

3-26-2020

## **DNA Base Excision Repair Modulates DNA Repeat Instability and Non-B Form DNA Structures**

Eduardo E. Laverde

*Florida International University*, [elave006@fiu.edu](mailto:elave006@fiu.edu)

Follow this and additional works at: <https://digitalcommons.fiu.edu/etd>



Part of the [Biochemistry Commons](#), [Laboratory and Basic Science Research Commons](#), and the [Molecular Biology Commons](#)

---

### **Recommended Citation**

Laverde, Eduardo E., "DNA Base Excision Repair Modulates DNA Repeat Instability and Non-B Form DNA Structures" (2020). *FIU Electronic Theses and Dissertations*. 4391.

<https://digitalcommons.fiu.edu/etd/4391>

This work is brought to you for free and open access by the University Graduate School at FIU Digital Commons. It has been accepted for inclusion in FIU Electronic Theses and Dissertations by an authorized administrator of FIU Digital Commons. For more information, please contact [dcc@fiu.edu](mailto:dcc@fiu.edu).

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

DNA BASE EXCISION REPAIR MODULATES DNA REPEAT INSTABILITY AND  
NON-B FORM DNA STRUCTURES

A dissertation submitted in partial fulfillment of

the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

BIOCHEMISTRY

by

Eduardo E. Laverde

2020

To: Dean Michael R. Heithaus  
College of Arts, Sciences and Education

This dissertation, written by Eduardo E. Laverde, and entitled DNA Base Excision Repair Modulates DNA Repeat Instability and non-B Form DNA Structures, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

---

Alejandro Barbieri

---

Yuk-Ching Tse-Dinh

---

Xiaotang Wang

---

Yuan Liu, Major Professor

Date of Defense: March 26, 2020

The dissertation of Eduardo E. Laverde is approved.

---

Dean Michael R. Heithaus  
College of Arts, Sciences and Education

---

Andrés G. Gil  
Vice President for Research and Economic Development  
and Dean of the University Graduate School

Florida International University, 2020

© Copyright 2020 by Eduardo E. Laverde

All rights reserved.

## DEDICATION

I dedicate this dissertation to my family for their love, encouragement, and support.

America M. Niño de Guzman

Eduardo R. Laverde

Nelly R. Laverde

Leticia J. Lopez.

Alexander X. Laverde

Alyson X. Laverde

Ruth S. Laverde

Isabel M. Laverde

## ACKNOWLEDGMENTS

I want to express my most sincere gratitude to all the people that directly or indirectly help me to achieve this work. First, I would like to thank my advisor/mentor Dr. Yuan Liu for her guidance, mentorship, help, and support through my undergraduate and graduate studies. She motivated me to pursue this magnificent and fulfilling career path in biochemistry and molecular biology. Under her tutelage, I developed many skills such as critical thinking, writing/oral communication, and experimental procedures that will serve me well in every step of my career.

I want to thank my committee members Dr. Alejandro Barbieri, Dr. Yuk-Ching Tse-Dihn, and Dr. Xiaotang Wang, for their guidance, comments, support, and collaboration through my Ph.D. training.

I want to thank my present and past lab members who helped me directly and indirectly through this period of my life. I would especially like to thank Dr. Meng Xu for her patience and guidance in training me when I first joined the lab. Also, I would like to thank Dr. Yanhao Lai for her training and guidance in most of my experimental procedures and methods. Her invaluable feedback helped me shape my ideas to complete this dissertation. I want to thank Dr. Yaou Ren, Dr. Zhongliang Jiang, Dr. Jill Beaver, Pawlos Tsegay, as well as other members of Dr. Liu's research lab and neighboring labs for their friendship and moral support during my Ph.D. studies.

I want to thank our collaborators. Dr. Samuel H. Wilson, Genome Integrity and Structural Biology Laboratory, National Institute of Environmental Health

Sciences (NIEHS), National Institutes of Health (NIH) for providing us with expression vectors and purified enzymes necessary for the completion of these projects. Dr. Joseph Jiricny for providing us with the FAN1 expression vector required for the fulfillment of the first aim. Dr. Catherine Freudenreich, Tufts University, critically analyze and provide feedback for the achievement of the second aim. Dr. Kenneth Fischbeck, National Institute of Neurological Disorders and Stroke (NINDS), for providing us with the AOA2 and fibroblast cell lines for the completion of the third aim. Dr. Fenfei Leng, Florida International University (FIU), for providing us with RNaseH1 necessary for control experiments of the second and third aims.

Finally, I would like to thank the university graduate (UGS) school and the National Institute of General Medical Science (NIGMS RISE program) for providing me with the “McNair Graduate Fellowship,” “Dissertation Year Fellowship,” and “2018 Summer Research Award”. These fellowships and grants helped me complete this dissertation on time.

ABSTRACT OF THE DISSERTATION  
DNA BASE EXCISION REPAIR MODULATES DNA REPEAT INSTABILITY AND  
NON-B FORM DNA STRUCTURES

by

Eduardo E. Laverde

Florida International University, 2020

Miami, Florida

Professor Yuan Liu, Major Professor

The human genome is constantly attacked by endogenous and exogenous sources of DNA damage that generates DNA base lesions and strand breaks leading to genome instability, cell death, and diseases. To combat these adverse effects, cells have evolved a robust DNA repair mechanism called “the DNA base excision repair (BER),” which efficiently removes DNA lesions maintaining genome stability. However, its underlying molecular mechanisms remain to be elucidated. In my dissertation research, I explored the molecular mechanism by which BER modulates trinucleotide repeats (TNR) by processing non-B form structures such as hairpins and R-loops through the coordination among BER enzymes and cofactors and the proteins from other DNA repair pathways. For the first time, we discovered that Fanconi anemia-associated nuclease 1 (FAN1) coordinated with a key BER enzyme, flap endonuclease 1 (FEN1), to prevent repeat expansion by promoting CAG repeat deletion during BER of oxidative DNA damage. We further demonstrated that the BER cofactor, PCNA/ubPCNA, coordinated with FAN1 to



attenuate CAG repeat expansion during BER in a CAG repeat hairpin. Exploring the mechanisms by which BER regulates TNR instability by processing bulky non-B form structures, R-loops. We found that an abasic lesion on the non-template strand of a CAG repeat R-loop was incised by AP endonuclease 1 (APE1), converting the R-loop into a double-flap intermediate containing an RNA:DNA hybrid. The R-loop intermediate inhibited DNA polymerase  $\beta$  (pol  $\beta$ ) DNA synthesis and stimulated FEN1 cleavage of the repeats leading to repeat deletion. We showed that FEN1 partially cleaved the RNA strand on a CAG repeat R-loop to facilitate pol  $\beta$  skip-over of a hairpin leading to repeat deletion. We further identified a new role of FEN1 of resolving R-loops through the BER pathway. We found that FEN1 efficiently cleaved RNA during the processing of DNA lagging strand and R-loops and demonstrated that FEN1 also used its flap tracking to track down to the DNA region to make a cleavage removing the RNA leading to the lagging strand maturation and R-loop resolution via BER. Our study provides new insights into the molecular mechanism by which BER maintains genome stability and integrity.

## TABLE OF CONTENTS

CHAPTER	PAGE
INTRODUCTION .....	1
A. DNA damage and repair .....	1
B. The DNA Base Excision Repair Pathway.....	8
C. BER of DNA base damage as a regulator of trinucleotide repeat instability..	14
D. R-loop biology and DNA Repair .....	18
OVERVIEW .....	25
CHAPTER 1: FANCONI ANEMIA-ASSOCIATED NUCLEASE 1 PREVENTS CAG REPEAT EXPANSION BY PROMOTING REPEAT DELETION DURING DNA BASE EXCISION REPAIR .....	26
ABSTRACT .....	26
INTRODUCTION .....	27
MATERIALS AND METHODS.....	31
RESULTS.....	40
DISCUSSIONS.....	52
CHAPTER 2: R-LOOPS MEDIATE TRINUCLEOTIDE REPEAT INSTABILITY VIA MODULATION OF COORDINATED DNA BASE EXCISION REPAIR ENZYMATIC ACTIVITIES .....	55
ABSTRACT .....	56
INTRODUCTION .....	57
MATERIALS AND METHODS.....	59
RESULTS.....	65
DISCUSSION .....	85
CHAPTER 3. FLAP ENDONUCLEASE 1 CLEAVES RNAs TO RESOLVE R- LOOPS THROUGH DNA BASE EXCISION REPAIR.....	91
ABSTRACT .....	91
INTRODUCTION .....	92
MATERIALS AND METHODS.....	96
RESULTS.....	100
DISCUSSIONS.....	114
REFERENCES .....	120
VITA .....	138

## LIST OF FIGURES

CHAPTER	PAGE
INTRODUCTION	
Figure I.1 DNA structures and hydrolytic attacks on DNA bases and sugars.....	2
Figure I.2 Types of DNA damage and DNA repair mechanism. ....	6
Figure I.3 Base excision repair pathway.....	10
Figure I.4 DNA glycosylases: in DNA repair and beyond .....	12
Figure I.5 Types of TNR related diseases. ....	15
Figure I.6 A balance between the addition and removal of TNRs via BER governs TNR stability.....	18
Figure I.7 Formation of an R-loop during gene transcription .....	18
Figure I.8 R-loop formation is favored by multiple factors.....	19
Figure I.9 The multicellular functions of R-loops.....	21
Figure I.10 The formation of R-loops in TNRs results in repeat instability. ....	23
CHAPTER 1	
Figure 1.1 FAN1 is recruited to CAG repeats in HD differentiate neurons after bromate treatment. ....	41
Figure 1.2 FAN1 promotes CAG repeat deletion during long-patch BER. ....	43
Figure 1.3 FAN1 cleavage of CAG repeats during BER. ....	45
Figure 1.4 FAN1 and FEN1 coordinate to cleave CAG repeats during BER .....	47
Figure 1.5 FAN1 and FEN1 cooperate to promote efficient BER on CAG repeats. ....	49
Figure 1.6 FAN1 promotes the removal of CAG repeats hairpin intermediate during BER. ....	51
Figure 1.7 A hypothetical model of FAN1 promoting or attenuating CAG repeat deletion through BER. ....	55

## CHAPTER 2

Figure 2.1 R-loop and R-loop intermediate annealing..	64
Figure 2.2 APE1 incision of an abasic site on TNR R-loops.....	67
Figure 2.3 Pol $\beta$ DNA synthesis on TNR R-loops .....	69
Figure 2.4 Pol $\beta$ DNA synthesis on the nicked duplex TNR and nicked TNR R-loops. ....	71
Figure 2.5 Removal of the RNA strand in TNR R-loops by bacterial RNase HI restores the efficient synthesis of TNRs by pol $\beta$ .....	72
Figure 2.6 FEN1 cleavage activity of TNRs on R-loops.....	75
Figure 2.7 FEN1 cleavage activity on the nicked TNR R-loops. ....	76
Figure 2. 8 Removal of the RNA strand in TNR R-loops prevents the cleavage of a long repeat flap by FEN1.....	77
Figure 2.9 Coordination of FEN1 flap cleavage and pol $\beta$ DNA synthesis in TNR R-loops during BER. ....	79
Figure 2.10 FEN1 cleaves the RNA strand of TNR R-loops. ....	81
Figure 2.11 BER in R-loops promotes TNR deletion. ....	83
Figure 2.12 BER in TNR R-loops lead to repaired products. ....	84
Figure 2.13 A hypothetical model of BER in R-loops leads to TNR deletion. ....	87

## CHAPTER 3

Figure 3.1 FEN1 endo-cleavage of RNA/DNA hybrids. ....	102
Figure 3.2 FEN1 exo-cleavage of RNA/DNA hybrids. ....	103
Figure 3.3 FEN1 cleavage of RNA and DNA tracking in an RNA/DNA hybrid..	105
Figure 3. 4 FEN1 recruitment to R-loops upon oxidative DNA damage.....	107
Figure 3.5 FEN1 cleavage of RNA on R-loops through BER.....	109
Figure 3.6 FEN1 rate of cleavage between flap and RNA R-loops intermediate.....	110

Figure 3.7 APE1 cleavage of RNA on R-loops intermediates through BER. ....	112
Figure 3.8 APE1 rate of cleavage of RNA on R-loops. ....	113
Figure 3.9 FEN1 and APE1 promotes the resolution of R-loops during BER. ..	114
Figure 3.10 A hypothetical of R-loops mitigation by BER. ....	116

## ABBREVIATIONS AND ACRYONYMS

°C	Degree Celsius
3-meA	3-methyladenine
5',8-cdA	5',8-cyclo-2'-deoxyadenosine
5'-dRP	5'-deoxyribose phosphate
5-hU	5-hydroxyuracil
5-meC	5-methylcytosine
6-FAM	6-carboxyfluorescein
A	Adenine
AAO	Age at onset
AID	Activation-induced cytosine deaminase
AOA2	Ataxia with oculomotor apraxia 2
ALS4	Amyotrophic lateral sclerosis type 4
APE1	AP Endonuclease 1
AP site	Apurinic/apyrimidinic site
APOBEC	Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like
AQR	Aquarius helicase
AR	Androgen receptor
ATP	Adenosine triphosphate
BER	Base excision repair
Bis-Tris	2,2-Bis(hydroxymethyl)-2,2',2'-nitrilotriethanol

bp	Base pair
BRCA1	Breast cancer type 1
BRCA2	Breast cancer type 2
BSA	Bovine serum albumin
C	Cytosine
cdPu	5',8-cyclo-2'-deoxypurines
CSR	Class switch recombination
Cordycepin	3'-deoxyadenosine
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
dRP	Deoxyribose phosphate
dsDNA	Double-stranded DNA
DTT	Dithiothreitol
8-oxoG	8-oxoguanine
<i>E. coli</i>	Escherichia coli
EDTA	Ethylenediaminetetraacetic acid
EME1	Essential meiotic structure-specific endonuclease 1
EXO1	Exonuclease 1
FaPy	Formamidopyrimidine
FA	Fanconi Anemia
FAN1	Fanconi-anemia associated nuclease 1

FANCM	Fanconi anemia complementation complex M
FEN1	Flap endonuclease 1
FMR1	Fragile X mental retardation 1
FPLC	Fast protein liquid chromatography
FRDA	Friedreich's ataxia
FXN	Frataxin
FXS	Fragile X syndrome
G	Guanine
GST	Glutathione S-transferase
HD	Huntington's disease
HEK	Human embryonic kidney
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His	Histadine
HMGB1	high-mobility group protein B1
HR	Homologous recombination
H <sub>2</sub> O <sub>2</sub>	hydrogen peroxide
ICL	Interstrand crosslink
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IR	Ionizing radiation
KCl	Potassium chloride
KBRO <sub>3</sub>	potassium bromate
LB	Lysogeny/Luria broth
LIG I	DNA ligase I



LIG III	DNA ligase III
LP-BER	Long patch-base excision repair
MAM	Methylazoxymethanol
MBD4	Methyl-binding domain glycosylase 4
MD	Myotonic dystrophy
MEF	Mouse embryonic fibroblast
MgCl <sub>2</sub>	Magnesium chloride
miRNA	microRNA
MLH	MutL homologue
MMR	Mismatch repair
MMS	Methyl methanesulfonate
mRNA	messenger RNA
MSH	MutS homologue
Mus81	MUS81 structure-specific endonuclease subunit
MutS $\beta$	MSH2/MSH3
MW	Molecular weight
MYH	MutY homologue
MNNG	N-methyl-N'-nitro-N-nitrosoguanidine
N	Nitrogen
NaCl	Sodium chloride
NaH <sub>2</sub> PO <sub>4</sub>	Sodium dihydrogen phosphate
NaOH	Sodium hydroxide
N <sup>7</sup> -mG	N <sup>7</sup> -methylguanine

N <sup>3</sup> -mA	N <sup>3</sup> -methyl adenine
NEIL	Endonuclease VIII-like
NER	Nucleotide excision repair
NHEJ	Nonhomologous end-joining
NP-40	Nonidet P-40
nM	nanomolar
nt	nucleotide
O <sub>2</sub>	Oxygen
O <sub>2</sub> <sup>-</sup>	Superoxide
O•	Oxygen radical
<sup>1</sup> O <sub>2</sub>	Singlet oxygen
OH•	hydroxyl radical
O <sup>6</sup> -mG	O <sup>6</sup> -methylguanine
OGG1	8-oxoguanine DNA glycosylase
PAGE	Polyacrylamide gel electrophoresis
PARP1	Poly(ADP-ribose) polymerase 1
PCNA	Proliferating cell nuclear antigen
pH	Potential of hydrogen
pI	Isoelectric point
PIP	PCNA-interacting protein
PMSF	Phenylmethylsulfonyl fluoride
pol β	DNA polymerase β
pol δ	DNA polymerase δ

pol ε	DNA polymerase ε
polyA	polyalanine
polyQ	polyglutamine
RAD2	RADiation sensitivity 2
RNA	Ribonucleic acid
ROS	Reactive oxygen species
ROOH	lipid peroxides
ROO•	peroxyl radicals
RPA	replication protein A
SAM	S-adenosylmethionine
SETX	Human senataxin
Sen1	Yeast senataxin
siRNA	small interference RNA
SMUG 1	Single-strand selective monofunctional uracil DNA glycosylase
SN-BER	Single-nucleotide base excision repair
ssDNA	single-stranded DNA
T	thymine
TDG	thymine DNA glycosylase
Tg	thymine glycol
THF	tetrahydrofuran
TL	Translesion synthesis
TNR	Trinucleotide repeat

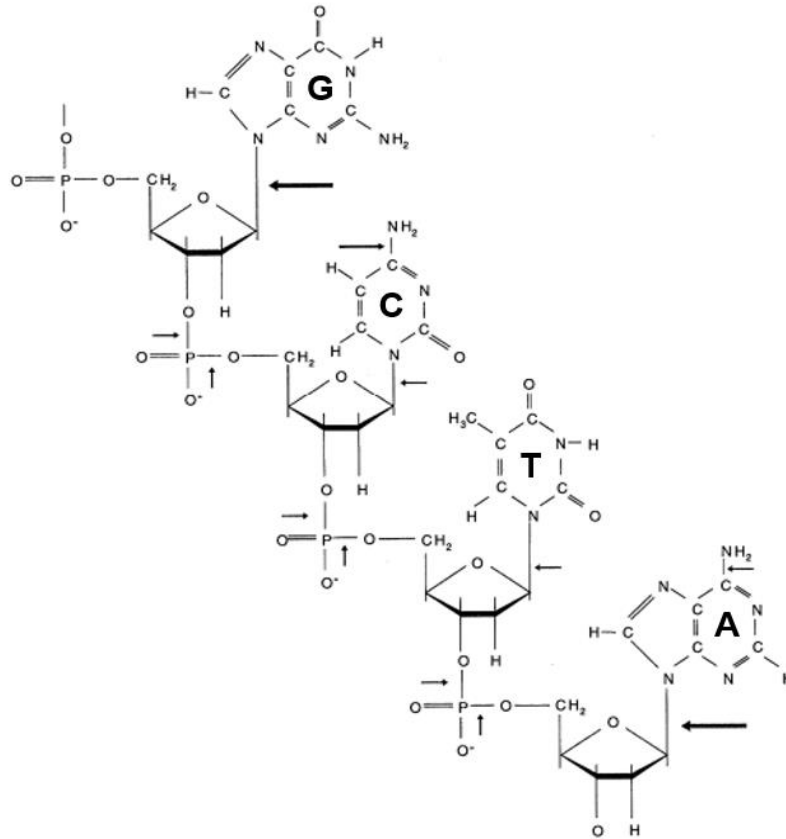
TOP1	DNA topoisomerase 1
TPC	topotecan
Tris	Tris(Hydroxymethyl)aminomethane
U	Uracil
U2OS	Human bone osteosarcoma
UBZ	Ubiquitin-binding domain
ub-PCNA	Ubiquitylated PCNA
UDG	Uracil DNA glycosylase
$\mu\text{M}$	micromolar
UDG	uracil DNA glycosylase
UNG	uracil DNA N-glycosylase
UTR	untranslated region
UV	ultraviolet
WT	wild-type
XPF	Xeroderma pigmentosum complementation group F
XPG	Xeroderma pigmentosum complementation group G
XRCC1	X-ray repair cross-complementing protein 1

## INTRODUCTION

### A. DNA damage and repair

The genetic information of all living creatures is encoded in long chains of deoxyribonucleic acid (DNA) sequences that form the genome. The genetic information is frequently passed through generations in a semi-conserved manner. Thus, the stability of the genome is crucial for the preservation of the species. Beneficial modifications in the genome facilitate the survival of organisms and their adaptation to the environment and are the driving force of evolution. However, deleterious alterations of the genome can lead to the death of organisms. Countless modifications in our genome are the product of DNA insults generated from endogenous or exogenous sources of DNA damaging agents, such as reactive oxygen species (ROS) from cellular respiration, metabolic processes, inflammatory response, environmental toxicants, among others. Most of DNA damage is caused by the simple, spontaneous hydrolysis of DNA bases. i.e., the hydrolysis of the *N*-glycosidic bond of purines, adenine (A) and guanine (G), called depurination [3] (Fig. 1.1). Depurination is the simplest DNA damage frequently generated under physiological conditions (acid/base catalysis). It is estimated that about 10,000 DNA purine bases are hydrolyzed per cell per day [7]. Moreover, spontaneous deamination of cytosines (C) and 5-methylcytosine (5-meC) that can convert the bases to uracil (U) and thymine (T), also frequently occurs under physiological conditions [8]. The mechanism of this conversion is the direct deamination by alkali-catalyzed hydrolysis and the nucleophilic attack by water on the protonated base in an acid-catalyzed reaction [9]. Cytosine deamination occurs

at a significant rate under physiological conditions. However, it is less frequent than depurination [10]. A damage-specific DNA glycosylase subsequently recognizes the DNA damage and ultimately repaired by DNA base excision repair (BER) [3].



