecular cell* **64**, 704-719 (2016).
159. N. Serbyn *et al.*, The Aspartic Protease Ddi1 Contributes to DNA-Protein Crosslink Repair in Yeast. *Molecular cell* **77**, 1066-1079.e1069 (2020).
160. Y. Sun *et al.*, Excision repair of topoisomerase DNA-protein crosslinks (TOP-DPC). *DNA repair* **89**, 102837 (2020).
161. H. Zhang, Y. Xiong, J. Chen, DNA-protein cross-link repair: what do we know now? *Cell Biosci* **10**, 3 (2020).
162. Y. Sun, L. K. Saha, S. Saha, U. Jo, Y. Pommier, Debulking of topoisomerase DNA-protein crosslinks (TOP-DPC) by the proteasome, non-proteasomal and non-proteolytic pathways. *DNA repair* **94**, 102926 (2020).
163. Y. Mao, S. D. Desai, C. Y. Ting, J. Hwang, L. F. Liu, 26 S proteasome-mediated degradation of topoisomerase II cleavable complexes. *J Biol Chem* **276**, 40652-40658 (2001).
164. H. Xiao *et al.*, The topoisomerase IIbeta circular clamp arrests transcription and signals a 26S proteasome pathway. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 3239-3244 (2003).
165. I. Alchanati *et al.*, The E3 ubiquitin-ligase Bmi1/Ring1A controls the proteasomal degradation of Top2alpha cleavage complex - a potentially new drug target. *PloS one* **4**, e8104 (2009).
166. A. Canela *et al.*, Topoisomerase II-Induced Chromosome Breakage and Translocation Is Determined by Chromosome Architecture and Transcriptional Activity. *Molecular cell* **75**, 252-266.e258 (2019).
167. N. Sciascia *et al.*, Suppressing proteasome mediated processing of topoisomerase II DNA-protein complexes preserves genome integrity. *eLife* **9** (2020).

168. K. C. Lee, R. L. Bramley, I. G. Cowell, G. H. Jackson, C. A. Austin, Proteasomal inhibition potentiates drugs targeting DNA topoisomerase II. *Biochem Pharmacol* **103**, 29-39 (2016).
169. Y. Mao, S. D. Desai, L. F. Liu, SUMO-1 conjugation to human DNA topoisomerase II isozymes. *J Biol Chem* **275**, 26066-26073 (2000).
170. J. Stingele, M. S. Schwarz, N. Bloemeke, P. G. Wolf, S. Jentsch, A DNA-dependent protease involved in DNA-protein crosslink repair. *Cell* **158**, 327-338 (2014).
171. M. Y. Balakirev *et al.*, Wss1 metalloprotease partners with Cdc48/Doa1 in processing genotoxic SUMO conjugates. *eLife* **4** (2015).
172. J. R. Mullen, C. F. Chen, S. J. Brill, Wss1 is a SUMO-dependent isopeptidase that interacts genetically with the Slx5-Slx8 SUMO-targeted ubiquitin ligase. *Mol Cell Biol* **30**, 3737-3748 (2010).
173. J. Stingele, B. Habermann, S. Jentsch, DNA-protein crosslink repair: proteases as DNA repair enzymes. *Trends Biochem Sci* **40**, 67-71 (2015).
174. J. P. Duxin, J. M. Dewar, H. Yardimci, J. C. Walter, Repair of a DNA-protein crosslink by replication-coupled proteolysis. *Cell* **159**, 346-357 (2014).
175. R. S. Maskey *et al.*, Spartan deficiency causes accumulation of Topoisomerase 1 cleavage complexes and tumorigenesis. *Nucleic acids research* **45**, 4564-4576 (2017).
176. M. Mórocz *et al.*, DNA-dependent protease activity of human Spartan facilitates replication of DNA-protein crosslink-containing DNA. *Nucleic acids research* **45**, 3172-3188 (2017).
177. J. Stingele *et al.*, Mechanism and Regulation of DNA-Protein Crosslink Repair by the DNA-Dependent Metalloprotease SPRTN. *Molecular cell* **64**, 688-703 (2016).

178. N. B. Larsen *et al.*, Replication-Coupled DNA-Protein Crosslink Repair by SPRTN and the Proteasome in *Xenopus* Egg Extracts. *Molecular cell* **73**, 574-588.e577 (2019).
179. D. Lessel *et al.*, Mutations in SPRTN cause early onset hepatocellular carcinoma, genomic instability and progeroid features. *Nature genetics* **46**, 1239-1244 (2014).
180. M. Svoboda, J. Konvalinka, J. F. Trempe, K. Grantz Saskova, The yeast proteases Ddi1 and Wss1 are both involved in the DNA replication stress response. *DNA repair* **80**, 45-51 (2019).
181. M. C. J. Yip, N. O. Bodnar, T. A. Rapoport, Ddi1 is a ubiquitin-dependent protease. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 7776-7781 (2020).
182. S. W. Yang *et al.*, A eukaryotic enzyme that can disjoin dead-end covalent complexes between DNA and type I topoisomerases. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 11534-11539 (1996).
183. J. J. Pouliot, K. C. Yao, C. A. Robertson, H. A. Nash, Yeast gene for a Tyr-DNA phosphodiesterase that repairs topoisomerase I complexes. *Science (New York, N.Y.)* **286**, 552-555 (1999).
184. J. J. Pouliot, C. A. Robertson, H. A. Nash, Pathways for repair of topoisomerase I covalent complexes in *Saccharomyces cerevisiae*. *Genes Cells* **6**, 677-687 (2001).
185. F. Cortes Ledesma, S. F. El Khamisy, M. C. Zuma, K. Osborn, K. W. Caldecott, A human 5'-tyrosyl DNA phosphodiesterase that repairs topoisomerase-mediated DNA damage. *Nature* **461**, 674-678 (2009).
186. Z. Zeng, F. Cortés-Ledesma, S. F. El Khamisy, K. W. Caldecott, TDP2/TTRAP is the major 5'-tyrosyl DNA phosphodiesterase activity in vertebrate cells and is critical for cellular resistance to topoisomerase II-induced DNA damage. *J Biol Chem* **286**, 403-409 (2011).

187. F. Gómez-Herreros *et al.*, TDP2-dependent non-homologous end-joining protects against topoisomerase II-induced DNA breaks and genome instability in cells and in vivo. *PLoS genetics* **9**, e1003226 (2013).
188. F. Gómez-Herreros *et al.*, TDP2 protects transcription from abortive topoisomerase activity and is required for normal neural function. *Nature genetics* **46**, 516-521 (2014).
189. Y. Pommier *et al.*, Tyrosyl-DNA-phosphodiesterases (TDP1 and TDP2). *DNA repair* **19**, 114-129 (2014).
190. K. C. Nitiss, M. Malik, X. He, S. W. White, J. L. Nitiss, Tyrosyl-DNA phosphodiesterase (Tdp1) participates in the repair of Top2-mediated DNA damage. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 8953-8958 (2006).
191. M. A. Borda, M. Palmitelli, G. Veron, M. Gonzalez-Cid, M. de Campos Nebel, Tyrosyl-DNA-phosphodiesterase I (TDP1) participates in the removal and repair of stabilized-Top2alpha cleavage complexes in human cells. *Mutation research* **781**, 37-48 (2015).
192. H. Interthal, J. J. Pouliot, J. J. Champoux, The tyrosyl-DNA phosphodiesterase Tdp1 is a member of the phospholipase D superfamily. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 12009-12014 (2001).
193. D. R. Davies, H. Interthal, J. J. Champoux, W. G. Hol, The crystal structure of human tyrosyl-DNA phosphodiesterase, Tdp1. *Structure* **10**, 237-248 (2002).
194. R. Hirano *et al.*, Spinocerebellar ataxia with axonal neuropathy: consequence of a Tdp1 recessive neomorphic mutation? *The EMBO journal* **26**, 4732-4743 (2007).
195. R. Gao, S. Y. Huang, C. Marchand, Y. Pommier, Biochemical characterization of human tyrosyl-DNA phosphodiesterase 2 (TDP2/TTRAP): a Mg(2+)/Mn(2+)-dependent phosphodiesterase specific