**Figure I.1 DNA structures and hydrolytic attacks on DNA bases and sugars [3].**

In many organisms, the majority of DNA base lesions arise from oxidative stress generated under normal cellular processes. The DNA molecule is susceptible to spontaneous depurination and depyrimidination via the attack by oxygen radicals during mitochondria cellular respiration during which mitochondrial

oxidative phosphorylation leads to the production of ATP, while it can give rise to ROS such as hydrogen peroxides ( $\text{H}_2\text{O}_2$ ), superoxide anion radicals ( $\text{O}_2^-$ ), hydroxyl radicals ( $\bullet\text{OH}$ ), lipid peroxides ( $\text{ROOH}$ ), peroxy radicals ( $\text{ROO}\bullet$ ) among others [11]. Currently, over 100 DNA modifications resulting from oxidative DNA damage have been identified as being caused by endogenous or exogenous sources of DNA damage [12]. The most common form of oxidative DNA damage created by ROS is 8-oxoguanine (8-oxoG). Reactive oxygen species can also create other types of DNA base lesions such as thymine glycol, N-substituted ureas, and hydantoins. In addition, ROS can induce lipid peroxidation and cause exocyclic adducts on DNA bases [13, 14]. Hydroxyl radicals can also result in bulky DNA base damage such as cyclopurines and interstrand crosslinks, which hinders DNA replication causing mutations during the repair [15, 16]. Furthermore, normal cellular metabolic intermediates such as S-adenosylmethionine (SAM) can result in alkylated DNA base damage. S-adenosylmethionine is a methyl group donor used as a cofactor in many cellular functions, including amino acid metabolism, RNA and DNA methylation, histone methylation, and immune response [17, 18]. S-adenosylmethionine can also donate its methyl group to DNA bases in the absence of methyltransferases leading to the alkylation of DNA bases. Specifically, SAM can cause N<sup>7</sup>-methylguanine (N<sup>7</sup>-mG), N<sup>3</sup>-methyl adenine (N<sup>3</sup>-mA), and O<sup>6</sup>-methyl guanine (O<sup>6</sup>-mG) [19]. N<sup>3</sup>-methyl adenine can inhibit RNA and DNA polymerases leading to genotoxic effects. 8-oxoguanine is highly mutagenic because it can base pair with adenines instead of cytosines during DNA replication and repair [20]. N<sup>7</sup>-methylguanine is not mutagenic but can generate single-strand

DNA (ssDNA) breaks [21]. On the other hand, O<sup>6</sup>-mG is mutagenic because it can base pair with thymine rather than cytosine residues [22]. N<sup>7</sup>-methylguanine and N<sup>3</sup>-mA can be recognized by DNA glycosylase and go through BER [23].

Environmental factors are the major sources of exogenous DNA damage, which can cause ssDNA, double-strand DNA (dsDNA) breaks, nucleotide misincorporation, DNA adducts, and DNA crosslinks. Multiple sources that can induce alkylating DNA base damage include anticancer drugs, plant byproducts, and antiseptic chemicals, including methyl methanesulfonate (MMS) and N-methyl-N'-nitro-N-nitrosoguanidine (MNNG), temozolomide, methylazoxymethanol (MAM), β-propionolactone, among others [24]. These types of alkylating DNA damaging agents can create N<sup>7</sup>-mG, N<sup>3</sup>-mA, as well as O<sup>6</sup>-mG [24]. It has been shown that MMS alkylation is 20,000 folds higher than from endogenous SAM reactivity [25]. Ultraviolet radiation (UVR) is another source of environmentally-induced DNA damage [26]. For instance, UV-A can generate singlet oxygen (<sup>1</sup>O<sub>2</sub>), which can indirectly photosensitize DNA bases and induce ssDNA breaks. [27]. Conversely, UV-B radiation can promote chemical modification in DNA bases and cause the formation of complex DNA damage such as pyrimidine dimers, which are recognized and repaired by NER. These DNA base lesions can result in nucleotide misincorporation during replication and repair [8]. Also, tobacco smoke can result in the formation of bulky DNA adducts such as O<sup>4</sup>-POB-dT and O<sup>2</sup>-POB-dT that are subjected to repair by NER [28].

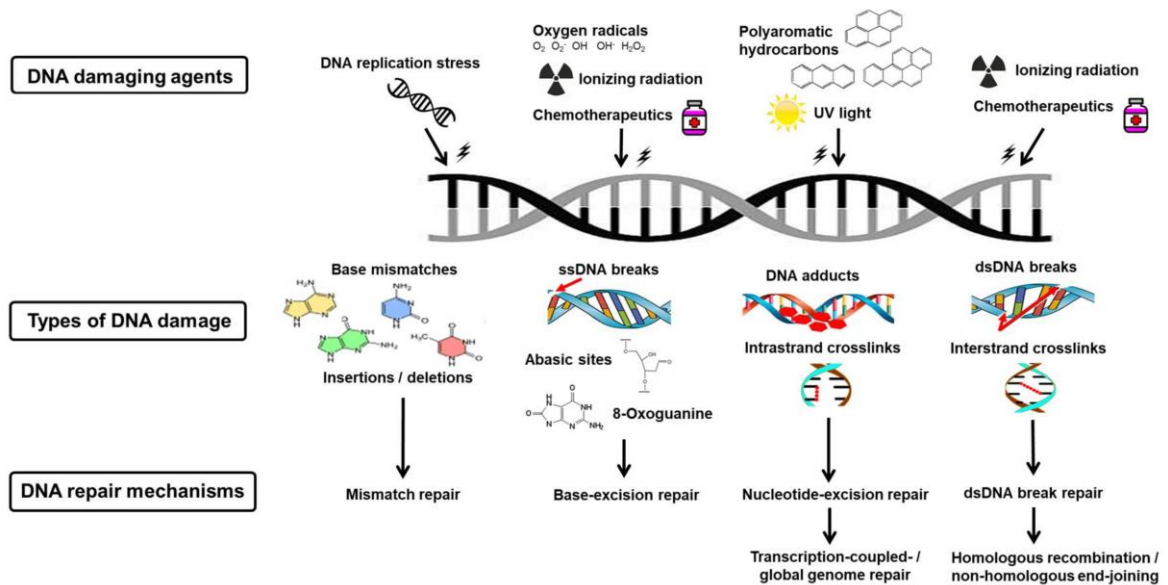
Failure to repair DNA base lesion can result in DNA replication errors leading to genome instability, mutations, and cell death. Damage to the DNA can



cause nucleotide misincorporation by DNA polymerases or DNA sequence misalignment during DNA replication and repair. High levels of oxidative or alkylating DNA damage can cause accumulation of ssDNA breaks, ultimately leading the formation of a double-strand break, chromosomal breakage, and cell death. As a physiological process, dsDNA breaks occur during meiosis when chromosomal crossover takes place between sister chromatids as well as during immunoglobulin class switch recombination, and transcriptions/replication when transcription clashes with the replication fork [29].

To maintain genome stability and integrity, cells have evolved different mechanisms to prevent and combat a variety of DNA lesions (Fig. 1.2) [1]. The front line of the mechanisms is to remove ROS before they can damage the DNA by cellular antioxidants such as catalases, superoxide dismutase, peroxidase, glutathione, cysteines, vitamin C, and vitamin B, among others. [30]. Removal of existing DNA damage by different DNA repair pathways is the second line of defense in combating the adverse consequences of DNA damage. Oxidized and alkylated DNA base lesions can be recognized by a damage-specific DNA glycosylase, 8-oxoGs, uracil, T/G mismatches, and alkylated DNA bases are recognized by 8-oxoguanine DNA glycosylase (OGG1), uracil DNA glycosylase (UNG), thymidine DNA glycosylase (TDG), and 3-alkyladenine-DNA (AAG), respectively leading to an abasic (AP) site or ssDNA break, which initiates the BER pathway. Misincorporated nucleotides can activate the DNA mismatch repair (MMR) pathway. The MMR can coordinate with MutYH, a DNA glycosylase to remove a misincorporated adenine as oppose to an 8-oxoG preventing mutations

[31]. On the other hand, for an O<sup>6</sup>-mG that misbasepairs with thymine, the MMR pathway serves as a direct signal for cell cycle arrest inducing a DNA damage response or apoptosis [32]. Furthermore, the MMR pathway can recognize and repair double helix misalignments created by small insertion or deletion loops during DNA replication and repair [33].



**Figure I.2 Types of DNA damage and DNA repair mechanism [1].**

Bulky DNA adducts that can distort the DNA helix, interfere with DNA base pairing, and obstruct gene transcription or DNA replication are recognized by nucleotide excision repair (NER) pathway. The NER resolves bulky adducts by removing a long fragment of DNA containing the lesion. The NER pathway scans for the distortions of the DNA helix and recognizes a distorted DNA containing a bulky DNA lesion. The removal of the DNA lesion involves the excision of 25-30 DNA fragment containing the damaged nucleotide leaving a large DNA gap that is

filled in by DNA polymerase  $\delta$  [34]. A nicked DNA is then sealed by DNA ligase. Interstrand crosslinks are recognized by the Fanconi Anemia (FA) repair pathway. The FA pathway is also involved in resolving a stalled replication fork [35].

A high level of ssDNA breaks or stalled replication fork can lead to dsDNA breaks that are subject to homologous recombination (HR) or nonhomologous end-joining (NHEJ) or NHEJ [36, 37]. Homologous recombination is a mechanism to repair dsDNA breaks in proliferating cells. During HR, the damaged DNA strand invades into the sister chromatid that is used as a template for the DNA synthesis for completion of the repair. The dsDNA breaks can also be repaired through the NHEJ pathway in both proliferating and nonproliferating cells. The NHEJ pathway does not require homology of the two DNA duplex strands and only requires a few base pairs that connect the two broken DNA ends, generating a nick that is sealed by DNA ligase IV (LIG IV). During NHEJ, the Ku 70/80 protein recognizes the broken ends of dsDNA strands recruiting other repair proteins including DNA protein kinase (DNA-PK), XRCC4 and LIG IV for rejoining the broken ends [38]

In a scenario where DNA lesions escape the detection by DNA repair machinery during DNA replication and repair and gene transcription allowing DNA damage accumulation on the replication forks and DNA repair and transcription intermediates, these lesions have to be bypassed by an error-prone translesion DNA synthesis (TLS) via a variety of translesion DNA polymerases [39] to avoid replication fork stalling, chromosomal breakage, DNA sequence rearrangements, and cell death maintaining the integrity of the genome.

## **B. The DNA Base Excision Repair Pathway**

The BER pathway is a highly conserved system that exists in simple organisms such as *E. coli* and yeast to multicellular eukaryotes such as humans. The BER pathway is responsible for repairing the most common form of DNA damage in mammalian cells, i.e., DNA base damage resulting from deamination, depurination, oxidation, alkylation. Base excision repair is also responsible for the repair of ssDNA breaks. Thus, BER is efficient in protecting cells from DNA damage by removing the most frequently produced DNA lesions, especially in mammals.

The BER pathway is mainly accomplished by five sequential steps during which the coordination of BER enzymatic activities is essential for the efficient removal of damaged bases [4] (Fig. 1.3). The first step of BER is the recognition and excision of the modified base to generate an abasic site. The second step is the incision of 5'-end of the abasic site to generate an ssDNA break intermediate with a 5'-deoxyribose phosphate (dRP) residue. In the third step, the dRP group is removed generating a one-nucleotide gap. In the fourth step, the gap is filled generating a nick, which will be sealed in the fifth step that completes the repair.

Base excision repair is initiated by the removal of a damaged base by a damage-specific DNA glycosylase, leaving an abasic site. Subsequently, AP endonuclease 1 (APE1) incises the 5'-end of the abasic site leading to ssDNA break intermediate with dRP residue. Then the repair diverges into two sub-pathways depending on the number of nucleotides replaced by DNA synthesis by DNA polymerase  $\beta$  (pol  $\beta$ ), designated as the single-nucleotide (SN)-BER and

long-patch (LP)-BER [40]. The SN-BER is initiated by the monofunctional glycosylase that leaves an abasic site, which is further incised by APE1 [41]. The cleaved abasic site forms an ssDNA break containing a 5'dRP native sugar, which is processed by the lyase activity of pol  $\beta$  through  $\beta$ -elimination [42, 43]. The lyase activity of pol  $\beta$  leaves a one-nucleotide gap that is filled by pol  $\beta$  DNA synthesis activity creating a nicked DNA that is recognized by either DNA ligase I (LIG I) or XRCC1/DNA ligase III $\alpha$  (LIG III $\alpha$ ) complex to complete the repair [44, 45]. However, if the sugar is oxidized or reduced, they cannot be removed by pol  $\beta$  5'-dRP lyase activity. The removal of the modified sugar will subject the LP-BER that involves the coordination between pol  $\beta$  or replication polymerases and flap endonuclease 1 (FEN1) [43, 46, 47]. The LP-BER is accomplished through the incorporation of 2-13 nucleotides by the DNA polymerases either via the "hit and run" mechanism mediated by pol  $\beta$  and FEN1 functional coordination [4] (Fig. 3) or strand displacement synthesis by pol  $\beta$  or replication polymerases and other repair DNA polymerases creating a 5'-flap containing the modified dRP group that is removed by FEN1, the multi-functional 5'-3' endo- and exonuclease [42, 48]. Other DNA polymerases that perform the strand displacement synthesis during BER include pol  $\delta$ , pol  $\epsilon$ , and pol  $\lambda$  through their coordination with BER cofactors, poly(ADP-ribose) polymerase 1 (PARP1) and proliferating cell nuclear antigen (PCNA) [47, 49-55]. The nick resulting from FEN1 flap cleavage is ligated by LIG I to complete LP-BER [4, 56].

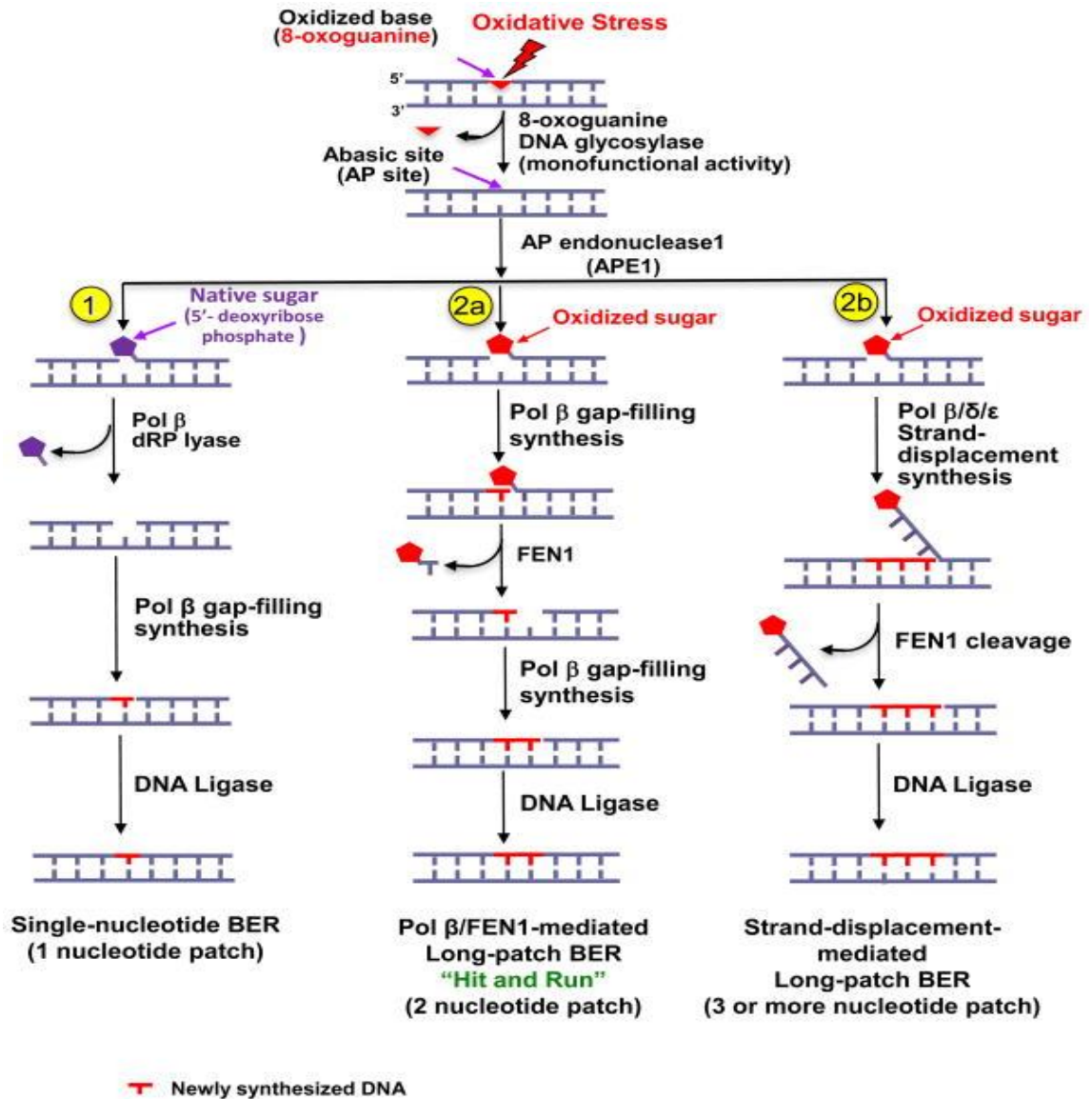
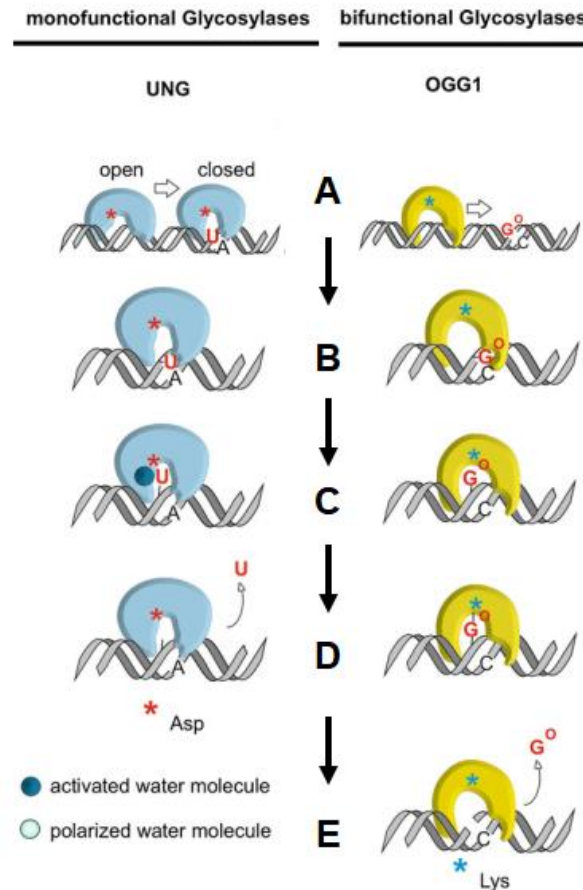


Figure I.3 Base excision repair pathway [4]

The BER pathway is initiated through the recognition and removal of damaged DNA bases by DNA glycosylases. The enzymes scan the genome for searching for specific DNA base damage by flipping each base out of the DNA helix. In mammals, there are eleven DNA glycosylases that specialize in the recognition and cleavage of a variety of damaged DNA bases. The uracil DNA

glycosylase (UNG), and the single-strand specific monofunctional uracil DNA glycosylase 1 (SMUG) specifically remove uracil and 5-fluorouracil (5-FU) generated during cancer therapy [57, 58]. The thymine DNA glycosylase (TDG) and the methyl-CpG binding domain glycosylase 4 (MBD4) remove thymine mismatches from G-T as well as uracil and 5-FU [59, 60]. The endonuclease III-like 1 (NTHL1) removes oxidized, ring-fragmented or saturated pyrimidines, such as formamidopyrimidine (FaPy), 5-hydroxyuracil (5-hU), and thymine glycol (Tg) [61, 62]. The MutY homolog DNA glycosylase (MutYH) removes adenosine (A) opposite 8-oxoG, guanine or cytosine and 2-hydroxyadenine (2-hA) opposite G [63]. 8-oxoguanine DNA glycosylase 1 removes the 8-oxoG and excise other oxidized pyrimidines or ring-fragmented purines such as FaPy [64, 65]. The 3-methyl-purine glycosylase (MPG) removes the alkylated purines like N<sup>3</sup>-mA, N<sup>7</sup>-mG, 3-methylguanine (3-mG), and hypoxanthine [66, 67]. The Nei-like (NEIL) DNA glycosylase 1 and 2 remove oxidized pyrimidines such as 8-oxoG, Tg, FaPy, and 5-hydroxycytosine (5-hC) on double-stranded DNA [68], while NEIL3 preferentially removes FaPy on single-stranded DNA [69]. The DNA glycosylases can be further classified as monofunctional and bifunctional enzymes based on the mechanism of enzymatic reactions. The monofunctional DNA glycosylases use a water molecule to perform a nucleophilic attack at the N-glycosidic bond between the modified base and the sugar-phosphate backbone leading to the formation of an abasic site [70]. Bifunctional glycosylase can cleave both the glycosidic bond and the DNA backbone 3'-end of the damaged base through a Schiff-base intermediate

formed by lysine [71] (Fig. I.4). The abasic site generated by DNA glycosylases on the DNA is later recognized and incised by APE1.