- for the repair of topoisomerase cleavage complexes. *J Biol Chem* **287**, 30842-30852 (2012).
196. S. Adhikari *et al.*, Characterization of magnesium requirement of human 5'-tyrosyl DNA phosphodiesterase mediated reaction. *BMC Res Notes* **5**, 134 (2012).
  197. R. Gao, Schellenberg, M.J., Huang, S.N., Abdelmalak, M., Marchand, C., Nitiss, K.C., Nitiss, J.L., Williams, R.S., Pommier, Y., Proteolytic Degradation of Topoisomerase II (Top2) Enables the Processing of Top2-DNA and Top2-RNA Covalent Complexes by Tyrosyl-DNA-Phosphodiesterase 2 (TDP2) *Journal of Biological Chemistry* **289**, 17960-17969 (2014).
  198. M. J. Schellenberg *et al.*, ZATT (ZNF451)-mediated resolution of topoisomerase 2 DNA-protein cross-links. *Science (New York, N.Y.)* **357**, 1412-1416 (2017).
  199. E. P. Mimitou, L. S. Symington, Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. *Nature* **455**, 770-774 (2008).
  200. Z. Zhu, W. H. Chung, E. Y. Shim, S. E. Lee, G. Ira, Sgs1 helicase and two nucleases Dna2 and Exo1 resect DNA double-strand break ends. *Cell* **134**, 981-994 (2008).
  201. R. A. Deshpande, J. H. Lee, S. Arora, T. T. Paull, Nbs1 Converts the Human Mre11/Rad50 Nuclease Complex into an Endo/Exonuclease Machine Specific for Protein-DNA Adducts. *Molecular cell* **64**, 593-606 (2016).
  202. E. Cannavo, P. Cejka, Sae2 promotes dsDNA endonuclease activity within Mre11-Rad50-Xrs2 to resect DNA breaks. *Nature* **514**, 122-125 (2014).
  203. M. Malik, J. L. Nitiss, DNA repair functions that control sensitivity to topoisomerase-targeting drugs. *Eukaryot Cell* **3**, 82-90 (2004).

204. E. Hartsuiker, M. J. Neale, A. M. Carr, Distinct requirements for the Rad32(Mre11) nuclease and Ctp1(CtIP) in the removal of covalently bound topoisomerase I and II from DNA. *Molecular cell* **33**, 117-123 (2009).
205. T. Aparicio, R. Baer, M. Gottesman, J. Gautier, MRN, CtIP, and BRCA1 mediate repair of topoisomerase II-DNA adducts. *The Journal of cell biology* **212**, 399-408 (2016).
206. K. Nakamura *et al.*, Collaborative action of Brca1 and CtIP in elimination of covalent modifications from double-strand breaks to facilitate subsequent break repair. *PLoS genetics* **6**, e1000828 (2010).
207. K. C. Lee *et al.*, MRE11 facilitates the removal of human topoisomerase II complexes from genomic DNA. *Biol Open* **1**, 863-873 (2012).
208. N. N. Hoa *et al.*, Mre11 Is Essential for the Removal of Lethal Topoisomerase 2 Covalent Cleavage Complexes. *Molecular cell* **64**, 1010 (2016).
209. H. Sasanuma *et al.*, BRCA1 ensures genome integrity by eliminating estrogen-induced pathological topoisomerase II-DNA complexes. *Proceedings of the National Academy of Sciences of the United States of America* **115**, E10642-e10651 (2018).
210. M. Sabourin *et al.*, Yeast recombination pathways triggered by topoisomerase II-mediated DNA breaks. *Nucleic acids research* **31**, 4373-4384 (2003).
211. N. Adachi, S. Iizumi, S. So, H. Koyama, Genetic evidence for involvement of two distinct nonhomologous end-joining pathways in repair of topoisomerase II-mediated DNA damage. *Biochem Biophys Res Commun* **318**, 856-861 (2004).
212. A. Kurosawa *et al.*, DNA ligase IV and artemis act cooperatively to suppress homologous recombination in human cells: implications for DNA double-strand break repair. *PloS one* **8**, e72253 (2013).