**Figure I.4 DNA glycosylases: in DNA repair and beyond [6]**

AP endonuclease 1 has a 5'-endonuclease activity, which incises the DNA backbone at the 5'-end of the abasic site [72, 73]. AP endonuclease 1's incision of the abasic site generates a 3'-hydroxyl (3'-OH) and a 5'-dRP termini, which can further be processed by other BER enzymes [73]. AP endonuclease 1 also possesses a 5'-3' exonuclease activity where it can remove the damage at the 3'-



ends, such as 3'-phosphoglycolate esters, 3'-phosphates, and 3'-deoxyribose residue generated by the hydrolysis of the phosphodiester bond by either the bifunctional DNA glycosylases or by oxygen radicals direct attack to the DNA backbone. These result in ssDNA break intermediates that are subject to the cleavage by APE1 3'-exonuclease [74]. The removal of these blocking groups the 3'-end by APE1 allows the extension of the upstream strand by DNA polymerases and nick ligation by DNA ligases. Furthermore, APE can also cleave 3' misincorporated nucleotides to serve as a proofreading function for DNA polymerase, which lacks this activity, such as DNA polymerase  $\beta$  (pol  $\beta$ ) during BER [75].

The functional coordination mediates efficient BER among the multiple BER proteins and cofactors (Fig. 1.4). APE1 can enhance OGG1 activity by increasing the turnover rate of OGG1 in generating an abasic site [76]. It has been shown that APE1 binding to an abasic site dislodges OGG1 stimulating the dissociation of the glycosylase from its product, an abasic site, thereby increasing the catalytic efficiency of the glycosylase.

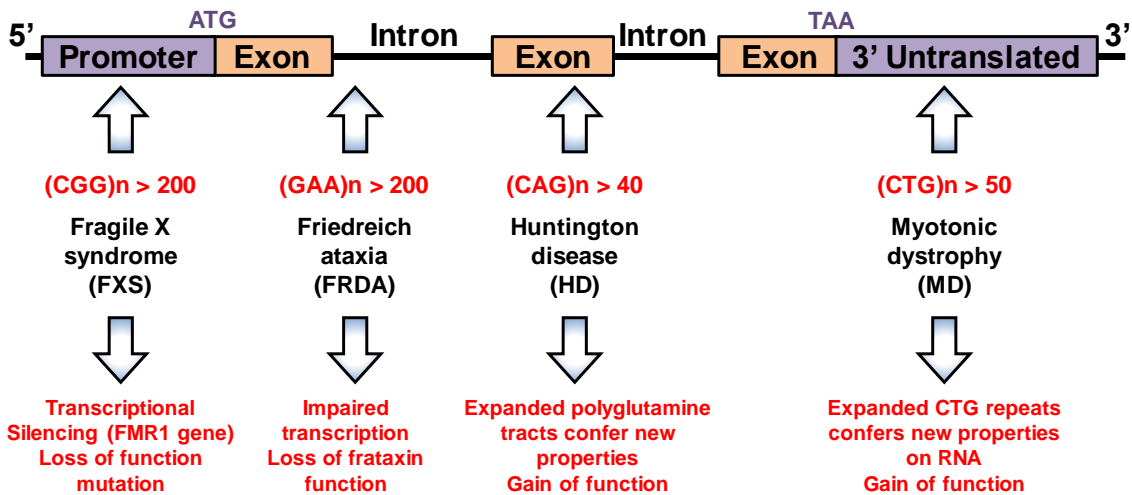
Moreover, APE1 recruits pol  $\beta$  and stimulates its dRP lyase activity and gap-filling DNA synthesis activity during SN-BER [77, 78]. Additionally, APE1 enhances the enzymatic activity of both FEN1 and LIG I during LP-BER by physically interacting with these proteins [79]. Similarly, PCNA stimulates FEN1 cleavage and multiple DNA polymerases syntheses to promote the removal of the damaged strand during LP-BER [51, 80]. Also, PARP1 can stimulate DNA synthesis and hence regulate SN-BER and LP-BER pathway by interfering with APE1/pol  $\beta$

interaction and pol  $\beta$  gap-filling synthesis during the repair [81]. Moreover, PARP1 can inhibit LP-BER by repressing the pol  $\beta$ /FEN1 “gap translation” mechanism [81]. Recent studies have also shown that MSH2 and MSH3, the MMR proteins, stimulate pol  $\beta$  DNA synthesis in trinucleotide repeats during BER [82]. Thus, these studies indicate that repair enzymes from other DNA repair pathways may crosstalk with the BER pathway to facilitate the removal of DNA base lesions and maintain genome stability. In addition, BER scaffolding proteins and cofactors such as PARP1 and X-ray repair cross-complementation protein (XRCC1) are also recruited to DNA base lesions to protect the ssDNA break from further recruiting the downstream BER enzymes [40, 83-85] increasing the efficiency of BER.

### **C. BER of DNA base damage as a regulator of trinucleotide repeat instability**

Trinucleotide repeats (TNR) are microsatellites DNA that occurs in a variety of genes involved in the regulation of gene expression and protein function [86]. TNR expansion is the cause of over 40 neurodegenerative diseases, including Huntington’s disease (HD), Friedreich’s ataxia (FRDA), myotonic dystrophy (DM), fragile X syndrome (FXS), among others [87, 88] (Fig. I.5). Trinucleotide repeats undergo dynamic length changes somatically due to their propensity of forming secondary structures such as hairpins, loops, G-quadruplexes, and bulky non-B form structures such as R-loops during DNA replication, repair, and gene transcription. These TNR structures modulate the activities of DNA repair proteins that, in turn, result in repeat instability. Thus, DNA repair appears to play a critical role in modulating TNR instability. Trinucleotide repeats are rich in Gs, As, and Cs

that form hotspots of DNA base damage. This guanine rich sequence allows BER to be actively involved in the repair of DNA base lesions in TNRs through which the instability of repeats is modulated.



**Figure I.5 Types of TNR related diseases.**

A role of BER in mediating TNR expansion is demonstrated by the fact that OGG1 deficiency in HD mice attenuates CAG repeat expansion [89]. It has been shown that during BER in TNRs, ssDNA breaks are generated through the removal of damaged DNA bases and incision of abasic sites by DNA glycosylases and APE1. This subsequently allows the dissociation of the repeats on the damaged strand from its template forming the secondary structures such as hairpins along with a multi-nucleotide gap. The gap is filled by pol  $\beta$  generating an extra number of TNRs. The hairpin inhibits FEN1 conventional flap cleavage and forces the enzyme to perform its alternate flap cleavage on a short flap attached to the hairpin generating a nick. LIG I then seals the nick and ligates the hairpin with the newly synthesized DNA strand resulting in repeat expansion [90] (Fig. I.6).

Studies from our group have found that BER in a TNR tract also results in repeat deletion [91]. We have demonstrated that a TNR hairpin can form on the template strand. Polymerase  $\beta$  then skips over the hairpin to perform DNA synthesis to displace the downstream repeats into a flap. This skip over allows FEN1 to remove more repeats than those synthesized by pol  $\beta$  resulting in repeat deletion [91]. We further demonstrate that the location of a DNA base lesion in a TNR tract governs the repeat deletion or expansion [92]. The BER of a base lesion located in the middle of TNRs preferentially leads to repeat deletion, whereas a base lesion located at the 5'-side of the repeat tract results in repeat expansion. On the other hand, BER of a base lesion located at the 3'-side of the repeats does not affect repeat instability [92]. Interestingly, we have discovered that BER of a base lesion located in a TNR hairpin loop leads to the opening of the hairpin converting it into a double-flap intermediate for which the 5'-flap is removed by FEN1, whereas the 3'-flap is removed by a 3'-flap endonuclease, Mus81/Eme1 leading the removal of a hairpin and prevention of repeat expansion [93]. Furthermore, we have shown that the regulatory role of BER in modulating TNR instability is also modified by the coordination between BER enzymes and cofactors. During BER in a base lesion in a hairpin loop, the 3'-exonuclease activity of APE1 can progressively cleave the 3'-end of the upstream strand. In coordinating with FEN1 cleavage of 5'-flap, this leads to the resolution of the hairpin. Moreover, APE1 can stimulate the ligation of a nick by LIG I, preventing repeat expansion [94]. Also, during BER in a TNR hairpin loop, PCNA can slide down to the downstream to facilitate the annealing of the 3'-flap to the template

strand. This 3' flap reannealing creates a 5'-TNR flap that is efficiently removed by FEN1 leading to the removal of the hairpin [95]. On the other hand, during BER in TNRs, MMR protein complex, MSH2-MSH3 binds to the hairpin or loops formed on the template strand and interacts with pol  $\beta$  to guide the polymerase to copy the repeats within the hairpin or loop, thereby promoting repeat expansion [82]. Besides, we also show that an oxidized sugar in a TNR tract can crosslink with the dRP lyase domain of pol  $\beta$  as well as LIG I to inhibit the activities of the enzymes resulting in the accumulation of ssDNA breaks in the TNR tracks [96], potentially promoting the repeat instability through DNA recombination. Also, we have identified a critical role of dRP lyase domain in preventing TNR slippage and deletion through interaction between several lysine residues in the dRP lyase domain with the dRP group [97]. All these findings indicate an active role of BER and DNA base damage in modulating TNR instability by coordinating with BER cofactors, including MMR proteins. Our results further suggest that there is a balance between the addition and removal of TNR mediated by BER that maintains the stability of repeats. Disruption of the balance through the coordination among BER enzymes and cofactors leads to repeat deletion or expansion (Fig. I.6).

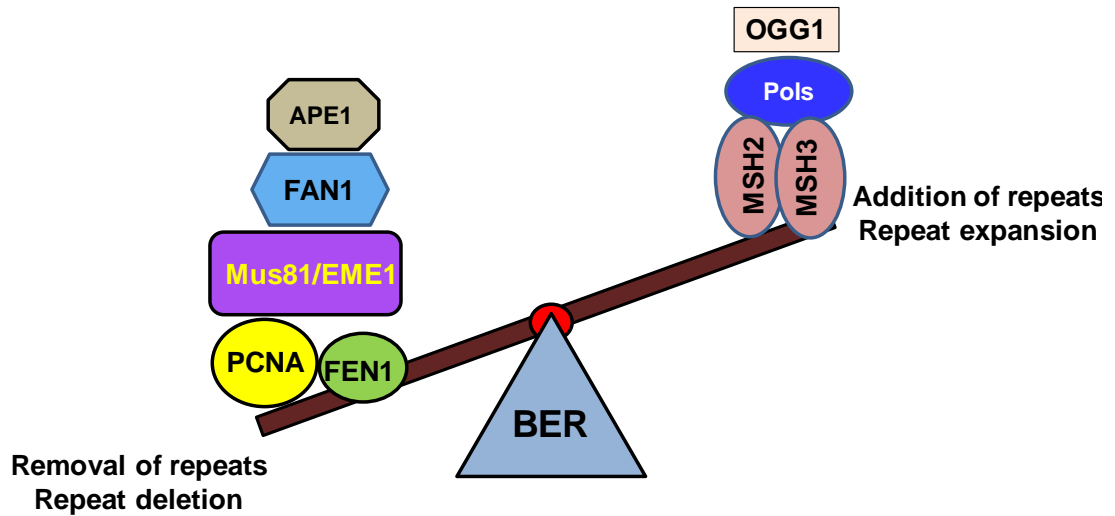


Figure I.6 A balance between the addition and removal of TNRs via BER governs TNR stability.

#### D. R-loop biology and DNA Repair

R-loops are three-stranded nucleic acid non-B form structure, which contains a displaced non-template ssDNA and an RNA:DNA hybrid generated during gene transcription [5, 98] (Fig. I.7). RNA:DNA hybrids are more stable than duplex DNA because they can adopt an intermediate conformation that resembles part of a dsRNA (A-form) and a dsDNA (B-form) [99]. R-loops occurs when the

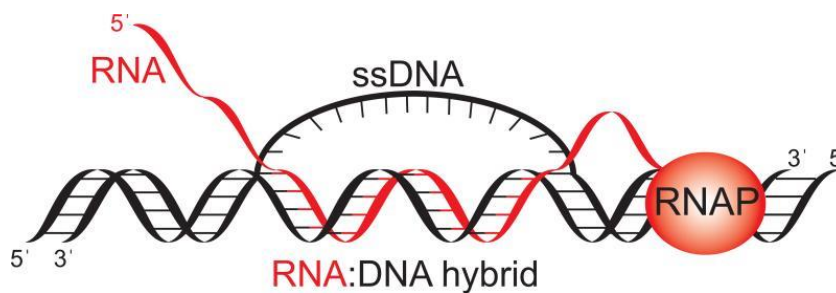
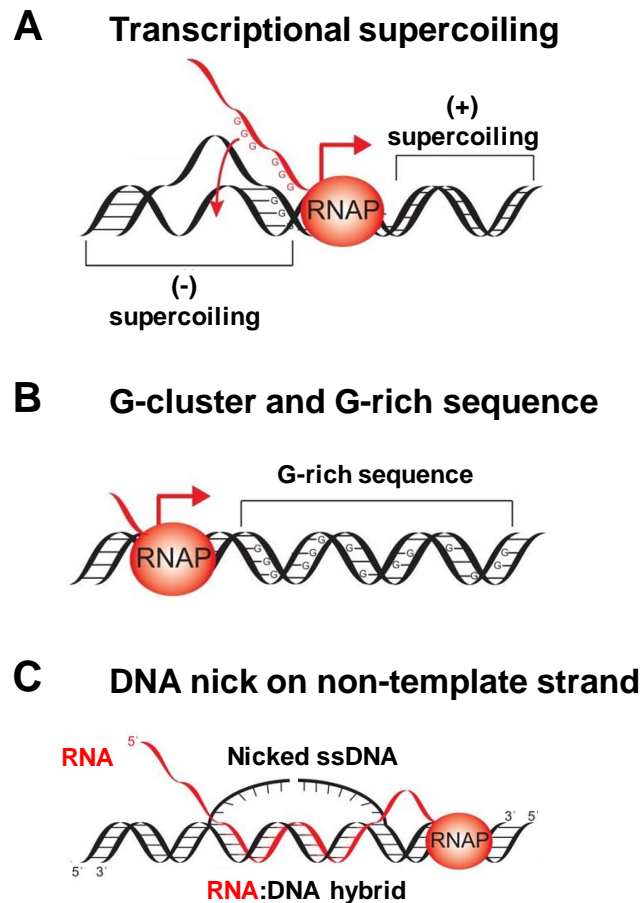


Figure I.7 Formation of an R-loop during gene transcription [2]

nascent RNA transcript synthesized by an RNA polymerase re-hybridizes with the template strand, designated the “thread-back” mechanism [100, 101]. This thread-back mechanism is supported by the crystal structure of the RNA polymerase II, showing that newly synthesized RNAs anneal back to the DNA template strand [102]. Several factors can influence the formation of R-loops. These include transcriptional supercoiling, the abundance of guanine-rich regions on the



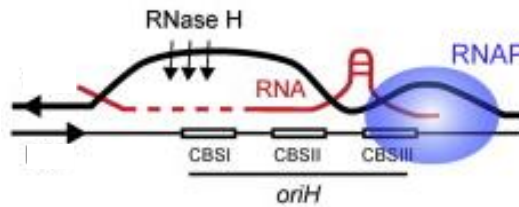
**Figure I.8 R-loop formation is favored by multiple factors [2].** A. Transcriptional supercoiling facilitates R-loop formation by intertwining the RNA:DNA hybrid. B. G-rich sequences facilitates R-loop formation by increasing the G:C base pairing. C. DNA nicks on the non-template strand facilitates R-loop formation by preventing DNA:DNA rehybridization.

template strand, and nicks on the non-template strands [101] (Fig. I.8). The G-rich sequences on the template strand facilitate the formation of R-loops. High G-cluster regions can also promote the stability and elongation of the RNA:DNA hybrid by G:C base pairing [100]. Transcriptional supercoiling can stabilize the RNA:DNA hybrid by intertwining the transcript to the template strand [101]. In addition, a nick formed on the non-template strand can favor RNA:DNA hybrid by preventing the non-template strand from reannealing to the template strand [101].

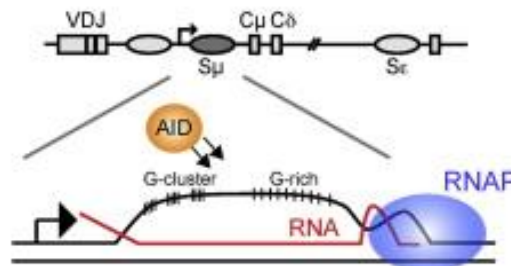
R-loops are involved in multiple cellular functions, including mitochondrial DNA replication, immunoglobulin (Ig) class-switch recombination (CSR), gene expression, and transcription termination [5, 98] (Fig. I.9). In mitochondria, DNA replication starts with the formation of an R-loop at the origin of replication [103]. The RNA:DNA hybrid is then processed by RNaseH1, which cleaves the part of the RNA strand. Subsequently, the residual RNA fragment serves as a primer for DNA synthesis by DNA polymerase  $\gamma$  during mitochondrial DNA replication [104]. Since the region of repetitive switch (S) in the Ig gene contains G-rich sequences, which can promote R-loop accumulation [105, 106], in Ig CSR, the ssDNA region of the R-loops is recognized by B cell-specific activation-induced deaminase (AID) that converts deoxycytidine to deoxyuridines that are then processed by the BER and MMR machinery leaving ssDNA breaks [107-109]. The ssDNA breaks on the S region of the Ig gene are then converted to dsDNA breaks that are subject to NHEJ [110]. During gene transcription, R-loop can accumulate at the pausing sites at the downstream of the poly(A) at the 3' untranslated region,



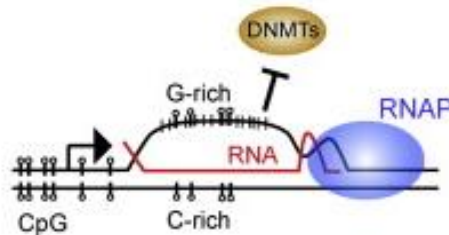
### A Mitochondrial DNA replication



### B Ig class-switch recombination



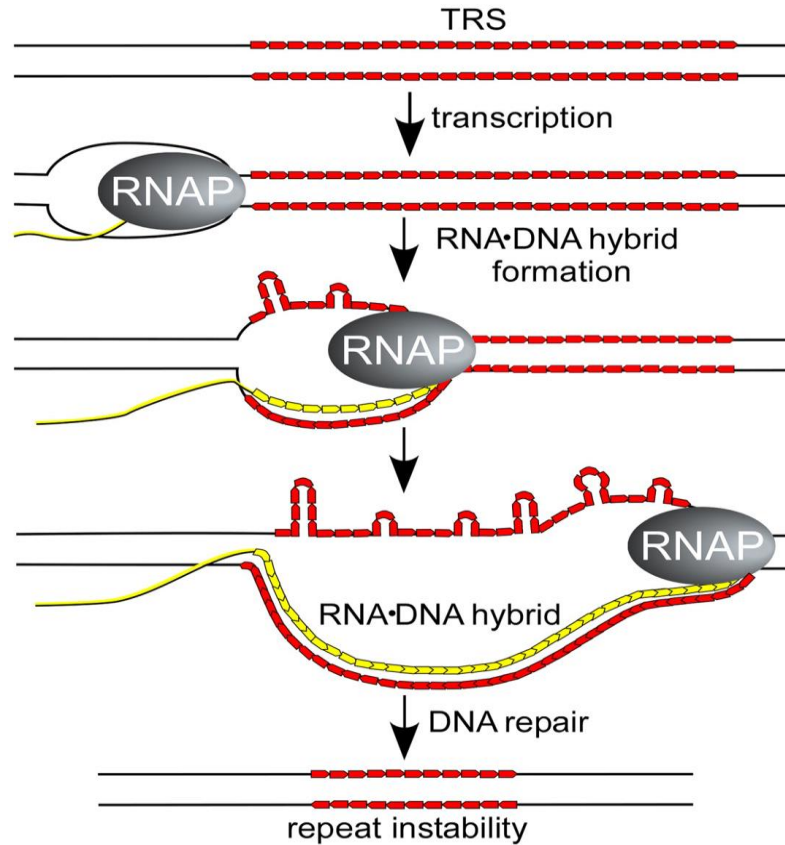
### C CpG-Island Promoter unmethylation



**Figure I.9 The multicellular functions of R-loops [5].** A. Shows R-loops involved in mitochondria DNA replication as the RNA, cleaved by RNase H, is used as primer to start replication. B. Shows R-loop involved in Immunoglobulin (Ig) class-switch recombination as the ssDNA is targeted of deamination by AID. C. Shows R-loop involved in promoting transcription of active genes by blocking repressive DNA methylation from DNA methyltransferases (DNMTs).

signaling the termination of transcription [111]. Also, R-loop can accumulate at the G-rich sequences downstream of non-methylated CpG promoters indicating that it may protect the promoter from being hypermethylated by DNA methyltransferases sustaining gene activation.

R-loops can also sporadically accumulate in the genome. This accumulation can result from the inactivation of RNA processing machinery leading to detrimental effects to cells and causing human diseases, such as neurodegenerative diseases and cancer [98]. Mutations in senataxin (SETX), a putative RNA:DNA helicase, resulting in progressive degeneration motor neurons in the brain and spinal cord [112, 113] and are associated with the neurodegenerative disorder, oculomotor apraxia 2 (AOA2) and amyotrophic lateral sclerosis type 4 (ALS4). SETX plays a role in transcription termination by resolving RNA:DNA hybrids at G-rich pausing sites and promoting the cleavage of the untranslated region of mRNAs [111]. SETX inactivation leads to a deficiency of transcription termination, R-loop accumulation, defective DNA damage response, and genome instability [114-116]. R-loops can readily accumulate at expanded TNRs and has been proposed to play a role in the pathology of TNR expansion diseases [117, 118]. R-loops accumulate at expanded GAA repeats at the frataxin gene in FRDA and recruit the repressive heterochromatinization on the repeats [119], thereby causing transcriptional silencing. *In vitro* and *in vivo* studies have also shown that R-loops formed in TNR tracts cause RNA polymerase arrest and promote TNR instability through DNA repair [117, 120, 121] (Fig. I.10). On the other hand, in cancer, BRCA1 and BRCA2 that are associated with breast and ovarian cancers prevent the formation of R-loops protecting chromosomes and maintaining genome stability [122, 123].