213. Y. Pommier, G. Capranico, A. Orr, K. W. Kohn, Distribution of topoisomerase II cleavage sites in simian virus 40 DNA and the effects of drugs. *Journal of molecular biology* **222**, 909-924 (1991).
214. M. Sander, T. S. Hsieh, Drosophila topoisomerase II double-strand DNA cleavage: analysis of DNA sequence homology at the cleavage site. *Nucleic acids research* **13**, 1057-1072 (1985).
215. J. R. Spitzner, M. T. Muller, A consensus sequence for cleavage by vertebrate DNA topoisomerase II. *Nucleic acids research* **16**, 5533-5556 (1988).
216. J. R. Spitzner, I. K. Chung, M. T. Muller, Eukaryotic topoisomerase II preferentially cleaves alternating purine-pyrimidine repeats. *Nucleic acids research* **18**, 1-11 (1990).
217. M. Cornarotti *et al.*, Drug sensitivity and sequence specificity of human recombinant DNA topoisomerases IIalpha (p170) and IIbeta (p180). *Molecular pharmacology* **50**, 1463-1471 (1996).
218. K. L. Marsh *et al.*, Amsacrine-promoted DNA cleavage site determinants for the two human DNA topoisomerase II isoforms alpha and beta. *Biochem Pharmacol* **52**, 1675-1685 (1996).
219. D. Strumberg, J. L. Nitiss, J. Dong, K. W. Kohn, Y. Pommier, Molecular analysis of yeast and human type II topoisomerases. Enzyme-DNA and drug interactions. *J Biol Chem* **274**, 28246-28255 (1999).
220. W. H. Gittens *et al.*, A nucleotide resolution map of Top2-linked DNA breaks in the yeast and human genome. *Nature communications* **10**, 4846 (2019).
221. S. Lee *et al.*, DNA cleavage and opening reactions of human topoisomerase II $\alpha$  are regulated via Mg<sup>2+</sup>-mediated dynamic bending of gate-DNA. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 2925-2930 (2012).

222. Y. Jang *et al.*, Selection of DNA Cleavage Sites by Topoisomerase II Results from Enzyme-Induced Flexibility of DNA. *Cell Chem Biol* **26**, 502-511.e503 (2019).
223. A. K. McClendon, N. Osheroff, DNA topoisomerase II, genotoxicity, and cancer. *Mutation research* **623**, 83-97 (2007).
224. F. Gómez-Herreros *et al.*, TDP2 suppresses chromosomal translocations induced by DNA topoisomerase II during gene transcription. *Nature communications* **8**, 233 (2017).
225. J. Olmedo-Pelayo, D. Rubio-Contreras, F. Gómez-Herreros, Canonical non-homologous end-joining promotes genome mutagenesis and translocations induced by transcription-associated DNA topoisomerase 2 activity. *Nucleic acids research* **48**, 9147-9160 (2020).
226. K. Szlachta *et al.*, Topoisomerase II contributes to DNA secondary structure-mediated double-stranded breaks. *Nucleic acids research* **48**, 6654-6671 (2020).
227. F. Gómez-Herreros, DNA Double Strand Breaks and Chromosomal Translocations Induced by DNA Topoisomerase II. *Front Mol Biosci* **6**, 141 (2019).
228. N. Stantial *et al.*, Trapped topoisomerase II initiates formation of de novo duplications via the nonhomologous end-joining pathway in yeast. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 26876-26884 (2020).
229. J. B. Leppard, J. J. Champoux, Human DNA topoisomerase I: relaxation, roles, and damage control. *Chromosoma* **114**, 75-85 (2005).
230. J. C. Wang, Moving one DNA double helix through another by a type II DNA topoisomerase: the story of a simple molecular machine. *Quarterly reviews of biophysics* **31**, 107-144 (1998).

231. R. Ceccaldi, B. Rondinelli, A. D. D'Andrea, Repair Pathway Choices and Consequences at the Double-Strand Break. *Trends Cell Biol* **26**, 52-64 (2016).
232. J. L. Nitiss, Targeting DNA topoisomerase II in cancer chemotherapy. *Nat. Rev. Cancer* **9**, 338-350 (2009).
233. N. G. Bush, K. Evans-Roberts, A. Maxwell, DNA topoisomerases. *EcoSal Plus* **6** (2015).
234. K. Evans-Roberts, A. Maxwell, DNA Topoisomerases. *EcoSal Plus* **3** (2009).
235. J. L. Nitiss, K. C. Nitiss, A. Rose, J. L. Waltman, Overexpression of type I topoisomerases sensitizes yeast cells to DNA damage. *J. Biol. Chem.* **276**, 26708-26714 (2001).
236. P. Pourquier *et al.*, Effects of uracil incorporation, DNA mismatches, and abasic sites on cleavage and religation activities of mammalian topoisomerase I. *J Biol Chem* **272**, 7792-7796 (1997).
237. P. Pourquier *et al.*, Induction of reversible complexes between eukaryotic DNA topoisomerase I and DNA-containing oxidative base damages. 7, 8-dihydro-8-oxoguanine and 5-hydroxycytosine. *J Biol Chem* **274**, 8516-8523 (1999).
238. J. Sekiguchi, S. Shuman, Site-specific ribonuclease activity of eukaryotic DNA topoisomerase I. *Molecular cell* **1**, 89-97 (1997).
239. N. Kim *et al.*, Mutagenic processing of ribonucleotides in DNA by yeast topoisomerase I. *Science (New York, N.Y.)* **332**, 1561-1564 (2011).
240. J. L. Sparks, P. M. Burgers, Error-free and mutagenic processing of topoisomerase 1-provoked damage at genomic ribonucleotides. *The EMBO journal* **34**, 1259-1269 (2015).

241. S. Y. Huang, S. Ghosh, Y. Pommier, Topoisomerase I alone is sufficient to produce short DNA deletions and can also reverse nicks at ribonucleotide sites. *J Biol Chem* **290**, 14068-14076 (2015).
242. J. E. Cho *et al.*, Parallel analysis of ribonucleotide-dependent deletions produced by yeast Top1 in vitro and in vivo. *Nucleic acids research* **44**, 7714-7721 (2016).
243. K. C. Nitiss, J. L. Nitiss, L. A. Hanakahi, DNA Damage by an essential enzyme: A delicate balance act on the tightrope. *DNA repair* **82**, 102639 (2019).
244. M. D. Megonigal, J. Fertala, M. A. Bjornsti, Alterations in the catalytic activity of yeast DNA topoisomerase I result in cell cycle arrest and cell death. *J. Biol. Chem.* **272**, 12801-12808 (1997).
245. N. A. Levin, M. A. Bjornsti, G. R. Fink, A novel mutation in DNA topoisomerase I of yeast causes DNA damage and RAD9-dependent cell cycle arrest. *Genetics* **133**, 799-814 (1993).
246. J. E. Cho, S. Jinks-Robertson, Topoisomerase I and genome stability: The good and the bad. *Methods in molecular biology (Clifton, N.J.)* **1703**, 21-45 (2018).
247. J. V. Walker *et al.*, A mutation in human topoisomerase II alpha whose expression is lethal in DNA repair-deficient yeast cells. *J. Biol. Chem.* **279**, 25947-25954 (2004).
248. K. Lehner, S. V. Mudrak, B. K. Minesinger, S. Jinks-Robertson, Frameshift mutagenesis: the roles of primer-template misalignment and the nonhomologous end-joining pathway in *Saccharomyces cerevisiae*. *Genetics* **190**, 501-510 (2012).
249. B. J. Thomas, R. Rothstein, Elevated recombination rates in transcriptionally active DNA. *Cell* **56**, 619-630 (1989).

250. A. L. Goldstein, J. H. McCusker, Three new dominant drug resistance cassettes for gene disruption in *Saccharomyces cerevisiae*. *Yeast* **15**, 1541-1553 (1999).
251. X. Guo, Y. F. Hum, K. Lehner, S. Jinks-Robertson, Regulation of hetDNA length during mitotic double-strand break repair in yeast. *Mol. Cell* **67**, 539-549 (2017).
252. U. Gueldener, J. Heinisch, G. J. Koehler, D. Voss, J. H. Hegemann, A second set of loxP marker cassettes for Cre-mediated multiple gene knockouts in budding yeast. *Nucleic acids research* **30**, e23 (2002).
253. B. Llorente, L. S. Symington, The Mre11 nuclease is not required for 5' to 3' resection at multiple HO-induced double-strand breaks. *Mol. Cell. Biol.* **24**, 9682-9694 (2004).
254. A. Stepanov, K. C. Nitiss, G. Neale, J. L. Nitiss, Enhancing drug accumulation in *Saccharomyces cerevisiae* by repression of pleiotropic drug resistance genes with chimeric transcription repressors. *Molecular pharmacology* **74**, 423-431 (2008).
255. J. Cho, S. Jinks-Robertson, Deletions associated with stabilization of the Top1 cleavage complex in yeast are products of the nonhomologous end-joining pathway. *Proc. Natl. Acad. Sci. USA* **116**, 22683-22691 (2019).
256. U. Guldener, J. Heinisch, G. J. Koehler, D. Voss, J. H. Hegemann, A second set of loxP marker cassettes for Cre-mediated multiple gene knockouts in budding yeast. *Nucleic Acids Res.* **30**, e23 (2002).
257. M. Jannatipour, Y. X. Liu, J. L. Nitiss, The top2-5 mutant of yeast topoisomerase II encodes an enzyme resistant to etoposide and amsacrine. *J. Biol. Chem.* **268**, 18586-18592 (1993).
258. Y. Mao, M. Sun, S. D. Desai, L. F. Liu, SUMO-1 conjugation to topoisomerase I: a possible repair response to topoisomerase-mediated DNA damage. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 4046-4051 (2000).