**Figure I.10 The formation of R-loops in TRS results in repeat instability.**

R-loops can result in different forms of genome instability, including mutations, recombination, chromosome rearrangements, and chromosome loss. The non-template strand of R-loops can serve as a substrate for AID [106, 124]. The ssDNA is also susceptible to endogenous and exogenous DNA damage [8, 125]. Thus, AID and other APOBEC family enzymes can convert dC into dU located in the non-template strand of an R-loop. Also, endogenous and exogenous DNA damaging agents can oxidize, alkylate, and deaminate DNA bases or induce depurination and depyrimidination of DNA bases located on the non-template strand of an R-loop. All these can initiate BER and generating ssDNA break intermediates leading to genome instability. Likewise, R-loops can lead to genome

instability through secondary structures such as G-quadruplex and hairpins, which form on the non-template DNA strand of R-loops. These structures can cause replication fork stalling and DNA strand breaks leading to mutations, recombination, and chromosome rearrangement.

To combat the deleterious consequences of persistent R-loops, cells adopt several mechanisms to resolve R-loops. The most important and well-characterized mechanism that resolves R-loop is through the direct removal of the RNA strand of R-loops by RNase H1 and H2 [126]. In addition, several helicases have been shown to be able to unwind RNA:DNA. These include Pif1 DNA helicase [127, 128], DHX9 (RHA) RNA helicase [129], and sen1/SETX [130]. Recent studies have shown that R-loop can also be resolved through DNA repair. It has been shown that NER nucleases xeroderma pigmentosum group G (XPG) and XPF are required to process the R-loops that accumulate in the absence of AQR or SETX or in the presence of the TOP1 inhibitor camptothecin [131] resulting in ssDNA breaks that are subsequently converted into dsDNA breaks. Moreover, a study from the Freudenreich's group has shown that AID-induced ssDNA breaks in the non-template strand in R-loops, leading to the deletion of CAG repeats in yeast suggesting that the ssDNA breaks on an R-loop may be subject to BER through which the R-loop is resolved. The pathways for resolving R-loops need to be elucidated.

## OVERVIEW

Base excision repair is an essential pathway that efficiently repairs DNA base lesions, the most frequently produced DNA damage in mammalian cells. The BER pathway also plays a critical role in active DNA demethylation to regulate epigenetic stability. Furthermore, BER is a crucial pathway in regulating genome stability through the processing of a variety of secondary structures formed in the repetitive DNA sequences in the human genome. Moreover, recent studies also indicate an essential role of BER enzymes in processing RNA during DNA replication, repair, and microRNA biogenesis. These studies suggest that the BER pathway can govern multicellular functions in addition to its role in DNA base lesion repair. However, the molecular mechanisms underlying BER-mediated cellular functions and genome and epigenome instability remain to be elucidated. To address these knowledge gaps, we explored the roles of the coordination between key BER enzymes and BER cofactors in modulating TNR instability through the processing R-loops. We then explored the roles of BER in resolving R-loops through the processing of RNA. Chapter I describes the study on how Fanconi anemia-associated nuclease 1 can process TNRs through BER to prevent TNR expansion by promoting repeat deletion. Chapter II describes the study on how BER in TNR R-loops can lead to TNR deletion. Chapter III describes the study on how BER can resolve an R-loop via the coordination between FEN1 and APE1 processing of the RNA strand during BER.

# **CHAPTER 1: FANCONI ANEMIA-ASSOCIATED NUCLEASE 1 PREVENTS CAG REPEAT EXPANSION BY PROMOTING REPEAT DELETION DURING DNA BASE EXCISION REPAIR**

## **ABSTRACT**

Huntington's disease (HD) is an inherited progressive neurodegenerative disorder and caused by the expansion of CAG repeats in the encoding region of the huntingtin (*HTT*) gene. It has been proposed that the root of these expansions is the formation of secondary structures by CAG repeats during DNA replication and repair. Previous studies have found that DNA base lesions and base excision repair (BER) can mediate CAG repeat instability through the BER core enzymes activities and cofactors. Here, we provide the first evidence that the Fanconi anemia-associated nuclease 1 (FAN1) was recruited to CAG repeats upon oxidative DNA damage to promote CAG repeat deletion through BER. We showed that FAN1 can complement FEN1 through its 5' flap endonuclease and 5' exonuclease activities. We further demonstrated that PCNA and ubPCNA coordinated with FAN1 to facilitate the attenuation of CAG repeat expansion during BER. Our studies suggest that FAN1 can effectively process the secondary structures by coordinating with FEN1 and PCNA during BER, thereby leading to CAG repeat deletions.

## INTRODUCTION

Huntington's disease (HD) is among over 40 trinucleotide repeat (TNR) expansion disorders caused by CAG repeat expansion in the coding region of the (*HTT*) huntingtin gene, resulting in the aggregation of mutant Huntingtin proteins in cerebral cortex neurons leading to progressive and devastating neurodegenerative symptoms [132]. The clinical symptoms for HD include involuntary movement, dementia, speech difficulties, and impaired balance [133]. In HD patients and mouse models, the age at onset of HD depends upon the length of the glutamine-encoding CAG repeat sequences, usually above 36-40 CAG repeats with the longer the repeats associated with earlier age at onset (AAO) [134]. In general, CAG repeat expansion can occur during germline transmission and somatic cellular differentiation indicating that the disease can affect individuals at any stage in life [88, 135]. However, the molecular mechanism underlying CAG instability is not well understood. It has been shown that CAG instability occurs by the formation of slipped strands such as hairpins and loops during DNA replication [136], repair [4], recombination [137], and gene transcription [138]. These lead to the integration of extra repeats into the genome resulting in repeat expansion [139]. Thus, somatic CAG expansion may play a crucial role in promoting the increased CAG repeat length up to the point of reaching the threshold leading to neurodegenerative symptoms of the diseases and thereby governing the age at onset of the disease.

It has been shown that oxidative DNA damage and DNA base excision repair are involved in modulating CAG repeat instability. CAG instability is supported by the post-mortem analysis of HD patients and *HTT* mouse model brains, which showed an increased level of oxidative DNA damage. The early study suggests that disease progression is associated with elevated levels of oxidized DNA bases, 8-oxoguanines (8-oxoGs) [140, 141]. Also, it has been shown that the age-dependent somatic CAG expansion in HD mouse models is associated with the repair of 8-oxoGs through BER [89]. The CAG repeats are GC rich sequences that are hotspots for endogenous and exogenous sources of oxidative DNA damaging agents [8], which can promote CAG expansion through a ‘toxic oxidation cycle’, through which CAG instability occurs by multiple rounds of DNA base damage and repair [88]. Furthermore, yeast models and *in vitro* characterization showed that BER enzymes play a role in mediating somatic CAG instability through DNA repair. It has been proposed that polymerase  $\beta$  (pol  $\beta$ ) plays a significant role in CAG repeat expansion through BER [48, 89, 90, 142]. On the other hand, flap endonuclease 1 (FEN1) prevents CAG repeat expansion and promotes CAG repeat deletion by removing the repeats through long-patch BER [92, 93, 95, 143]. Thus, CAG repeat instability occurs through BER by the loss of coordination between pol  $\beta$  synthesis and FEN1 cleavage [89, 90] and by BER enzyme stoichiometry, which modulates the outcome repeat instability [144, 145].



The enzymatic activities and coordination of BER differ within human individuals depending on age and genetic variations, suggests that somatic CAG repeat instability is governed by the BER capacity of the human body and tissues [146]. Furthermore, it has been suggested that DNA enzymatic activities rather than the DNA damage levels are associated with somatic CAG expansions [145]. In HD mouse models, somatic CAG repeat instability is prevented by a deficiency in 8-oxoguanine glycosylase (OGG1) and mismatch repair proteins, MSH2 or MSH3 mismatch repair protein delaying the occurrence of the mutant huntingtin phenotype [89, 147, 148]. This is because the MSH2 and MSH3 complex can stabilize the downstream hairpin structure, which can inhibit FEN1 5'-flap cleavage, thereby leading to CAG repeat expansion through long-patch BER [82]. Moreover, MSH2 and MSH3 interact and stimulate pol  $\beta$  synthesis of repeats preventing the pol  $\beta$  bypass synthesis over the hairpin formed on the template strand leading to repeat expansion [82, 92, 143]. On the other hand, APE1 can prevent CAG repeat expansion via its 3' exonuclease activity as well as by stimulating DNA ligase I (LIG I) during BER in a hairpin loop. Similarly, Mus81/EME1 3'-flap endonuclease can process the 3'-flap leading the removal of the hairpin via the coordination with FEN1 [93]. Furthermore, the BER cofactor, proliferating cells nuclear antigen (PCNA), a sliding clamp that participates in DNA replication and repair can promote CAG repeat deletion in duplex DNA and prevent repeat expansion by stimulating FEN1 cleavage and the removal of a CAG repeat hairpin structure during BER [95]. Thus, these factors can influence the instability on CAG repeats through BER in the context of the location of DNA base lesions

within the CAG repeat tracts [92, 144] while the activities of BER enzymes and cofactors are modulated during the repair [82, 94-96].

Recent genome-wide association studies (GWAS) have identified single nucleotide polymorphism (SNPs) in the Fanconi-anemia associated nuclease 1 (FAN1) gene as a factor associated with the variation of the age at onset of HD [149, 150]. Furthermore, increased FAN1 expression levels in HD patient stem cells and mouse models is associated with delay of the AAO and slower progression of HD [151]. FAN1 is an evolutionarily conserved nuclease whose biological functions are still unclear. FAN1 was identified to be related to DNA repair based on domain homology with other DNA repair proteins [152], its interaction with mismatch repair proteins [153, 154], and the results from RNA interference screening for sensitivity to DNA interstrand crosslinking agents [155]. FAN1 is a complex protein that possesses multiple structural domains, a RAD18-like ubiquitin-binding zinc finger (UBZ) domain, SAP DNA binding domain, and a virus-type replication repair nuclease (VRR\_nuc) domain [155]. FAN1 exhibits a 5' flap endonuclease activity as a dimer, which is involved in resolving interstrand crosslinks [154-157] and a 5' exonuclease activity as a monomer, which can process DNA overhangs during homologous recombination (HR) [154-158]. The FAN1 UBZ domain is essential for the interaction with the ID complex in the Fanconi anemia (FA) pathway and the interaction with ubiquitylated PCNA (ubPCNA) at the stalled replication fork. FAN1 can physically interact with MLH1, MLH3, and PMS2. However, its biological function has not been elucidated [153].

Recent studies have demonstrated that FAN1 can bind to expanded *HTT* CAG repeats and prevent CAG repeat expansion independently of its nuclease activity [151]. However, it remained to be understood how FAN1 can prevent CAG repeats without using its nuclease activity. Here, we tested if FAN1 can use its 5'-endo/exonuclease to remove CAG repeats leading to repeat deletion. We showed that FAN1 bound to the expanded *HTT* CAG repeats upon oxidative DNA damage. We found that FAN1 can complement FEN1 to remove CAG repeats leading to repeat deletion during BER. Furthermore, we found that FAN1 promoted CAG repeat deletion in duplex DNA and prevented CAG repeat expansion in hairpin structures. Our studies provide the first evidence of FAN1 nuclease can prevent CAG repeat expansion and promote CAG deletion that may potentially delay the AAO of HD.

## **MATERIALS AND METHODS**

### **Materials**

The DNA oligonucleotides were from Integrated DNA Technologies (IDT, Coralville, IA, USA). The deoxynucleotides 5'-triphosphates (dNTPs) were purchased from Sigma-Aldrich (St. Louis, MO). The radionucleotides Cordycepin 5'-triphosphate [ $\alpha^{32}\text{P}$ ] (5000 Ci/mmol) and [ $\gamma^{32}\text{P}$ ]ATP (3000 Ci/mmol) and were purchased from Perkin Elmer Inc. (Boston, MA, USA). Terminal deoxynucleotidyl transferase (TdT) and T4 polynucleotide kinase (PNK) were purchased from Thermo Fisher Scientific (Waltham, MA, USA). Micro Bio-Spin 6 chromatography columns were purchased from Bio-Rad Laboratories (Hercules, CA, USA). All

chemical reagents were purchased from Sigma-Aldrich (St. Louis, MO, USA) and Thermo Fisher Scientific (Waltham, MA, USA).

### **Expression and purification of BER proteins**

The plasmid (pET15b) that expresses recombinant human LIG I was transformed in *E. coli* BL21(AI). One colony of the transformant was inoculated into a 5 ml Luria broth (LB) and incubated at 37 °C overnight. The cell culture then was transferred into a 6 liters LB medium and incubated at 37 °C at 225 rpm until OD<sub>600</sub> reached 0.8. The protein expression was induced with the addition of 1 mM IPTG and 0.1% L-arabinose for 18 hours at 16 °C and collected by centrifugation at 3000 rpm for 30 min at 4 °C. Cells were pelleted and resuspended in lysis buffer which contained 50 mM Tris-HCl pH 7.5, 50 mM NaCl, 1 mM EDTA, 1 mM DTT, 0.1% Nonidet P-40, 1 mM PMSF, and one table of Roche protease inhibitor (Roche, Indianapolis, IN) and subjected to French Press. The cell lysates were centrifuged at 12,000 rpm for 30 min at 4 °C. The supernatant was loaded onto a 20-ml cellulose phosphate (Sigma-Aldrich, St. Louis, MO, USA). Protein fractions were eluted using a linear gradient of NaCl from 50 mM to 1 M. Peak fractions were combined and dialyzed into a buffer containing 50 mM Tris-HCl pH 8.0, 500 mM NaCl, 10 mM imidazole, 7 mM 2-mercaptoethanol, 0.5% inositol, and 1 mM PMSF. Samples were loaded onto a 5-ml Nickel-Nitrilotriacetic acid (Ni-NTA) column. LIG I was eluted using a linear gradient of imidazole from 10 mM to 1M. The peak fractions were combined and dialyzed into a buffer containing 30 mM HEPES pH 7.0, 30 mM KCl, 0.1% inositol, 1 mM DTT, 1 mM EDTA, and 1 mM

PMSF. Samples were then loaded onto a 10-ml Q Sepharose column and eluted using a linear gradient of KCl from 30 mM to 1 M. Purified LIG I was aliquoted and frozen at  $-80\text{ }^{\circ}\text{C}$  until further use.

The plasmid (pET24b) that expresses human recombinant FEN1 was transformed in *E. coli* BL21(DE3). One colony of the transformant was inoculated into a 5 ml Luria broth (LB) and incubated at  $37\text{ }^{\circ}\text{C}$  overnight. The cell culture then was transferred into a two liters LB medium and incubated at  $37\text{ }^{\circ}\text{C}$  at 225 rpm until  $\text{OD}_{600}$  reached 0.6. The protein expression was induced with the addition of 1 mM IPTG for 3 hours at  $37\text{ }^{\circ}\text{C}$  and collected by centrifugation at 3000 rpm for 30 min at  $4\text{ }^{\circ}\text{C}$ . The collected cells were resuspended in FEN1 lysis buffer (30 mM HEPES pH 7.5, 30 mM KCl, 0.5% inositol, 1 mM DTT, 1 mM EDTA, 1 mM phenylmethylsulphonyl fluoride (PMSF) and one table of Roche protease inhibitor) and lysed with a French press cell disruptor (Glen Mills, Clifton, NJ, USA). The lysate was subject to centrifugation at 12,000 rpm for 30 minutes at  $4^{\circ}\text{C}$  and the supernatant was loaded into a 10-ml Q Sepharose column. The flow-through was collected and immediately loaded into a 5-ml CM Sepharose column. FEN1 protein was eluted into a KCl linear gradient from 30 mM to 2 M. The peak fractions were analyzed by SDS-PAGE and dialyzed into FEN1 hydrophobic buffer (30 mM HEPES pH 7.5, 1.7 M  $(\text{NH}_4)_2\text{SO}_4$ , 30 mM KCl, 0.5% inositol, and 1 mM PMSF). The combined fractions were loaded into a 5-ml phenyl Sepharose column and eluted with an  $(\text{NH}_4)_2\text{SO}_4$  linear gradient from 1.7 M to 0 M. The peak fractions were analyzed by SDS-PAGE gel and tested for enzymatic activity and

contamination. Clean fractions were combined and dialyzed into lysis buffer. The samples were loaded into a 1-ml Mono-S column (GE Healthcare, Piscataway, NJ, USA), and eluted using a linear gradient of KCl from 30 mM to 1 M. The purified FEN1 was dialyzed into FEN1 storage buffer (30 mM HEPES pH 7.5, 50 mM KCl, 1 mM DTT, 1 mM EDTA, and 20% glycerol), aliquoted, and frozen at  $-80^{\circ}\text{C}$  until further use.

The plasmid (pT7) that expresses human recombinant APE1 was transformed in *E. Coli* BL21(DE3). One colony of the transformant was inoculated into a 5 ml Luria broth (LB) and incubated at  $37^{\circ}\text{C}$  overnight. The cell culture then was transferred into a 3 liters LB medium and incubated at  $37^{\circ}\text{C}$  at 225 rpm until  $\text{OD}_{600}$  reached 0.8. The protein expression was induced with the addition of 0.5 mM IPTG for 3.5 hours at  $37^{\circ}\text{C}$  and collected by centrifugation at 3000 rpm for 30 min at  $4^{\circ}\text{C}$ . The collected cells were resuspended in APE1 lysis buffer (50 mM HEPES pH 7.5, 30 mM NaCl, 0.5% inositol, 1 mM DTT, 1 mM EDTA, 1 mM PMSF, and one table of Roche protease inhibitor) and lysed with a French press cell disruptor (Glen Mills, Clifton, NJ, USA). The lysate was subject to centrifugation at 12,000 rpm for 30 minutes at  $4^{\circ}\text{C}$  and the supernatant was loaded into a 10-ml Q Sepharose column. The flow-through was collected and immediately loaded into a 10-ml SP Sepharose column. APE1 protein was eluted into a KCl linear gradient from 30 mM to 1 M. The peak fractions were analyzed by SDS-PAGE and dialyzed into APE1 hydrophobic buffer (30 mM HEPES pH 7.5, 1 M  $(\text{NH}_4)_2\text{SO}_4$ , 30 mM KCl, 0.5% inositol, and 1 mM PMSF). The combined sample was loaded into a 5-ml

phenyl Sepharose column and eluted with a  $(\text{NH}_4)_2\text{SO}_4$  linear gradient from 1 M to 0 M. The peak fractions were analyzed by SDS-PAGE gel and tested for specific activity and contamination. Clean fractions were combined and dialyzed into APE1 lysis buffer. The APE sample was loaded into a 1-ml Mono-S column (GE Healthcare, Piscataway, NJ, USA), and eluted using a linear gradient of KCl (30 mM to 1 M). THE purified APE1 was dialyzed into APE1 storage buffer (50 mM HEPES pH 7.5, 50 mM KCl, 1 mM DTT, 1 mM EDTA, and 20% glycerol), aliquoted, and frozen at  $-80^\circ\text{C}$  until further use.

The plasmid (pGEX) that expresses the human recombinant FAN1-GST tag was expressed. FAN1-GST plasmid was transformed into Rosetta (DE3) (Millipore Sigma, Burlington, MA). A single colony was inoculated into a 5 mL LB media containing 100 $\mu\text{g/ml}$  ampicillin and 1% glucose and incubated at  $37^\circ\text{C}$  overnight. The culture was transferred to 0.5 L of LB media and incubated at  $37^\circ\text{C}$  at 225 rpm until  $\text{OD}_{600}$  reached 0.8. The protein expression was induced by first heat shock the culture for 5 minutes at  $42^\circ\text{C}$  and later by the addition of 1 mM IPTG and the incubation for 16 hours at  $18^\circ\text{C}$  at 225 rpm. The cells were collected by centrifugation at 3000 rpm for 30 min at  $4^\circ\text{C}$ . The pellet was resuspended in lysis buffer (20 mM Tris-HCl pH 7.3, 150 mM NaCl, 0.1% NP40, 1 mM EDTA, 1 mM DTT, 30 mM spermidine, 1 mM PMSF, and one tablet of cComplete protease inhibitor) and lysed by using a French press. The lysate was subject to centrifugation at 12,000 rpm for 30 minutes at  $4^\circ\text{C}$ . The supernatant was loaded into a 1 mL Glutathione Sepharose column (GE Healthcare, Piscataway, NJ, USA).

FAN1-GST and the resin were incubated for 2 hours at 4 °C. The column was washed 6 times with 5 mL of FAN1 wash buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 10% glycerol, 1 mM EDTA, 1 mM 2-mercaptoethanol, 1 mM PMSF, and one tablet of cOmplete protease inhibitor). The FAN1 protein was eluted by using 5 mL of FAN1 elution buffer (20 mM Tris-HCl pH 8.0, 20 mM glutathione, 150 mM NaCl, 10% glycerol, 1 mM EDTA, 1 mM 2-mercaptoethanol, 1 mM PMSF, and one tablet of cOmplete protease inhibitor) and incubated for 10 minutes for each 1 mL. The peak fractions were analyzed by SDS-PAGE and tested for enzymatic activity and contamination. Clean fractions were combined, aliquoted, and frozen at -80°C until further use.

### **Construction oligonucleotide substrates**

The DNA oligonucleotide substrates containing a tetrahydrofuran (THF) residue, an analog of an abasic site were designed to mimic a scenario where an abasic lesion occurs in the middle of a (CAG)<sub>20</sub> repeat tract. The guanine at the tenth unit of (CAG)<sub>20</sub> repeats was substituted with either a THF or uracil residue. The sequences of the oligonucleotides are listed in Table 1. (CAG)<sub>20</sub> duplex substrates were constructed by annealing the THF containing oligonucleotide to its template strand at a molar ratio of 1:3. The (CAG)<sub>20</sub> hairpin substrate was constructed by annealing the upstream and downstream primers with the template at a molar ration of 1:3:5:5 in a total of 25 µl annealing reaction. The (CAG)<sub>20</sub> duplex and hairpin substrates were annealed by denaturation at 95 °C for 10 min and cooling down to room temperature.