259. S. Worland, J. Wang, Inducible overexpression, purification, and active-site mapping of DNA topoisomerase-II from the yeast *Saccharomyces cerevisiae*. *J. Biol. Chem.* **264**, 4412-4416 (1989).
260. A. T. Rogojina, J. L. Nitiss, Isolation and characterization of mAMSA-hypersensitive mutants. Cytotoxicity of Top2 covalent complexes containing DNA single strand breaks. *J. Biol. Chem.* **283**, 29239-29250 (2008).
261. J. L. Nitiss, E. Soans, A. Rogojina, A. Seth, M. Mishina, Topoisomerase assays. *Curr. Protoc. Pharmacol.* **57**, 3.3.1-3.3.27 (2012).
262. M. D. Rose, P. Novick, J. H. Thomas, D. Botstein, G. R. Fink, A *Saccharomyces cerevisiae* plasmid bank based on a centromere-containing shuttle vector. *Gene* **60**, 237-243 (1987).
263. J. L. Nitiss *et al.*, Amsacrine and etoposide hypersensitivity of yeast cells overexpressing DNA topoisomerase II. *Cancer research* **52**, 4467-4472 (1992).
264. D. E. Lea, C. A. Coulson, The distribution of the numbers of mutants in bacterial populations. *J. Genet.* **49**, 264-285 (1949).
265. R. M. Spell, S. Jinks-Robertson, "Determination of Mitotic Recombination Rates by Fluctuation Analysis in *Saccharomyces cerevisiae*" in Genetic Recombination: Reviews and Protocols, A. S. Waldman, Ed. (Humana Press, Totowa, NJ, 2004), 10.1385/1-59259-761-0:003, pp. 3-12.
266. X. Guo *et al.*, SMRT Sequencing for Parallel Analysis of Multiple Targets and Accurate SNP Phasing. *G3 (Bethesda, Md.)* **5**, 2801-2808 (2015).
267. K. R. Nair, Table of confidence interval for the median in samples from any continuous population. *Sankhya: The Indian Journal of Statistics* **4** (1940).
268. A. Moore *et al.*, Genetic control of genomic alterations induced in yeast by interstitial telomeric sequences. *Genetics* **209**, 425-438 (2018).

269. Y. Hsiung, S. H. Elsea, N. Osheroff, J. L. Nitiss, A mutation in yeast TOP2 homologous to a quinolone-resistant mutation in bacteria. Mutation of the amino acid homologous to Ser83 of *Escherichia coli gyrA* alters sensitivity to eukaryotic topoisomerase inhibitors. *J. Biol. Chem.* **270**, 20359-20364 (1995).
270. K. J. Dornfield, D. M. Livingston, Effects of controlled *RAD52* expression on repair and recombination in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **11**, 2013-2017 (1991).
271. J. Anand, Y. Sun, Y. Zhao, K. C. Nitiss, J. L. Nitiss, Detection of topoisomerase covalent complexes in eukaryotic cells. *Methods in molecular biology (Clifton, N.J.)* **1703**, 283-299 (2018).
272. D. Subramanian, C. S. Furbee, M. T. Muller, ICE bioassay. Isolating in vivo complexes of enzyme to DNA. *Methods in molecular biology (Clifton, N.J.)* **95**, 137-147 (2001).
273. G. N. Giaever, L. Snyder, J. C. Wang, DNA supercoiling *in vivo*. *Biophysical chemistry* **29**, 7-15 (1988).
274. S. M. Vos, E. M. Tretter, B. H. Schmidt, J. M. Berger, All tangled up: how cells direct, manage and exploit topoisomerase function. *Nat. Rev. Mol. Cell Biol.* **12**, 827-841 (2011).
275. M. Sander, T. Hsieh, Double strand DNA cleavage by type II DNA topoisomerase from *Drosophila melanogaster*. *J. Biol. Chem.* **258**, 8421-8428 (1983).
276. K. Lehner, S. V. Mudrak, B. K. Minesinger, S. Jinks-Robertson, Frameshift mutagenesis: the roles of primer-template misalignment and the nonhomologous end-joining pathway in *Saccharomyces cerevisiae*. *Genetics* **190**, 501-510 (2012).
277. J. M. Daley, P. L. Palmboos, D. Wu, T. E. Wilson, Nonhomologous end joining in yeast. *Annu. Rev. Genet.* **39**, 431-451 (2005).