## Enzymatic assay and repeat size measurement

The BER reactions in the presence of FAN1 were performed by incubating various types of oligonucleotide substrates with purified APE1, pol  $\beta$ , FAN1, FEN1 and LIG I in BER reaction buffer containing 50 mM Tris (pH 7.5), 50 mM KCl, 0.1 mg/ml BSA, 0.1 mM EDTA, and 0.01% NP-40. The BER reconstitution reactions of (CAG)<sub>20</sub> duplex containing an abasic site or a uracil was performed by incubating 10 nM APE1, 5 nM pol  $\beta$ , 50 nM FAN1, and 25 nM LIG I with 25 nM (CAG)<sub>20</sub> repeat-containing a THF substrate and 0.5 units of uracil DNA glycosylase (UDG) with 25 nM (CAG)<sub>20</sub> repeat-containing a uracil substrates. All reaction mixtures (20  $\mu$ l) were assembled on ice in BER reaction buffer in the presence of 50  $\mu$ M dNTPs, 5 mM Mg<sup>2+</sup>, 5 mM ATP, and indicated concentrations of BER enzymes and substrates. The reaction mixtures were incubated at 37 °C for 15 min or 30 min. Subsequently, the reaction mixtures were stopped with a 2x stopping buffer containing 95% deionized formamide and 10 mM EDTA. To determine the length of the repaired products during BER of an abasic site, the repaired products were isolated, and PCR amplified. The amplified repaired products were separated by capillary electrophoresis. The size of CAG repeats were determined by DNA fragment analysis according to the method described previously [159].

**Table 1. Oligonucleotide sequence**

Oligonucleotides	nt	Sequence
Downstream/ Damaged strand		
D1.1	50	5'- <b>pF</b> CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG TA CGT ACA CTT ACT CAT TGC-3'
D1.2	99	5'-CGA GTC ATC TAG CAT CCG TA CAG CAG CAG CAG CAG CAG CAG CAG CAG <b>CAF</b> CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG TA CGT AGA CTT ACT CAT TGC-3'
D1.3	100	5'-CGA GTC ATC TAG CAT CCG TA CAG CAG CAG CAG CAG CAG CAG CAG CAG <b>CAU</b> CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG TA CGT AGA CTT ACT CAT TGC-3'
Template strand		
T1.1	100	5'- <b>B</b> GCA ATG AGT AAG TCT ACG TA CTG TAC GGA TGC TAG ATG ACT CG-3'
T1.2	62	5'-G GCA ATG AGT AAG TCT ACG TA CTG CTG CTG CTG CTG CTG CTG TAC GGA TGC TAG ATG ACT CG-3'
Upstream strand		
U1	49	5'-CGA CTC ATC TAG CAT CCG TA CAG CAG CAG CAG CAG CAG CAG CAG CAG CA- 3
PCR Primer		
P1.1	20	5'-CGA GTC ATC TAG CAT CCG TA-3'
P1.2	20	5'- <b>6</b> CAA TGA GTA AGT CTA CGT A-3'
P1.3	23	5'-GCT CAG GTT CTG CTT TTA CCT GC-3'
P1.4	18	5'-TGC AGG GTT ACC GCC ATC-3'

**B:** Biotin**U:** Uracil**F:** Tetrahydrofuran, THF**p:** Phosphate**6:** 6-Carboxyfluorescein

## **Cell culture**

Human HD induced pluripotent stem cells (iPSCs)(GM23225) were purchased from Coriell Institute for Medical Research (Camden, NJ) and cultured in Ham's F12 Medium/Dulbecco Modified Eagles Medium, 1:1 mixture with 2mM L-glutamine, 10 ng/ml basic fibroblast growth factor, and 20% knock-out serum replacement. Cells were grown at 37°C under 5% CO<sub>2</sub>. Cells were differentiated into neuronal cells as described previously [160].

Chromatin immunoprecipitation (ChIP) was performed as previously described [82] with minor modifications. Briefly, cells were treated with 10 mM KBrO<sub>3</sub> for 2 hours and untreated cells were used as a negative control. Cells were cross-linked using 1% formaldehyde at 37°C for 30 min. The crosslinking reaction was stopped by adding glycine to 125 mM. The DNA was sheared by sonication for 15 cycles of 30 s ON and 30 s OFF at 4°C with Bioruptor ultrasonicator (Diagenode, Denville, NJ). The supernatant was diluted by 10-fold. The lysate was divided in equal aliquots, added either IgG or FAN1-ab (ab222206, Abcam Inc., Cambridge, UK), and incubated with protein A/G agarose beads. The ChIP crosslinking reversal was performed using 0.2 M NaCl and incubation at 65°C for 6 h. DNA was subject to proteinase K digestion and phenol/chloroform extraction. The recovered DNA was used for quantitative PCR.

## **Quantitative real-time PCR and data analysis**

The quantitative PCR reaction was performed by using SYBR Green Supermix kit (Bio-Rad Laboratories). The DNA samples from the ChIP were

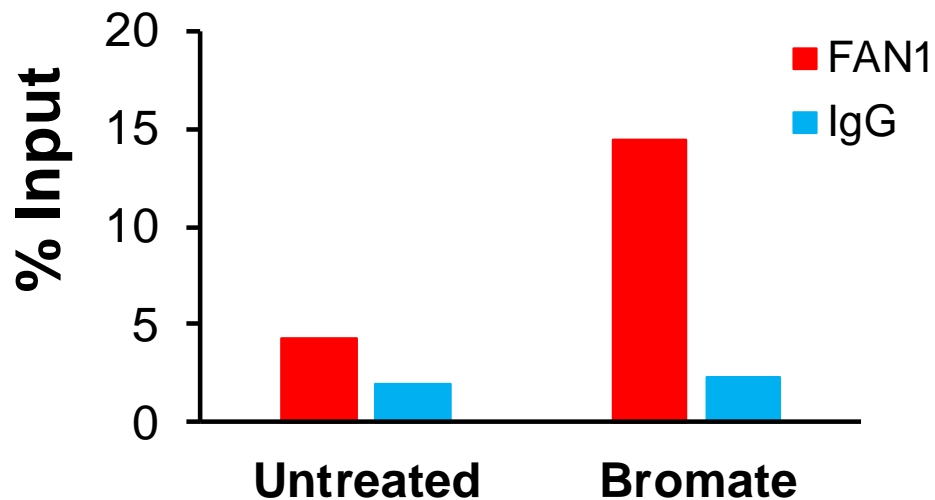
amplified using a CFX Connect Real-Time PCR Detection System (Bio-Rad Laboratories). The *HTT* gene region containing CAG repeats were amplified by using P1.3 and P1.4 primers (Table 1). The amplification was carried out by the following PCR procedure: Denature for 98°C for 2 min, denaturation at 98°C for 20s, annealing at 51°C for 1 min, extension at 72°C for 2 min, in a total of 40 cycles. The data analysis was done as previously described [82].

## RESULTS

### **FAN1 is recruited to CAG repeats upon oxidative DNA damage**

FAN1 interacts with several DNA repair enzymes including MSH2/MSH3, PMS1, PMS2, PCNA, and FANCD1 complex [154-156, 161]. These interactions suggest that FAN1 is involved in different DNA repair pathways removing DNA damage and maintaining genome stability. Previous results from Smorgorzewska et al. 2010 have shown that FAN1 knockdown in human bone osteosarcoma (U2OS) cells partially increased sensitivity to MMS and CPT, suggesting that FAN1 is also involved in the repair of ssDNA breaks by BER [155]. To further explore whether FAN1 can participate in the repair of oxidative DNA damage in CAG repeats, we initially determined if the enzyme can be recruited to the repeats in HD neurons differentiated from the iPSCs from an HD patient using ChIP assay. Differentiated HD neurons were treated with 10 mM potassium bromate (KBrO<sub>3</sub>) for 2 h. Untreated neurons were used as control. The results showed that similar to the recruitment of DNA BER enzymes to oxidative DNA damage induced by KBrO<sub>3</sub> in CAG repeats in HD patient cells [82], FAN1 was recruited to the expanded

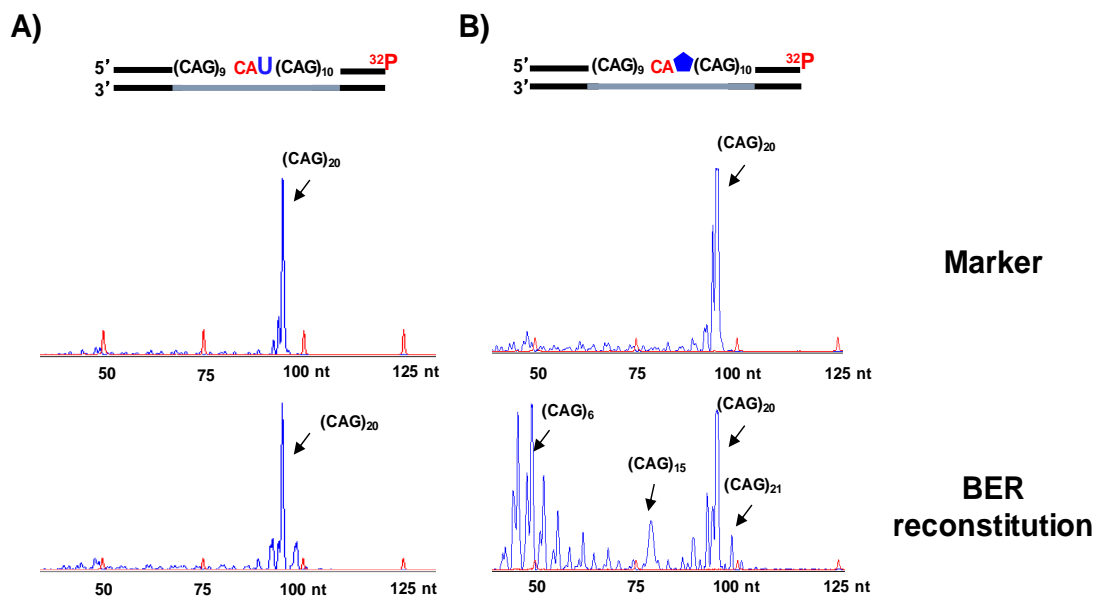
CAG repeats in HD neurons upon the treatment with 10 mM KBrO<sub>3</sub> (Fig. 1.1). Only a low level of recruitment of FAN1 was detected in untreated neurons (Fig. 1.1). IgG gave a low basal level of the signal from nonspecific binding (Fig. 1.1). Since KBrO<sub>3</sub> specifically induces 8-oxoGs, the results indicate that FAN1 was recruited to 8-oxoGs in the CAG repeats in HD neurons.



**Figure 1.1 FAN1 is recruited to CAG repeats in HD differentiate neurons after bromate treatment.** HD neurons were treated with 10 mM KBrO<sub>3</sub> (right) for 2 h. Untreated neurons were used as control (left). An antibody against FAN1 was used to pull down the DNA fragments containing CAG repeats that were crosslinked with FAN1 in KBrO<sub>3</sub>-treated and untreated HD neurons (red). IgG was used as control (blue) for detecting the signal from non-specific protein binding of the antibody. (The results were provided by Dr. Yanhao Lai at Dr. Liu's Laboratory)

## **FAN1 promotes CAG repeat deletion through BER**

Previous studies by our group and others have shown that BER is actively involved in promoting TNR repeat deletion through the removal of more TNRs by FEN1 than those synthesized by pol  $\beta$  [4, 48, 82, 89, 90, 92, 95]. Previous studies have shown that the secondary structures generated by TNRs can inhibit FEN1 cleavage, leading to repeat expansion [90]. Since FAN1 has an efficient 5'-flap endonuclease activity and 5'-3' exonuclease activity [152, 154-158], it is possible that FAN1 endo/exonuclease may remove the secondary structures formed in the CAG repeat tract. To test this hypothesis, we examined whether FAN1 can remove CAG repeats to modulate CAG repeats instability by coordinating with FEN1 cleavage of the repeats through BER. We initially determined if FAN1 can lead to CAG repeat instability through BER. We found that BER reconstituted with the presence of FAN1 in the (CAG)<sub>20</sub> repeats containing uracil or tetrahydrofuran (THF), an analog of modified abasic site or sugar that is subject to the LP-BER pathway. The results showed that BER of the uracil in the CAG repeats in the presence of FAN1 predominantly led to the repaired products containing the full-length of the (CAG)<sub>20</sub> repeats with a small amount of the (CAG)<sub>19</sub> and (CAG)<sub>20</sub> products (Fig. 1.2A). In contrast, BER of the THF in the (CAG)<sub>20</sub> repeat tract with FAN1 predominantly led to large amount of deletion products (Fig. 1.2B). Additional analysis of the size of the repaired products showed that FAN1 led to the products with full-length and repeat deletion containing (CAG)<sub>5-7</sub> and (CAG)<sub>15</sub> and a small amount of the (CAG)<sub>21</sub> expansion product. These results suggest that FAN1 predominantly promoted CAG repeat deletion through the long-patch BER.

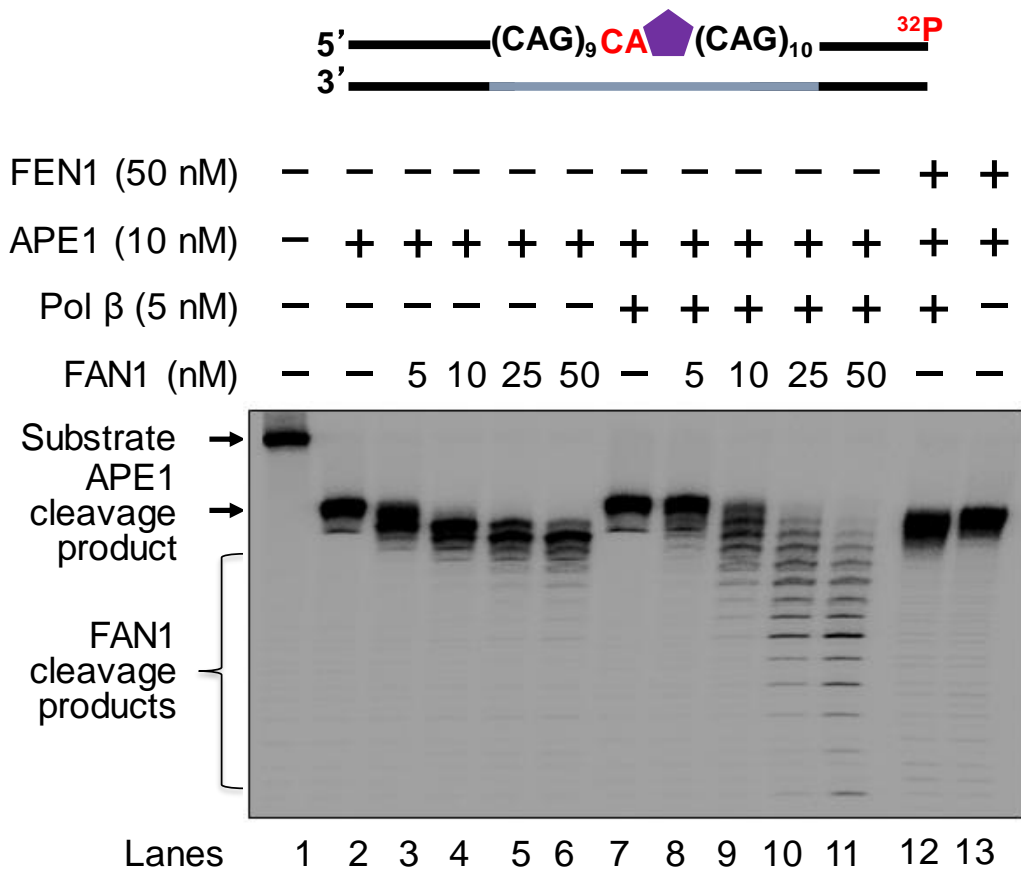


**Figure 1.2 FAN1 promotes CAG repeat deletion during long-patch BER.** To test whether FAN1 can modulate CAG repeat instability we reconstituted base excision repair (BER) in CAG repeats duplex containing a damage in the middle of the (CAG)<sub>20</sub> repeats. The repaired products were isolated, and PCR amplified. The PCR products were analyzed by capillary electrophoresis and DNA fragment analysis. (A) BER of a uracil in the (CAG)<sub>20</sub> repeat substrate containing an uracil in the middle of the repeats. The repaired products show the full length of the CAG repeats containing a small amount of (CAG)<sub>19</sub> deletion product and (CAG)<sub>21</sub> expansion product. (B) BER of THF in (CAG)<sub>20</sub> in the middle of the repeats. The repaired fragment products show large CAG repeat deletions ranging from (CAG)<sub>5</sub> to (CAG)<sub>19</sub> repeats and small deletion and expansion products (CAG)<sub>19</sub> and (CAG)<sub>21</sub>. All experiments were performed in triplicate.

## **FAN1 efficiently cleaves CAG repeats during BER**

Trinucleotide repeat instability is mediated by the formation of non-B form secondary structures such as loops, hairpins, and quadruplex structures during DNA replication [162-165] and repair [4, 90]. During BER, FEN1 flap cleavage of TNRs balances with the synthesis of the repeats by pol  $\beta$  can determine whether the repeats are deleted or expanded [4, 90]. To further examine whether the deletion products result from FAN1 cleavage of CAG repeats during BER, we examined FAN1 cleavage activity during BER of an abasic site (THF) in the middle of (CAG)<sub>20</sub> tract. Our results showed that increasing concentrations of FAN1 efficiently cleaved 1-3 CAG repeats in the absence of pol  $\beta$  (Figure 1.3, lanes 3-6) compared with the FEN1 cleavage that only removed 1 CAG repeat (Figure 1.3, lane 13). FAN1 cleavage of the repeats was significantly increased in the presence of 5 nM pol  $\beta$  (Figure 1.3, lane 8-11) indicating that pol  $\beta$  performed DNA strand displacement synthesis and created a repeat flap that was efficiently cleaved by FAN1 activity. However, FEN1 cleavage was only slightly stimulated by pol  $\beta$  (Figure 1.3, lane 12). These results suggest that during BER, CAG repeats on the downstream strand formed the secondary structures that inhibited FEN1 flap cleavage, but not FAN1 cleavage activity.

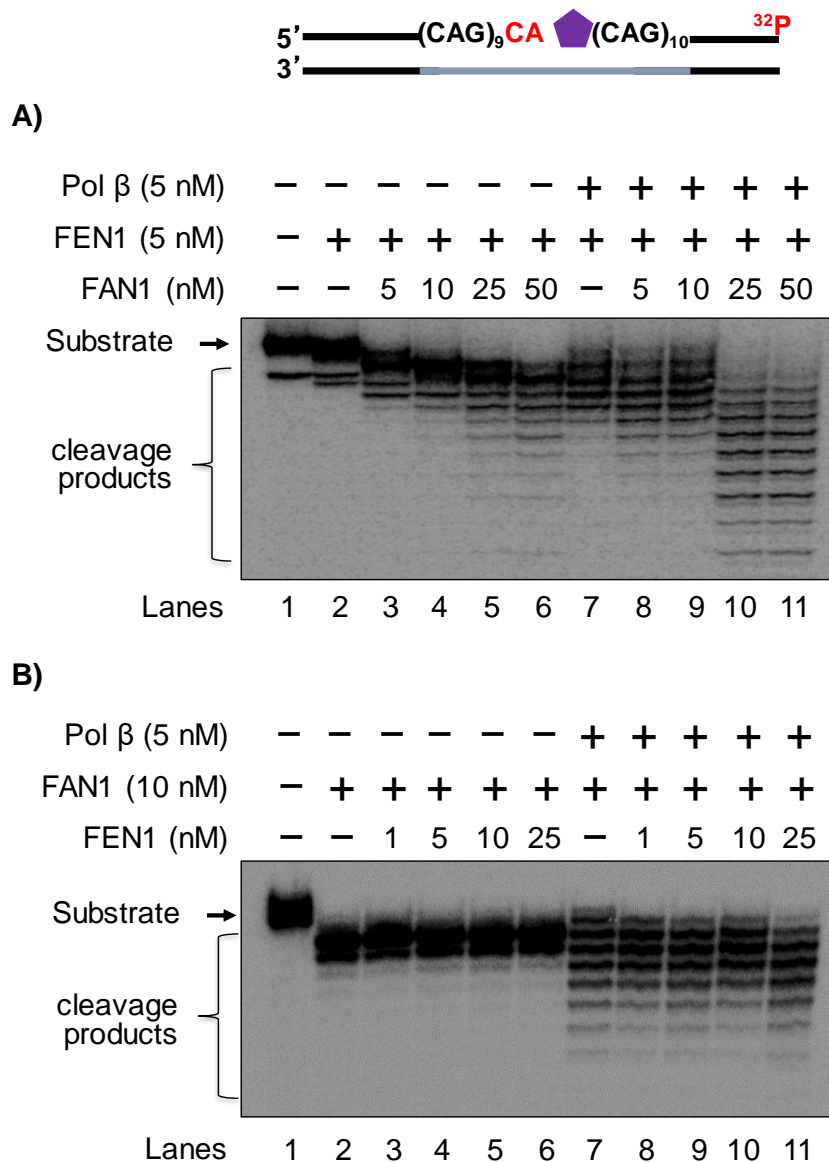




**Figure 1.3 FAN1 cleavage of CAG repeats during BER.** To test FAN1 activity on CAG repeats during BER, we performed FAN1 DNA cleavage assay on DNA substrates that contained a (CAG)<sub>20</sub> with a THF in the middle of the repeats. The substrate (lane 1) was pre-nicked by APE1 (lane 2). Lanes 3-6 show FAN1 cleavage at increasing concentrations in the absence of pol β. Lane 7 shows pol β and substrate. Lanes 8-12 show FAN1 cleavage at increasing concentrations in the presence of pol β. Lane 12 shows FEN1 cleavage in the presence of pol β and lane 13 shows FEN1 cleavage in the absence of pol β). All these experiments were performed in triplicates.

### **FAN1 coordinates with FEN1 to cleave CAG repeats during BER.**

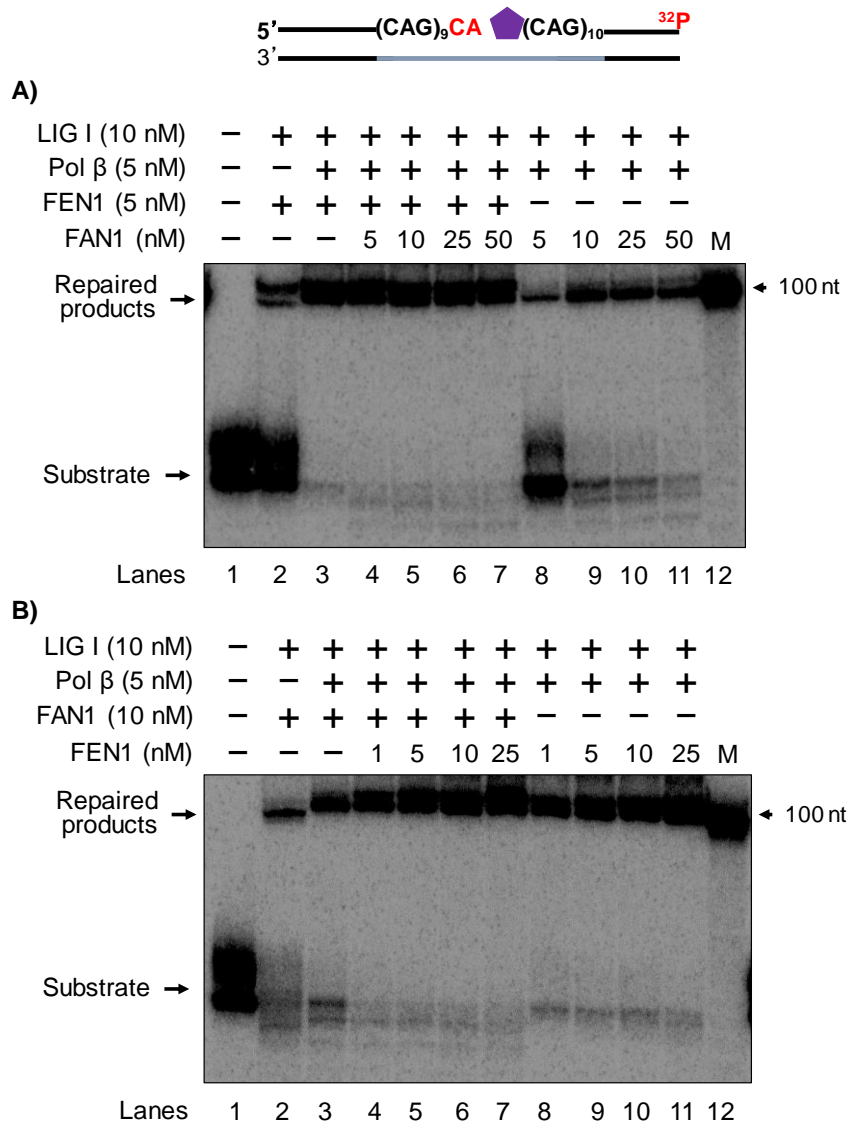
To explore whether FAN1 can coordinate with FEN1 to remove CAG repeats promoting repeat deletion during BER, we initially examined the FEN1 cleavage of CAG repeats during BER of an abasic site in the duplex (CAG)<sub>20</sub> substrate in the presence and absence of pol  $\beta$  with increasing concentrations of FAN1 (Fig. 1.4A). The results showed FEN1 at 5 nM alone exhibited a weak cleavage activity on CAG repeats on the (CAG)<sub>20</sub> substrate in the absence of pol  $\beta$  and FAN1 (Fig. 1.4A, lane 2). Increasing concentrations of FAN1 resulted in the cleavage of multiple numbers of CAG repeats (Fig. 1.4A, lanes 3-6). Similar cleavage activity of FEN1 on the CAG repeats in the presence of FAN1 and pol  $\beta$  (5 nM) was detected (Fig. 1.4A, lanes 8-11) compared to that with the FEN1 alone in the presence of 5 nM pol  $\beta$  (Fig. 1.4A, lane 7). In the presence of pol  $\beta$ , high concentrations of FAN1 at 25 nM-50 nM removed all the (CAG)<sub>10</sub> repeats at the downstream strand (Fig. 1.4A lanes 10-11). On the other hand, FAN1 alone showed efficient cleavage on the CAG repeats on the substrate (Fig. 1.4B, lane 2). Increasing concentrations of FEN1 had little effects on the cleavage product of FAN1 (Fig. 1.4B, lanes 3-6). Similar results were shown in the presence of pol  $\beta$  (Fig. 1.4B, lanes 8-11) compared to FAN1 and pol  $\beta$  alone (Fig. 1.4B, lane 7). These results suggest that CAG repeats formed hairpin structures that inhibited FEN1 cleavage of the repeats. The results further indicate that FAN1 employed its 5' flap endonuclease and 5' exonuclease activity to process hairpin structures, thereby disrupting the structures and stimulating the removal of the repeats.



**Figure 1.4 FAN1 and FEN1 coordinate to cleave CAG repeats during BER.** To test whether FAN1 and FEN1 can coordinate to remove CAG repeats during BER, we examined the cleavage of FAN1 and FEN1 in the presences and absence of pol β in a nicked substrate containing a THF in the upstream strand of the CAG repeats. A) FAN1 cFEN1 cleavage during BER. Lane 1 represent substrate only. Lane 2 represents FEN1 cleavage of the CAG substrate. Lanes 3-6 represents FEN1 cleavage in the presence of increasing concentration of FAN1. Lane 7 shows FEN1 cleavage in the presence of pol β. Lane 8-11 shows FEN1 cleavage in the presence of pol β with increasing concentration of FAN1. B) FEN1 complementation of FAN1 cleavage during BER. Lane 1 represent substrate only. Lane 2 represents FAN1 cleavage of the CAG substrate. Lanes 3-6 represents FAN1 cleavage in the presence of increasing concentration of FEN1. Lane 7 shows FAN1 cleavage in the presence of pol β. Lane 8-11 shows FAN1 cleavage in the presence of pol β with increasing concentration of FEN1. All experiments were performed in triplicate.

### **FAN1 and FEN1 coordinate to promote efficient BER on CAG repeats.**

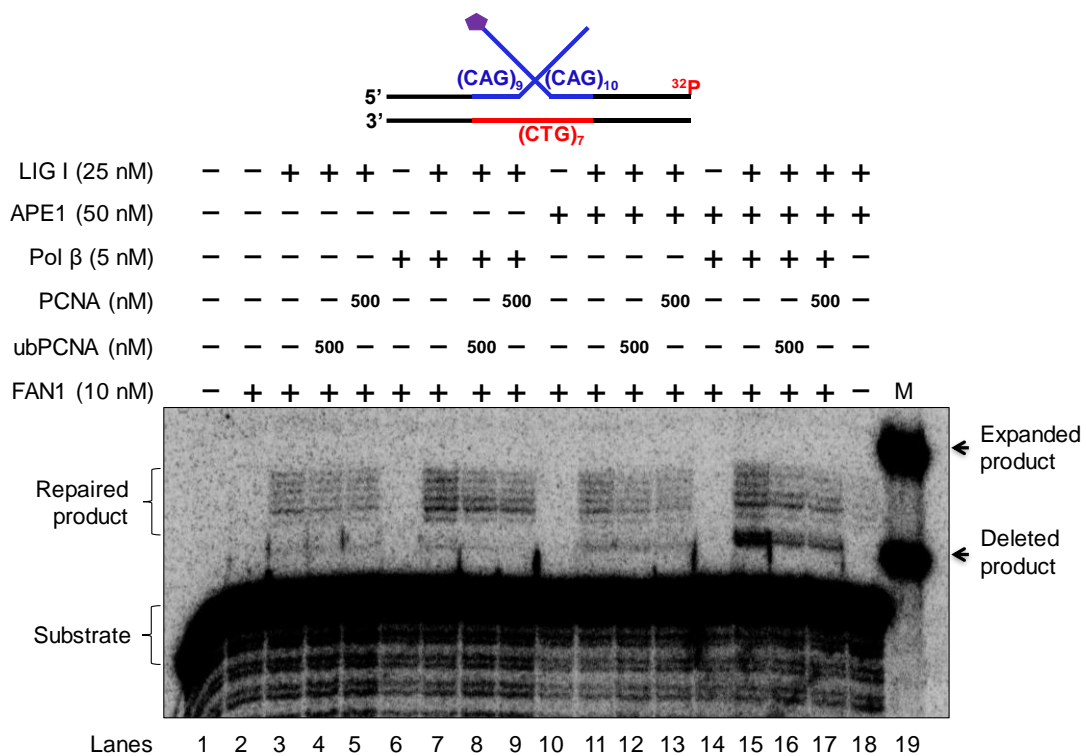
Since FEN1 is the major enzyme involved in the removal of modified bases during long patch BER [165-167] and plays a critical role in preventing TNR expansion during BER, we further determined if FEN1 can cooperate with FAN1 during BER in CAG repeats to facilitate the formation of the repaired products. We reconstituted BER of an abasic site in the (CAG)<sub>20</sub> repeat substrate in the presence and absence of FEN1 or FAN1. The results showed that both FEN1 and FAN1 promoted the production of the repaired products during BER in the (CAG)<sub>20</sub> substrate (Fig. 1.5A and 1.5B). FEN1 efficiently promoted the production of the repaired products independently of increasing concentrations of FAN1 (Fig. 1.5A, lanes 3-7 and 1.5B, lanes 8-11). On the other hand, FAN1 cleavage led to less repaired products in the absence of FEN1 (Fig. 1.5B, lane 3 and Fig. 1.5A, lanes 8-11). The results indicate that FAN1 and FEN1 cooperated in removing CAG repeats, facilitating the production of the repaired products during BER.



**Figure 1.5 FAN1 and FEN1 cooperate to promote efficient BER on CAG repeats.** To test whether FEN1 and FAN1 can coordinate with each other to promote efficient BER on CAG repeats, we performed BER reconstitution in the presence and absence of either FAN1 or FEN1 in a nicked substrate containing a THF in the upstream strand of the CAG repeats. A) BER in the presence and absence of FEN1 with increasing concentration of FAN1. Lane 1 represents substrate only. Lane 2 shows FEN1 and LIGI only. Lanes 3 shows BER in the absence of FAN1. Lanes 4-7 show BER with increasing concentrations of FAN1. Lane 8-11 show BER in the absence of FEN1 with increasing concentration of FAN1. B) BER in the presence and absence of FAN1 with increasing concentration of FEN1. Lane 1 represents substrate only. Lane 2 shows FAN1 and LIGI only. Lanes 3 shows BER in the absence of FEN1. Lanes 4-7 show BER with increasing concentrations of FEN1. Lane 8-11 show BER in the absence of FAN1 with increasing concentration of FEN1. Lane 12 for A and B represents the size marker. All experiments were performed in triplicate.

## **FAN1 attenuates CAG repeat expansion during BER in a CAG repeat hairpin.**

Early studies have shown that FAN1 interacts with ubPCNA [161]. However, it is unknown how the ubPCNA-FAN1 interaction can influence FAN1 activity, its coordination with other enzymes, and its cleavage on CAG repeats during the repair. To explore the role of FAN1 in processing hairpin structures during BER, we reconstituted BER with a double-flap CAG repeat substrate that mimics an intermediate generated by the 5'-incision of an abasic site by APE1 in a CAG repeat hairpin in the presence of BER cofactors, PCNA, and ubiquitinated-PCNA (ubPCNA). The results show that FAN1 promoted the formation of the repaired products that are shorter than the expanded product by removing the downstream CAG repeat flaps in the presence or absence of PCNA or ubPCNA (Fig. 1.6, lanes 3-5). In the presence of pol  $\beta$ , the shortened repaired products resulting from FAN1 cleavage were increased (Fig. 1.6, lanes 7-9). The presence of PCNA and ubPCNA reduced the amount of the repaired products and the formation of non-expanded products and deletion products (Fig. 1.6, lanes 7-9, lanes 15-17). The presence of pol  $\beta$  led to an increase in the quantity of repaired products (Fig. 1.6, lane 15). The results indicate that FAN1 removed the CAG repeats in the downstream strand to promoting the formation of nonexpanded and deletion products. Since the removal of a hairpin structure requires the removal of the upstream flap by a nuclease such as APE1 [94], the results suggest coordination between FAN1 and APE1 3'-5' exonuclease activity. Our results also showed that PCNA and ubPCNA prevented the generation of the repair products with longer repeats attenuating repeat expansion during BER in a CAG hairpin.



**Figure 1.6 FAN1 promotes the removal of CAG repeats hairpin intermediate during BER.** The effect of FAN1 on the removal on CAG repeat hairpin containing an abasic site during BER was examine by reconstituting BER in a substrate containing a hairpin intermediate with an abasic site located in the upstream strand. Lane 1 is the substrate only. Lane 2 corresponds to FAN1. Lanes 3-5 correspond to FAN1 and LIG I in the presence and absence of either PCNA or ubPCNA. Lane 6 corresponds to FAN1 an pol β. Lanes 7-9 corresponds to FAN1, LIG I, and pol β in the presence and absence of PCNA and ubPCNA. Lane 10 represents APE1 and FAN1. Lane 11-13 corresponds to APE1, FAN1, LIG I in the presence and absence of PCNA and ubPCNA. Lane 14 presents FAN1, APE1, and pol β. Lanes 15-18 corresponds to FAN1, APE1, pol β, and LIG I in the presence and absence of PCNA and ubPCNA. Lane 19 corresponds to the size marker top band represents full hairpin size and lower band represents the removal of the hairpin/double flap intermediates. All experiments were performed in triplicates.

## DISCUSSIONS

In the present study, we explored the role of FAN1 in modulating CAG repeat instability during BER. We further characterized FAN1 cleavage of CAG repeats during BER of an abasic site. Moreover, we determined the coordination between FAN1 and FEN1 cleavage in modulating CAG repeat instability during BER. We showed that FAN1 was recruited to CAG repeats upon oxidative DNA base damage induced by  $\text{KBrO}_3$  (Fig. 1.1). Also, we demonstrated that FAN1 promoted CAG repeat deletion during BER in duplex DNA containing CAG repeats during long-patch BER (Fig. 1.2B). Additional analysis on FAN1 cleavage on CAG repeats duplex showed that FAN1 efficiently cleaved CAG cleavage using its 5'-exonuclease cleavage activity and that pol  $\beta$  stimulated FAN1 5' flap cleavage during BER (Figure 1.3). We further demonstrated that FAN1 cleaved a long CAG repeat flap independently of FEN1 cleavage (Fig. 1.4). However, we found that FEN1 was necessary for generating the repaired products in the presence of FAN1 (Fig 1.5). Lastly, we showed that FAN1 attenuated CAG repeat expansion during the repair of a hairpin intermediate structure through BER (Fig. 1.6). Our results supported a model during which reactive oxygen species (ROS) induces DNA base lesions such as 8-oxoGs on the non-template strand of a duplex CAG repeat tract or hairpin. The 8-oxoG DNA glycosylase 1 (OGG1) removes the 8-oxoG in the duplex or the hairpin loop, leaving an abasic site on the non- template strand. APE1 cleaves 5' site of the abasic site to generate ssDNA break or a nick on the duplex DNA (Fig. 1.7A) or a hairpin loop (Fig. 1.7B). During this process, a 5'-flap is generated by pol  $\beta$  DNA synthesis during BER in a duplex CAG repeats, and

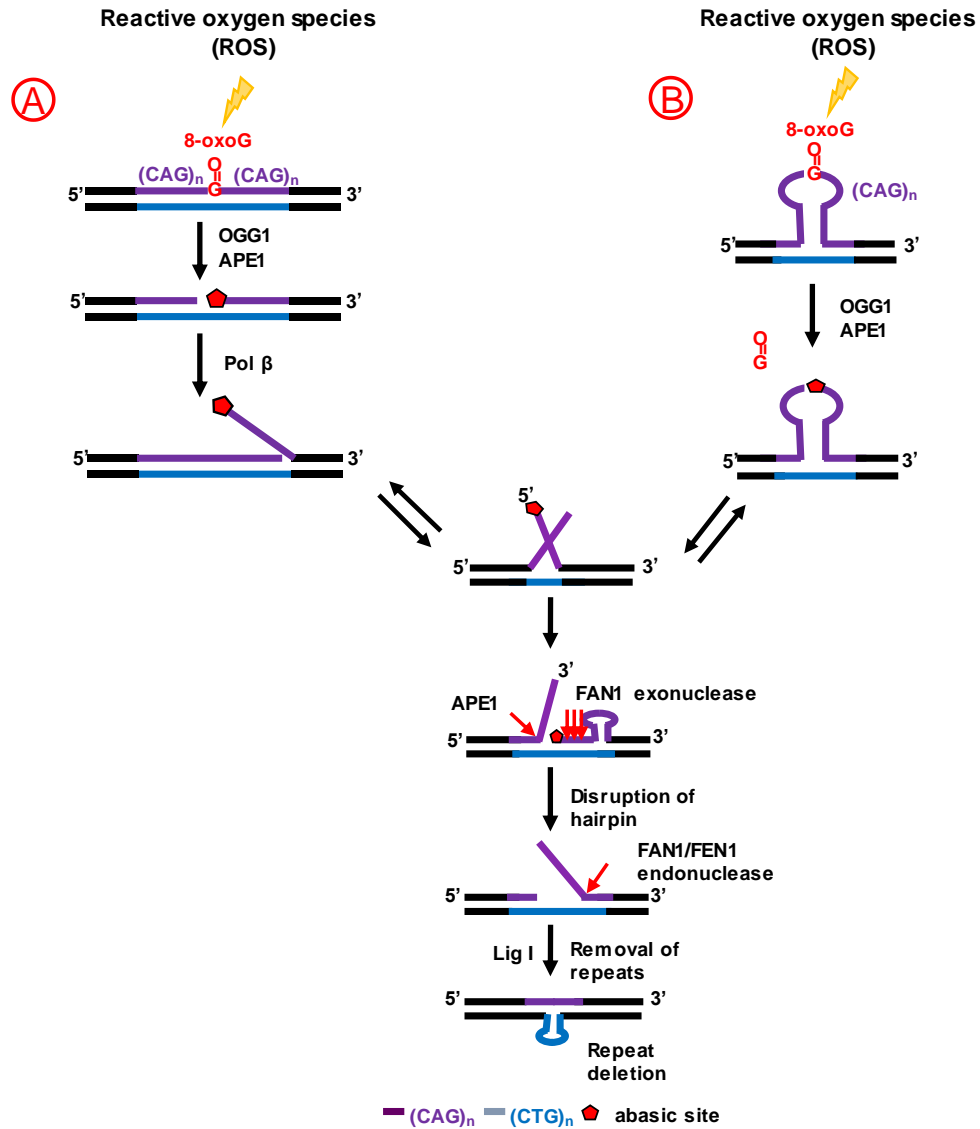


double flap intermediate is generated from the incision of CAG repeat hairpin. The CAG flaps can form hairpin structures that inhibit FEN1 flap cleavage. The hairpin formation allows FAN1 to use its 5' exonuclease activity to process the end of the hairpin and destabilize the hairpin converting it into a 5'-flap. Subsequently, FAN1 uses its 5'-endonuclease activity to remove the repeat flap leading to the formation of deletion products and prevention or attenuation of repeat expansion (Fig. 1.7).

FAN1 is a novel protein independently discovered by four different groups to be involved as an FA repair pathway because of its involvement in resolving DNA interstrand crosslinks [152, 154-156]. However, up to date, no FAN1 mutations have been found in FA patients. Interestingly, FAN1 mutations have been associated with karyomegalic interstitial nephritis (KIN) [168], colorectal cancer [169], aging [170], and schizophrenia and autism [171] suggesting that FAN1 has other cellular functions besides its function in multiple repair pathways. Although FAN1 has been initially identified to be involved in ICL resistance, however, it has been proposed that the enzyme may be involved in homologous recombination through its 5' to 3' exonuclease activity [157]. The similarity of FAN1 activity and substrate specificity with other nucleases such as FEN1, EXO1, SLX1 suggests that FAN1 may act as a supplementary nuclease because of its roles in a repair pathway that has not been identified. Moreover, the multicellular functions of FAN1 remain to be elucidated. Recent findings on the association of FAN1 single nucleotide polymorphism with HD age of onset suggests that FAN 1 plays an essential role in modulating the progression of TNR instability and polyglutamine diseases [150]. Here, we provided the first evidence that FAN1 can

be recruited to *HHT* gene upon oxidative DNA damage (Fig. 1.1), suggesting that FAN1 can participate in BER to modulate the stability of expanded CAG repeats. We also showed that FAN1 promoted repeat deletion through the coordination with BER enzymes during long-patch BER (Fig. 1.2B). However, FAN1 only slightly affect the repeat instability during single-nucleotide BER. It is possible that the native sugar was efficiently removed by pol  $\beta$ 's dRP lyase domain [82], which acted faster than FAN1 cleavage during the repair (Fig. 1.2A).

Our studies also indicate that FAN1 can coordinate with FEN1 in promoting CAG repeat deletion during BER in CAG repeats. FAN1 can interact with ubiquitinated PCNA. However, it does not interact with PCNA [161]. Conversely, FEN1 has been shown to interact and is stimulated by PCNA [172], but inhibited by ubPCNA [173]. The coordination among FAN1, FEN1, PCNA, and ubPCNA in modulating TNR instability during BER needs to be elucidated. Moreover, TNRs readily form stable secondary structures that can modify the activities of FAN1 and FEN1 [174]. It is of importance that further studies are needed to identify the roles of FAN1 and FEN1 and their coordination with BER cofactors and enzymes of other DNA repair pathways interactions in preventing TNR expansion through DNA damage and repair.