278. E. L. Baldwin, A. C. Berger, A. H. Corbett, N. Osheroff, Mms22p protects *Saccharomyces cerevisiae* from DNA damage induced by topoisomerase II. *Nucleic Acids Res.* **33**, 1021-1030 (2005).
279. S. Moreau, J. R. Ferguson, L. S. Symington, The nuclease activity of Mre11 is required for meiosis but not for mating type switching, end joining, or telomere maintenance. *Mol Cell Biol* **19**, 556-566 (1999).
280. Y. Habraken, P. Sung, L. Prakash, S. Prakash, Structure-specific nuclease activity in yeast nucleotide excision repair protein Rad2. *J Biol Chem* **270**, 30194-30198 (1995).
281. J. Roca, J. C. Wang, The capture of a DNA double helix by an ATP-dependent protein clamp: a key step in DNA transport by type II DNA topoisomerases. *Cell* **71**, 833-840 (1992).
282. J. Roca, J. C. Wang, DNA transport by a type II DNA topoisomerase: evidence in favor of a two-gate mechanism. *Cell* **77**, 609-616 (1994).
283. J. M. Berger, S. J. Gamblin, S. C. Harrison, J. C. Wang, Structure and mechanism of DNA topoisomerase II. *Nature* **379**, 225-232 (1996).
284. J. Roca, J. M. Berger, S. C. Harrison, J. C. Wang, DNA transport by a type II topoisomerase: direct evidence for a two-gate mechanism. *Proc. Natl. Acad. Sci. USA* **93**, 4057-4062 (1996).
285. B. H. Schmidt, N. Osheroff, J. M. Berger, Structure of a topoisomerase II-DNA-nucleotide complex reveals a new control mechanism for ATPase activity. *Nat. Struct. Mol. Biol.* **19**, 1147-1154 (2012).
286. T. T. Paull, 20 Years of Mre11 biology: No end in sight. *Mol. Cell* **71**, 419-427 (2018).
287. E. Cannavo, G. Reginato, P. Cejka, Stepwise 5' DNA end-specific resection of DNA breaks by the Mre11-Rad50-Xrs2 and Sae2 nuclease ensemble. *Proc. Natl. Acad. Sci. USA* **116**, 5505-5513 (2019).

288. L. S. Symington, J. Gautier, Double-strand break end resection and repair pathway choice. *Annu. Rev. Genet.* **45**, 247-271 (2011).
289. A. Sfeir, L. S. Symington, Microhomology-Mediated End Joining: A Backup Survival Mechanism or Dedicated Pathway? *Trends Biochem Sci* **40**, 701-714 (2015).
290. S. N. Guzder, Y. Habraken, P. Sung, L. Prakash, S. Prakash, Reconstitution of yeast nucleotide excision repair with purified Rad proteins, replication protein A, and transcription factor TFIIH. *J Biol Chem* **270**, 12973-12976 (1995).
291. C. J. Barnes, A. F. Wahl, B. Shen, M. S. Park, R. A. Bambara, Mechanism of tracking and cleavage of adduct-damaged DNA substrates by the mammalian 5'- to 3'-exonuclease/endonuclease RAD2 homologue 1 or flap endonuclease 1. *J Biol Chem* **271**, 29624-29631 (1996).
292. L. B. Alexandrov *et al.*, The repertoire of mutational signatures in human cancer. *Nature* **578**, 94-101 (2020).
293. A. Boot, S. G. Rozen, Distinctive indel mutational signature in tumors carrying TOP2A p.K743N. *bioRxiv* 10.1101/2020.05.22.111666, 2020.2005.2022.111666 (2020).
294. M. Petljak, L. B. Alexandrov, Understanding mutagenesis through delineation of mutational signatures in human cancer. *Carcinogenesis* **37**, 531-540 (2016).
295. J. Kim *et al.*, Somatic ERCC2 mutations are associated with a distinct genomic signature in urothelial tumors. *Nature genetics* **48**, 600-606 (2016).
296. S. Nik-Zainal *et al.*, Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* **534**, 47-54 (2016).

297. S. Mimaki *et al.*, Hypermutation and unique mutational signatures of occupational cholangiocarcinoma in printing workers exposed to haloalkanes. *Carcinogenesis* **37**, 817-826 (2016).
298. N. K. Hayward *et al.*, Whole-genome landscapes of major melanoma subtypes. *Nature* **545**, 175-180 (2017).
299. P. Polak *et al.*, A mutational signature reveals alterations underlying deficient homologous recombination repair in breast cancer. *Nature genetics* **49**, 1476-1486 (2017).
300. C. Pleguezuelos-Manzano *et al.*, Mutational signature in colorectal cancer caused by genotoxic pks(+) *E. coli*. *Nature* **580**, 269-273 (2020).
301. R. Poplin *et al.*, A universal SNP and small-indel variant caller using deep neural networks. *Nat Biotechnol* **36**, 983-987 (2018).
302. G. I. Lang, A. W. Murray, Estimating the per-base-pair mutation rate in the yeast *Saccharomyces cerevisiae*. *Genetics* **178**, 67-82 (2008).
303. M. L. García-Rubio, A. Aguilera, Topological constraints impair RNA polymerase II transcription and causes instability of plasmid-borne convergent genes. *Nucleic acids research* **40**, 1050-1064 (2012).
304. K. C. Dong, J. M. Berger, Structural basis for gate-DNA recognition and bending by type IIA topoisomerases. *Nature* **450**, 1201-1205 (2007).
305. M. Antoniou-Kourounioti, M. L. Mimmack, A. C. G. Porter, C. J. Farr, The Impact of the C-Terminal Region on the Interaction of Topoisomerase II Alpha with Mitotic Chromatin. *International journal of molecular sciences* **20** (2019).
306. M. J. Lippert *et al.*, Role for topoisomerase 1 in transcription-associated mutagenesis in yeast. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 698-703 (2011).

307. T. Takahashi, G. Burguiere-Slezak, P. A. Van der Kemp, S. Boiteux, Topoisomerase 1 provokes the formation of short deletions in repeated sequences upon high transcription in *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 692-697 (2011).
308. C. Concannon, R. S. Lahue, The 26S proteasome drives trinucleotide repeat expansions. *Nucleic acids research* **41**, 6098-6108 (2013).
309. Ž. Skok, N. Zidar, D. Kikelj, J. Ilaš, Dual Inhibitors of Human DNA Topoisomerase II and Other Cancer-Related Targets. *J Med Chem* **63**, 884-904 (2020).
310. H. L. Klein *et al.*, Guidelines for DNA recombination and repair studies: Cellular assays of DNA repair pathways. *Microb Cell* **6**, 1-64 (2019).

## Biography

Nicole Stantial attended Assumption University, where she received a B.A. in Biotechnology and Molecular Biology and minored in Chemistry. At Assumption, Nicole worked in the lab of Dr. David Crowley, and her work resulted in a first author publication, "Transcription-coupled Repair of UV Damage in the Halophilic Archaea". After graduating in 2015, Nicole started graduate school of Duke University in the Program in Cell and Molecular Biology (CMB). Nicole joined the lab of Dr. Sue Jinks-Robertson and the Department of Molecular Genetics and Microbiology (MGM). Nicole was co-first author of the publication "Trapped topoisomerase II initiates formation of de novo duplications via the nonhomologous end-joining pathway in yeast". Nicole received the Best Speaker Award from the MGM departmental annual retreat in 2018 and 2020, a Conference Travel Award from Duke University The Graduate School in 2018, and the Chairman's Travel Award from MGM at Duke University in 2018. Nicole was elected to co-chair the 2020 Gordan Research Seminar on Mutagenesis (cancelled due to COVID-19). She was chosen to participate in the Duke-NUS Pre-doctoral Research Exchange Program in 2020 but was unable to go to Singapore because of COVID-19. Finally, Nicole received the Jo Rae Wright Award from the Graduate and Professional Student Council at Duke University in 2020 for her work on combating food insecurity on campus.