**Figure 1.7 A hypothetical model of FAN1 promoting or attenuating CAG repeat deletion through BER.** Reactive oxygen species (ROS) can cause damage on CAG duplex or expanded CAG hairpin structures creating an 8-oxoguanine (8-oxoG). 8-oxoG is recognized by 8-oxoguanine DNA glycosylase (OGG1) leaving an abasic site that is incised by APE1 at the 5'-end. This cleavage will leave a nick on the duplex (A) or a double flap intermediate on a hairpin (B). On a duplex pol β will strand displace that damage strand creating a flap. This will result in a double flap equilibration on CAG repeats and the formation of a hairpin on the downstream strand which will block FEN1 cleavage. The resolution of this secondary structure will require FAN1 5' exonuclease activity to disrupt the hairpin structure and FEN1 to remove the flap. APE1 could remove the upstream strand facilitating the formation of a hairpin on the template strand. CAG repeat deletion occurs when the strand is ligated in the presence of a secondary structure on the template strand and attenuation of expansion occurs when the double flap hairpin is removed leading to unexpanded repaired products.

## **CHAPTER 2: R-LOOPS MEDIATE TRINUCLEOTIDE REPEAT INSTABILITY VIA MODULATION OF COORDINATED DNA BASE EXCISION REPAIR ENZYMATIC ACTIVITIES**

### **ABSTRACT**

Trinucleotide repeat (TNR) expansion is the cause of over 40 neurodegenerative diseases. Somatic TNR expansion and deletion can be modulated by DNA damage and repair and gene transcription. Recent studies have pointed towards a role for R-loops formed during gene transcription in causing TNR instability, and the base excision repair (BER) pathway mediates CAG repeat deletions induced by R-loops in yeast. However, it remains unknown how BER in R-loops can modulate TNR expansion and deletions. In this study, we explored the mechanisms by which BER modulates TNR instability through R-loops. We found that APE1 incised an abasic site in a TNR R-loop, creating a double-flap intermediate within an RNA:DNA hybrid that subsequently inhibited pol  $\beta$  synthesis of TNRs. The R-loop intermediate further stimulated FEN1 cleavage on the non-template strand 5' flap. Surprisingly, we found that the RNA strand in TNR-R-loops was also partially cleaved by FEN1. The partially cleaved annealed RNA allowed pol  $\beta$  to skip over a TNR loop on the template strand to perform DNA synthesis. Consequently, this resulted in the efficient removal of repeats by FEN1 and inadequate synthesis of TNRs by pol  $\beta$ , leading to repeat deletion. Our results indicate that BER in a TNR R-loop preferentially causes repeat deletion through the disruption of the balance between the addition and removal of TNRs.

## INTRODUCTION

Trinucleotide repeats (TNRs) expansions are responsible for over 40 human neurodegenerative and neuromuscular diseases including Huntington's disease (HD) (CAG/CTG), myotonic dystrophy type 1 (MD1) (CTG/CAG), fragile X syndrome (FXS) (CGG/CCG), and Friedreich's ataxia (FRDA) (GAA/TTC) among others [86, 88, 175]. The molecular basis underlying TNRs instability is the formation of non-B form or secondary DNA structures, including hairpins, loops, triplexes, and G-quadruplex during DNA replication [136], repair [4], recombination [137], and gene transcription [138]. The TNR tracts are rich in guanines, thereby forming hotspots for DNA base lesions. The most common form of DNA damage includes alkylated and oxidized DNA bases. It has been proposed that cycles of repeated oxidative DNA damage, can promote somatic TNR expansions through a "toxic oxidation cycle" [88] presumably through multiple rounds of repeat expansion resulting from the repair of base lesions by DNA base excision repair (BER). Thus, somatic TNR expansion may play a crucial role in allowing TNR length to reach the threshold that leads to the evident neurodegenerative symptoms of TNR diseases, thereby governing the age at onset of the diseases. Recent studies from our group and other groups have shown that BER of a variety of base lesions plays an active role in modulating TNR instability by inducing large repeat deletions and small expansions [82, 91-95]. Furthermore, we demonstrated that the location of a DNA base lesion in a TNR tract determines the outcome of the repeat deletion or expansion [92].

Gene transcription at expanded repeated sequences can lead to the formation of R-loops [5, 176]. R-loops are generated when a nascent RNA strand hybridizes back to its DNA template to create an RNA:DNA hybrid [5, 176]. The RNA:DNA hybrid within an R-loop is usually longer than the transient RNA:DNA hybrid generated by RNA priming during DNA replication. R-loops are detected in bacteria, yeast, and humans and are implicated to be involved in many cellular processes [5, 176]. Some physiological roles of R-loops are roles in DNA replication initiation at mitochondrial and prokaryotic origins of replication, class switch recombination at immunoglobulin genes, modulation of DNA methylation at CpGs, and transcription termination. Thus, the formation of R-loops is essential for cellular function. However, R-loop accumulation can also result in DNA damage and genome instability [5, 176]. Defective RNA processing can result in the accumulation of R-loops leading to activation of the DNA damage response [126, 177]. It is also suggested that R-loops can act as a mutagenic intermediate disrupting genome integrity.

R-loops can accumulate on expanded TNRs [120, 121, 175, 178-180], and their presence may be further promoted by the deficiency of the senataxin helicase that disrupts the RNA:DNA hybrid [181] or depletion of RNase H enzymes that cleave the RNA in the hybrid [126]. Moreover, R-loops on the expanded GAA, CAG, CTG, and CGG result in a guanine-rich single-stranded region on the non-template strand, which can be damaged by endogenous and exogenous DNA base damaging agents. A previous study from Su and Freudenreich (2017) has shown that the yeast cytosine deaminase, Fcy1, causes R-loop associated CAG

repeat contractions, and these were dependent on BER [120], indicating that BER is responsible for TNR contractions in R-loops. However, the molecular mechanisms by which BER promotes TNR deletion remains unknown.

We hypothesized that BER of DNA base damage in the non-template strand of a TNR R-loop results in repeat deletion by promoting the removal of TNRs but preventing the addition of the repeats. To test the hypothesis, we examined the activities of BER enzymes in TNR R-loops during BER and their impact on the stability of the repeats. We found that APE1 incised an abasic site in (GAA)<sub>20</sub> and (CAG)<sub>20</sub> repeat R-loops creating a double-flap intermediate that inhibited pol  $\beta$  DNA synthesis activity significantly. In contrast, FEN1 cleavage of GAA and CAG repeats during BER in TNR R-loops was significantly stimulated. Surprisingly, we found that the RNA strand in TNR R-loops was also partially processed by FEN1, promoting repeat deletion. The results indicate that BER of a DNA base lesion in R-loops promotes TNR deletion by altering the balance between the addition and removal of TNRs through the modulation of the activities of BER core enzymes and processing of the RNA strand.

## **MATERIALS AND METHODS**

### **Materials**

The DNA and RNA oligonucleotides were synthesized by Integrated DNA Technologies (IDT, Coralville, IA, USA). The deoxynucleotides 5'-triphosphates (dNTPs) were purchased from Sigma-Aldrich (St. Louis, MO). The radionucleotides  $\alpha$ -<sup>32</sup>P-Cordycepin 5-triphosphate (5000 Ci/mmol) and  $\gamma$ -<sup>32</sup>P-ATP

(3000 Ci/mmol) and were purchased from Perkin Elmer Inc. (Boston, MA, USA). Micro Bio-Spin 6 chromatography columns were purchased from Bio-Rad Laboratories (Hercules, CA, USA). Terminal deoxynucleotidyl transferase and T4 polynucleotide kinase were purchased from Thermo Fisher Scientific (Waltham, MA, USA). Diethyl pyrocarbonate (DEPC) was purchased from MP Biomedicals (Santa Ana, CA, USA). All chemical reagents were purchased from Thermo Fisher Scientific (Waltham, MA, USA) and Sigma-Aldrich (St. Louis, MO, USA).

### **Purification of BER enzymes**

Recombinant human AP endonuclease 1 (APE1), flap endonuclease 1 (FEN1), and DNA ligase I (LIG I) were expressed and purified as described previously [94, 95]. Human recombinant pol  $\beta$ -(His)<sub>6</sub>-tag was expressed and purified as previously reported [182] with minor modifications. Briefly, pol  $\beta$  expression in two liters of cell culture was induced at OD of 0.6 and 37 °C with 1mM isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) (VWR International, Radnor, PA) for 3.5 hours. Cells were pelleted and resuspended with Buffer 1 containing 30 mM HEPES, pH 8.0, 500 mM NaCl, 30 mM imidazole, 1 mM EDTA, 1 mM DTT, 1 mM PMSF, and one protease inhibitor cocktail tablet (Roche, St. Louis, MO). Cells were lysed using a French Press (GlenMills, Clifton, NJ). The soluble fraction of the cell lysates, the supernatant was collected by centrifugation at 12,000 rpm for 30 min. The supernatant was loaded into a Ni-NTA agarose column (Qiagen, Hilden, Germany), followed by the wash with 5 column volume (5 CV) by Buffer 2 containing 30 mM HEPES, pH 8.0, 1 M NaCl, 30 mM imidazole, 1 mM EDTA, 1



mM DTT, 1 mM PMSF, and one protease inhibitor cocktail tablet (Roche, St. Louis, MO) and then by 5 CV of Buffer 1. Pol  $\beta$  was eluted by imidazole gradient from 30 mM to 600 mM with Buffer 3 containing 30 mM HEPES, pH 8.0, 500 mM NaCl, 600 mM imidazole, 1 mM EDTA, 1 mM DTT, and 1 mM PMSF. The peak fractions were combined and dialyzed into 30 mM HEPES, pH 7.5, 1.7 M  $(\text{NH}_4)_2\text{SO}_4$ , 0.5% inositol, and 1 mM PMSF. The dialyzed samples were loaded into a phenyl sepharose 6 fast flow column (GE Healthcare Bio-Science, Uppsala, Sweden) and eluted with a 30 mM  $(\text{NH}_4)_2\text{SO}_4$  reverse gradient. The fractions were tested for enzymatic activity and contamination of E. Coli DNA polymerases and nucleases. The fractions with a high level of enzymatic activity and low level of polymerase and nuclease contamination were combined and dialyzed into the buffer containing 30 mM HEPES, pH 7.5, 30 mM KCl, 1 mM EDTA, 1 mM DTT, and 1 mM PMSF. The dialyzed proteins were loaded on a Mono Q column (GE Healthcare Bio-science, Uppsala, Sweden) and eluted by a 1 M KCl gradient. The peak protein fractions with the high level of enzymatic activity were combined and dialyzed into the storage buffer containing 30 mM HEPES, pH 7.5, 50 mM KCl, 20% glycerol, and 1 mM PMSF. The fractions were aliquoted and stored at -80 °C.

### **Construction of R-loop oligonucleotide substrates**

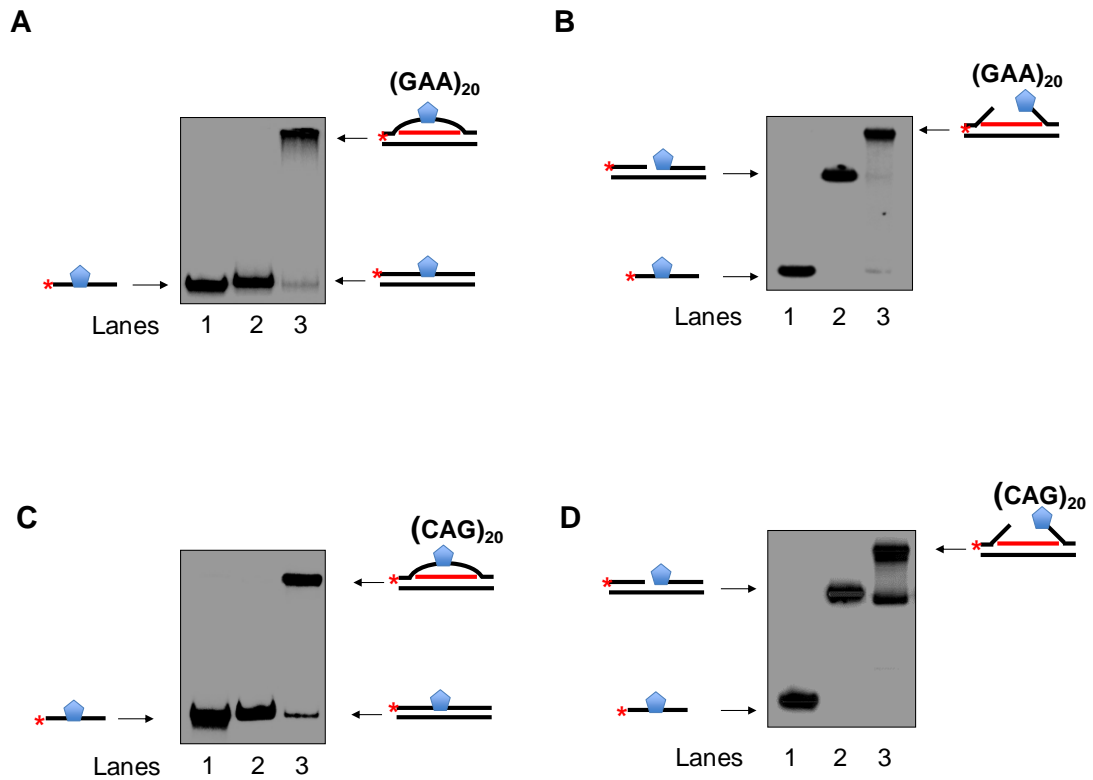
The DNA oligonucleotide substrates containing a tetrahydrofuran (THF) residue, an analog of an abasic site, were designed to mimic a scenario where an abasic lesion occurs in the middle of a  $(\text{GAA})_{20}$  and  $(\text{CAG})_{20}$  repeat tract. The guanine at the tenth unit of  $(\text{GAA})_{20}$  or  $(\text{CAG})_{20}$  repeats was substituted with the

THF residue. The sequences of the oligonucleotides are listed in Table 2. The (GAA)<sub>20</sub> and (CAG)<sub>20</sub> duplex substrates were constructed by annealing the THF containing oligonucleotide to its template strand at a molar ratio of 1:3. The (GAA)<sub>20</sub> and (CAG)<sub>20</sub> R-loop substrates were constructed by annealing the damage-containing strand, the template strand, the RNA strand containing (rGAA)<sub>20</sub> or (rCAG)<sub>20</sub> at a molar ratio of 1:3:15. The (GAA)<sub>20</sub> and (CAG)<sub>20</sub> substrates containing a nick were constructed by annealing the upstream and downstream primer with the template strand at a molar ratio of 1:3:5. The double-flap (GAA)<sub>20</sub> and (CAG)<sub>20</sub> substrates with an RNA:DNA hybrid that mimics nicked-R-loop intermediates were constructed by annealing the upstream, downstream, the RNA strands with (rGAA)<sub>20</sub> or (rCAG)<sub>20</sub> at a ratio of 1:3:15:5 in a total of 10 µl annealing reaction. Reagents and buffers used for constructing R-loops were prepared with 0.1% of DEPC treated water. The duplex and nicked-duplex substrates were constructed by denaturation at 95 °C for 10 min and subsequent cooling down to room temperature. The R-loop substrates and nicked-R-loop substrates were constructed by denaturation at 96 °C for 10 min, cooling down to 52°C, then immediately freezing on dry ice with 100% ethanol. The R-loop substrates constructed were verified using 8% native polyacrylamide gel (Fig. 2.1).

**Table 2. Oligonucleotide sequence**

Oligonucleotides	nt	Sequence
Downstream /Damaged strand		
D1	52	5'- <b>p</b> FAA GAA GAA GAA GAA GAA GAA GAA GAA GAA GAA GAA TA CGT AGA CTT ACT CAT TGC-3'
D2	50	5' - <b>p</b> F CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG TA CGT ACA CTT ACT CAT TGC-3'
D3	99	5'-CGA GTC ATC TAG CAT CCG TA GAA GAA GAA GAA GAA GAA GAA <b>F</b> AA GAA GAA GAA GAA GAA GAA GAA GAA GAA GAA TA CGT AGA CTT ACT CAT TGC-3'
D4	99	5'-CGA GTC ATC TAG CAT CCG TA CAG CAG CAG CAG CAG CAG CAG CAG <b>F</b> CAG CAG CAG CAG CAG CAG CAG CAG CAG TA CGT AGA CTT ACT CAT TGC-3'
Template strand		
T1	100	5'- <b>B</b> GCA ATG AGT AAG TCT ACG TA TTC TTC TTC TTC TTC TTC TTC TTC TTC TTC TTC TTC TTC TTC TA CGG ATG CTA GAT GAC TCG-3'
T2	100	5'- <b>B</b> GCA ATG AGT AAG TCT ACG TA CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG TAC GGA TGC TAG ATG ACT CG-3'
Upstream strand		
U1	47	5'-CGA GTC ATC TAG CAT CCG TA GAA GAA GAA GAA GAA GAA GAA GAA-3'
U2	49	5'-CGA CTC ATC TAG CAT CCG TA CAG CAG CAG CAG CAG CAG CAG CAG CA-3
RNA strand		
R1	60	5'-rGrArA rGrArA rGrArA rGrArA rGrArA rGrArA rGrArA rGrArA rGrArA rGrArA rGrArA rGrArA rGrArA rGrArA rGrArA rGrArA-3'
R2	60	5'-rCrArG rCrArG rCrArG rCrArG rCrArG rCrArG rCrArG rCrArG rCrArG rCrArG rCrArG rCrArG rCrArG rCrArG rCrArG rCrArG-3'
PCR Primer		
P1	20	5'-CGA GTC ATC TAG CAT CCG TA-3'
P2	20	5'- <b>6</b> CAA TGA GTA AGT CTA CGT A-3'

**B:** Biotin**F:** Tetrahydrofuran, THF**p:** Phosphate**r:** Ribonucleotide**6:** 6-Carboxyfluorescein



**Figure 2.1 R-loop and R-loop intermediate annealing.** R-loop substrates containing an (rGAA)<sub>20</sub>/(TTC)<sub>20</sub> (A) or (rCAG)<sub>20</sub>/(CTG)<sub>20</sub> (C) and R-loop intermediate containing an (rGAA)<sub>20</sub>/(TTC)<sub>20</sub> (B) or (rCAG)<sub>20</sub>/(CTG)<sub>20</sub> (D). RNA:DNA hybrid were constructed by annealing the <sup>32</sup>P-labeled non-template DNA strand and RNA strand with the template strand at a molar ratio of 1:3:15 for the 3 strands and 1:3:5:15 for the 4 strands. The primers were incubated at 95°C for 10 min and cooled down to 52°C, and immediately frozen on dry ice with 100% ethanol. Substrates were <sup>32</sup>P-labeled at the 5'-end of the strand with an abasic site and detected using native 8% polyacrylamide gel and phosphorimager. Lane 1 represents the single-strand non-template DNA strand. Lane 2 represents the duplex DNA. Lane 3 represents the R-loop or R-loop intermediate substrates.

## **Enzymatic assay and repeat size measurement**

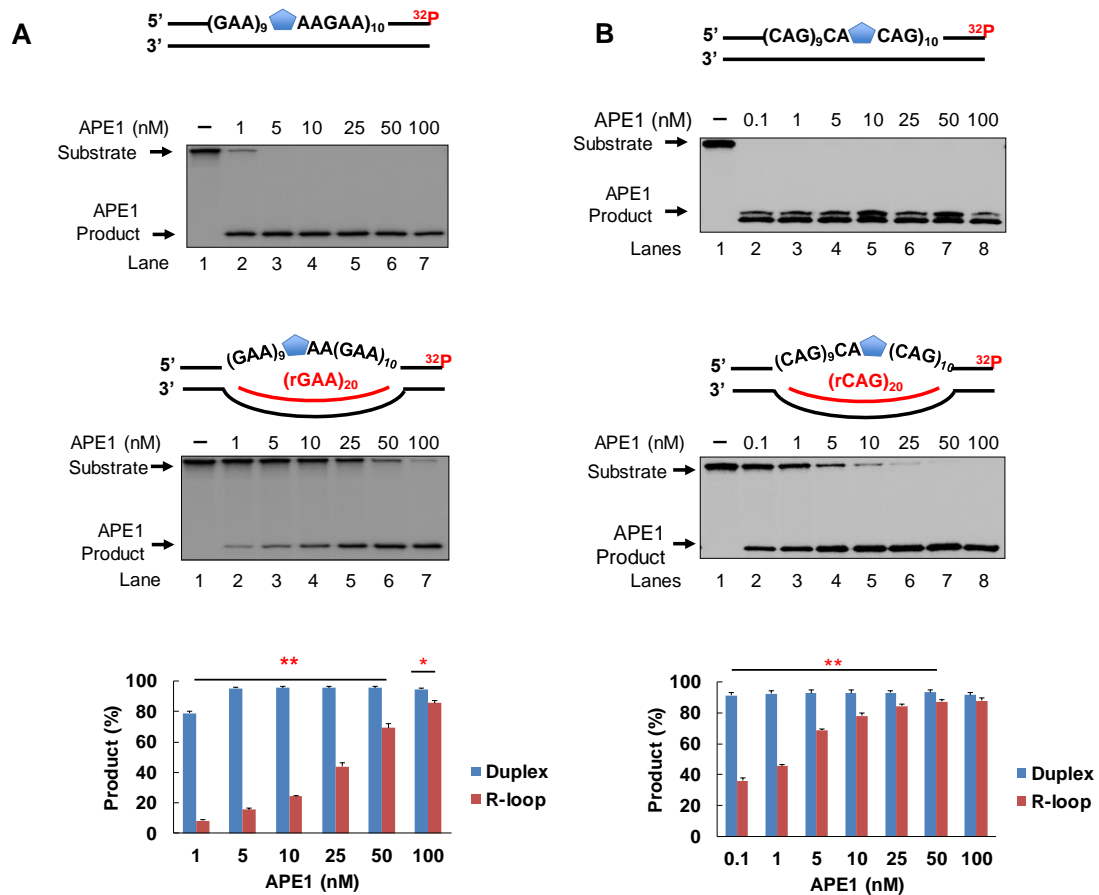
The BER reactions were performed by incubating various types of oligonucleotide substrates with purified APE1, pol  $\beta$ , or FEN1 in BER reaction buffer containing 30 mM HEPES (pH 7.5), 50 mM KCl, 0.1 mg/ml BSA, 0.1 mM EDTA, and 0.01% NP-40. BER reconstitution reactions with TNR-R-loop and TNR-duplex containing an abasic site was performed by incubating 50 nM APE1, 10 nM pol  $\beta$ , 10 nM FEN1, and 20 nM LIG I, with 5 nM (GAA)<sub>20</sub> or (CAG)<sub>20</sub> repeat-containing substrates. All reaction mixtures (20  $\mu$ l) were assembled on ice in BER reaction buffer in the presence of 50  $\mu$ M dNTPs, 5 mM Mg<sup>2+</sup>, 2 mM ATP, and indicated concentrations of BER enzymes and substrates. The reaction mixtures were incubated at 37 °C for 30 min. Subsequently, the reaction mixtures were stopped with a 2x stopping buffer containing 95% deionized formamide and 10 mM EDTA. The repaired products were isolated and amplified by PCR to determine the repeat size changes during BER of the abasic site. The amplified repaired products were separated by capillary electrophoresis. The size of GAA and CAG repeats were determined by DNA fragment analysis according to the method described previously [159].

## **RESULTS**

### **APE1 can incise an abasic site in TNR R-loops during BER.**

To elucidate the molecular mechanism by which TNR R-loops induce repeat instability through BER, we first examined if APE1 can incise the 5'-end of an abasic site (THF) located in the middle of (GAA)<sub>20</sub> and (CAG)<sub>20</sub> R-loops (Fig 2.2).

We found that APE1 at 0.1 nM-5 nM efficiently incised the abasic site located in the duplex (GAA)<sub>20</sub> and (CAG)<sub>20</sub> repeat substrates, though incision of the (GAA)<sub>20</sub> repeat duplex was less efficient than the (CAG)<sub>20</sub> repeat substrate (Figures 1A and 1B, the top and bottom panels). In contrast, APE1 at 1 nM and 5 nM exhibited poor 5'-incision on the abasic site in the (GAA)<sub>20</sub> repeat R-loop (Fig. 2.2A, the panel in the middle, lanes 2-3), though increasing concentrations from 10 nM to 100 nM led to increased incision of the abasic site with all the abasic sites incised at 100 nM (Figure 2.1A, the panel in the middle, lanes 4-7, and the bottom panel). For the abasic site in the (CAG)<sub>20</sub> R-loop, APE1 at 0.1 nM-5 nM generated a significant amount of incision product (Fig. 2.2 B, the panel in the middle, lanes 2-4, and the bottom panel). With increasing concentrations of the enzyme from 10 nM to 100 nM, the abasic site was completely converted to the products (Figure 2.1B, the panel in the middle, lanes 5-8, and the bottom panel). The results showed that the APE1 cleavage activity on the abasic site in the non-template single-strand DNA in the (GAA)<sub>20</sub> and (CAG)<sub>20</sub> R-loops was significantly less efficient compared with its activity in the duplex repeats. However, high concentrations of APE1 at 50 nM and 100 nM incised all the abasic sites in the TNR R-loops. These results are consistent with our previous studies showing that APE1 incises an abasic site in a single-stranded CAG repeat hairpin loop that is attenuated [93].



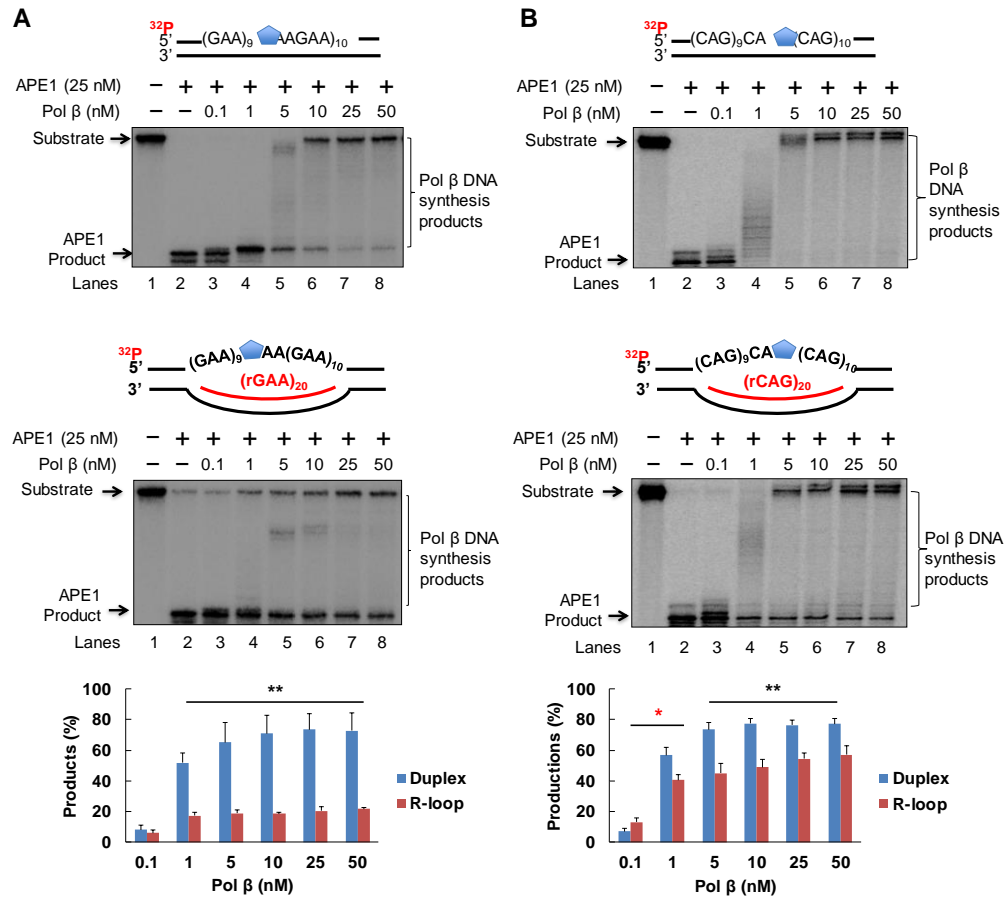
**Figure 2.2 APE1 incision of an abasic site on TNR R-loops.** APE1 incision of an abasic site in (GAA)<sub>20</sub> and (CAG)<sub>20</sub> repeats was performed on either duplex DNA substrates containing (GAA)<sub>20</sub> or (CAG)<sub>20</sub> repeats or R-loop substrates containing (rGAA)<sub>20</sub>/(TTC)<sub>20</sub> or (rCAG)<sub>20</sub>/(CTG)<sub>20</sub> with an abasic site located in the middle of the repeats. Substrates (10 nM) were <sup>32</sup>P-labeled at the 5'-end of the strand containing an abasic site and incubated with various concentrations of APE1 (1 nM-100 nM) at 37°C for 30 min. Substrates were separated from the product in 15% urea-denaturing polyacrylamide gels. Substrates and products were detected by a phosphorimager. Substrates are schematically illustrated above the gels. (A) APE1 5'-incision of an abasic site on the duplex DNA substrate containing (GAA)<sub>20</sub> repeats or the (GAA)<sub>20</sub> repeat R-loop substrate. Lane 1 represents the substrate only. Lanes 2-7 represent APE1 incision activity at the concentrations of 1 nM-100 nM (B) APE1 5'-incision of an abasic site on the duplex DNA substrate containing (CAG)<sub>20</sub> repeats or the (CAG)<sub>20</sub> repeat R-loop substrate. Lane 1 indicates substrate only. Lanes 2-8 indicate the reactions with APE1 at the concentrations of 0.1 nM-100 nM. The quantification of the APE1 incision product from (A) and (B) is shown below the gels. “\*\*” indicates a significant difference in the products between the duplex DNA substrate and R-loop substrate ( $P < 0.05$ ). “\*\*\*” indicates a significant difference with  $P < 0.01$ .

## **The DNA synthesis activity of DNA polymerase $\beta$ synthesis is inhibited during BER in TNR R-loops**

Our previous studies have shown that pol  $\beta$  can bypass loop structures on the TNR template strand to promote repeat deletion on duplex DNA [91]. Also, we demonstrated that weak DNA synthesis activity of pol  $\beta$  containing the R137Q polymorphism leads to small CAG repeat deletions [159]. Thus, we determined the DNA synthesis activity of pol  $\beta$  at the different concentrations (0.1 nM to 50 nM) in the TNR R-loops harboring the (rGAA)<sub>20</sub>/(TTC)<sub>20</sub> and (rCAG)<sub>20</sub>/(CTG)<sub>20</sub> with an abasic site in the middle of the non-template strand and compared the activity with that from the (GAA)<sub>20</sub> and (CAG)<sub>20</sub> duplex substrates (Figure 1.3). In this assay, the abasic site is first cleaved by APE1, then the addition of pol  $\beta$  will allow extension from the incised base to the end of the template. If the template strand is engaged in an R-loop, access to pol  $\beta$  could be impeded. We found that a low concentration of pol  $\beta$  at 0.1 nM failed to insert a nucleotide at both duplex and R-loop substrates containing (GAA)<sub>20</sub> and (CAG)<sub>20</sub> repeats (Fig. 2A and 3B, lane 3). With increasing concentrations of pol  $\beta$  from 1 nM to 50 nM, pol  $\beta$  DNA synthesis products were significantly increased with all the substrates (Fig. 1.3A and 1.3B, lanes 4-8). However, at the tested concentrations, pol  $\beta$  produced significantly less DNA synthesis products on the TNR R-loop substrates (~20% for the GAA and 40% for CAG repeat R-loop) than those generated from the duplex substrates (50%-70%) (Fig. 2.3A and 2.3B, the bar charts below the gels. Nonetheless, pol  $\beta$  was able to extend the cleaved substrate even when the template strand was



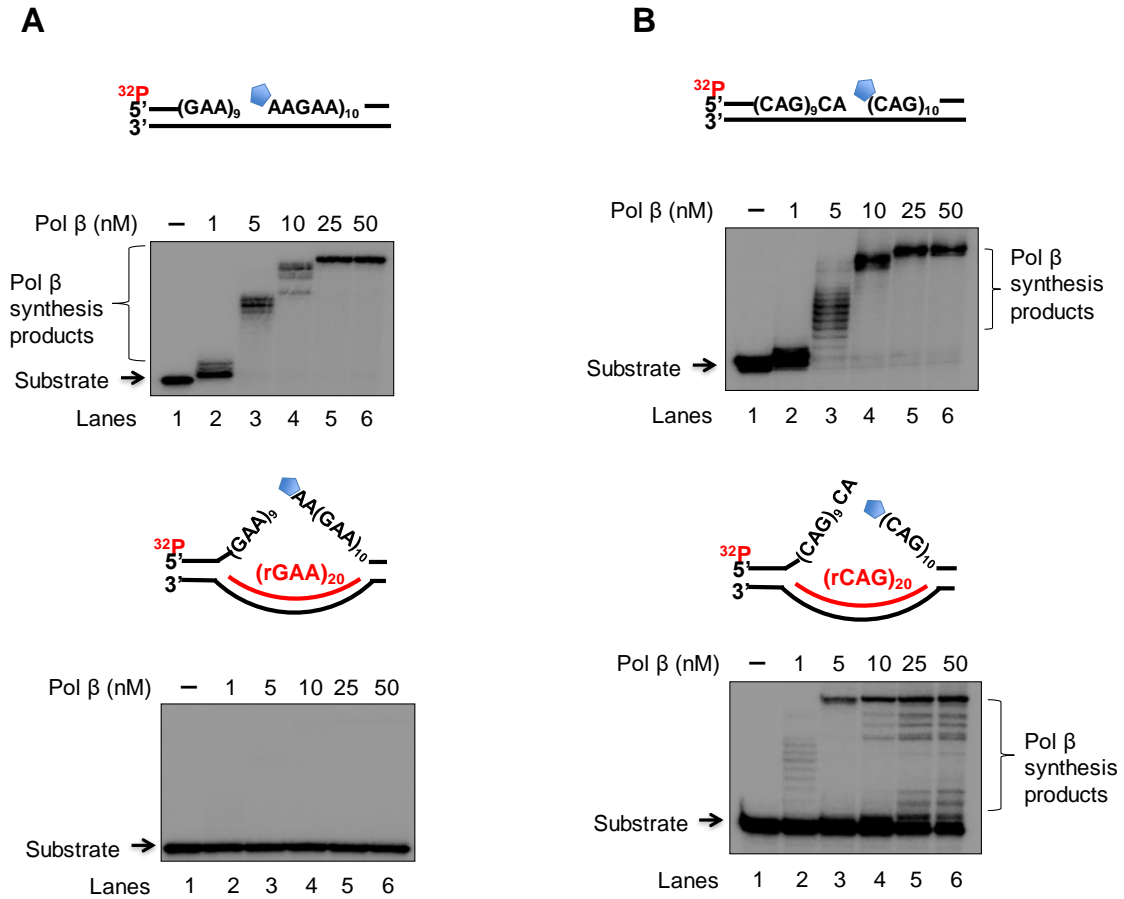
engaged in an R-loop, indicating that pol  $\beta$  can utilize an R-loop substrate for repair synthesis.



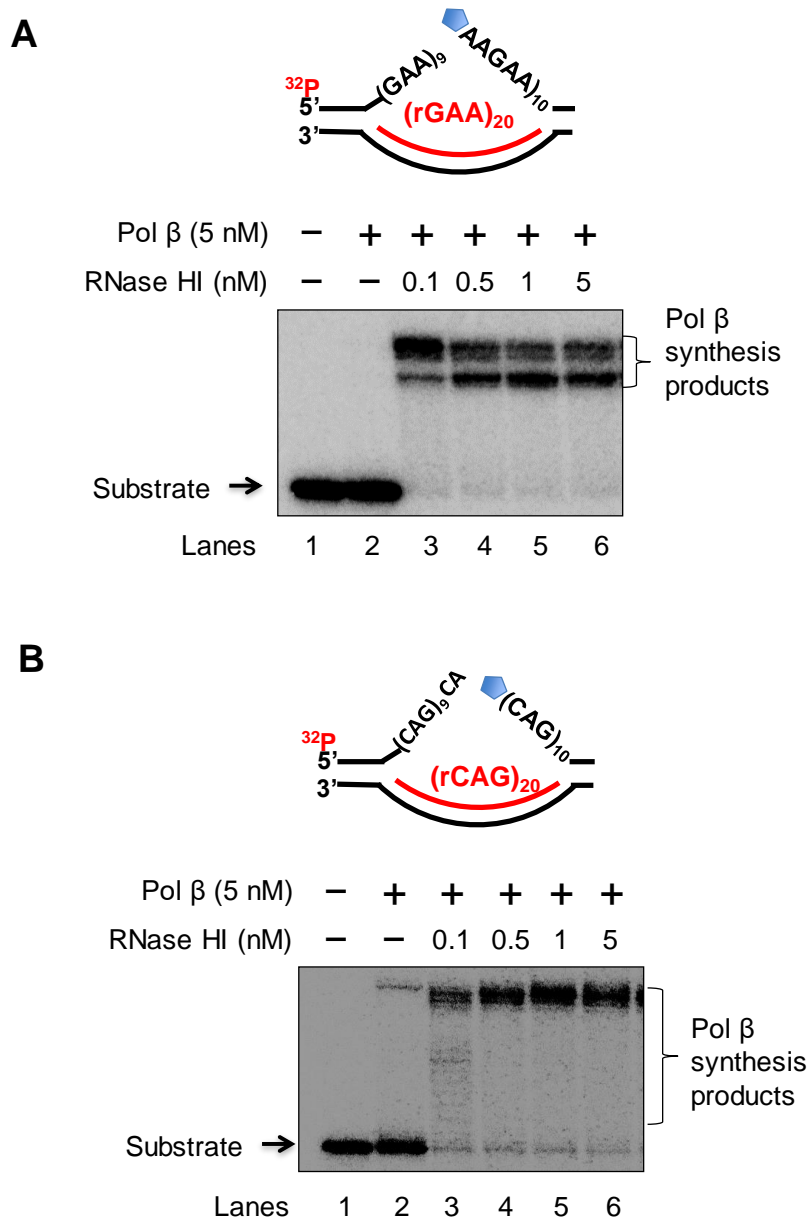
**Figure 2.3 Pol  $\beta$  DNA synthesis on TNR R-loops.** Pol  $\beta$  DNA synthesis activity in TNR R-loops was determined by incubating various concentrations of pol  $\beta$  (0.1 nM-50 nM) with 10 nM duplex DNA or R-loop substrates containing (GAA)<sub>20</sub> or (CAG)<sub>20</sub> with an abasic site in the repeats at 37°C for 30 min. Substrates (10 nM) were <sup>32</sup>P-labeled at the 5'-end of the strand containing an abasic site and incubated with 25 nM APE1 and increasing concentrations of pol  $\beta$  (0.1 nM-50 nM). Substrates and products were separated in a 15% urea-denaturing polyacrylamide gel and detected by phosphorimager. Substrates are schematically illustrated above the gels. Pol  $\beta$  DNA synthesis products are indicated by arrows. (A) Pol  $\beta$  DNA synthesis on the duplex DNA or R-loop containing (GAA)<sub>20</sub> repeats substrates with an abasic lesion in the middle of the repeats. (B) Pol  $\beta$  DNA synthesis on the duplex DNA or R-loop substrates containing (CAG)<sub>20</sub> repeats with an abasic site embedded in the middle of the repeats. Lane 1 represents the substrate only. Lane 2 indicates the reaction with 25 nM APE1. Lanes 3-8 indicate the reactions with APE1 and different concentrations of pol  $\beta$  (0.1 nM-50 nM). The quantification of the pol  $\beta$  DNA synthesis products is illustrated below the gels. “\*” indicates a significant difference in the products between the duplex DNA and R-loop substrate with  $P < 0.05$ . “\*\*” indicates a significant difference with  $P < 0.01$ .

To characterize the reduced DNA synthesis activity of pol  $\beta$  on the TNR R-loops, we examined pol  $\beta$  DNA synthesis activity on nicked (CAG)<sub>20</sub> or (GAA)<sub>20</sub> duplex and R-loop substrates that harbor a 3'- and 5'- repeat flap on the non-template strand in the presence of a (rGAA)<sub>20</sub>/(TTC)<sub>20</sub> or (rCAG)<sub>20</sub>/(CTG)<sub>20</sub>. These substrates mimic the intermediates with an abasic site in TNR R-loops that is incised by APE1 (Fig. 2.4). The results showed that pol  $\beta$  at the concentrations ranging from 1 nM to 50 nM performed efficient DNA synthesis on the duplex substrates containing (GAA)<sub>20</sub> or (CAG)<sub>20</sub> repeats (Figure 2.4A and 2.4B, the top panels). However, no pol  $\beta$  DNA synthesis products were detected with the substrate containing the GAA repeat flaps (Figure 2.4A, the bottom panel, lanes 2-6) indicating that pol  $\beta$  DNA synthesis was completely inhibited by the presence of the 3'-repeat flaps. Pol  $\beta$  DNA synthesis on nicked-(CAG)<sub>20</sub> R-loop with the 3'- and 5'-flap generated much fewer products than those generated from the duplex repeat substrate (Fig. 2.4B, the bottom panel, lanes 2-6) indicating that the DNA synthesis activity was also significantly inhibited by the presence of the RNA:DNA hybrid. These results suggest that the inhibition of pol  $\beta$  DNA synthesis activity resulted from a 3'-flap formed at the R-loop intermediates. To further confirm this, we used *E. coli* RNase HI to remove the RNA strand of the TNR R-loops substrates and examined the pol  $\beta$  DNA synthesis activity. We found that the removal of RNA in the TNR R-loops significantly stimulated pol  $\beta$  DNA synthesis activity (Fig. 2.5), indicating the necessity of the RNA strand of the R-loop for the inhibition of pol  $\beta$

activity. The results suggest that Pol  $\beta$  synthesis activity in the TNR R-loops was significantly inhibited by the formation of a 3'-flap in the R-loops during BER.



**Figure 2.4 Pol  $\beta$  DNA synthesis on the nicked duplex TNR and nicked TNR R-loops.** Pol  $\beta$  DNA synthesis on the nicked duplex TNRs or R-loops resulting from APE1 incision of an abasic site was determined by incubating the nicked TNR duplex or R-loop substrates with pol  $\beta$  at 1 nM-50 nM at 37°C for 30 min. Substrates were  $^{32}\text{P}$ -labeled at the 5'-end of the strand containing an abasic site. Substrates and products were separated in a 15% urea-denaturing polyacrylamide gel and detected by phosphorimager. Substrates are schematically illustrated above the gels. Pol  $\beta$  DNA synthesis products are indicated by arrows. (A) Pol  $\beta$  DNA synthesis on the duplex and R-loop substrates containing (GAA)<sub>20</sub> repeats with an abasic site in the middle of the repeats precut by APE1 incision. (B) Pol  $\beta$  DNA synthesis on the duplex and R-loop substrates containing (CAG)<sub>20</sub> repeats precut by APE1 incision. Lane 1 represents the substrate alone. Lanes 2-6 indicate the reactions with different concentrations of pol  $\beta$  at 1 nM-50 nM.

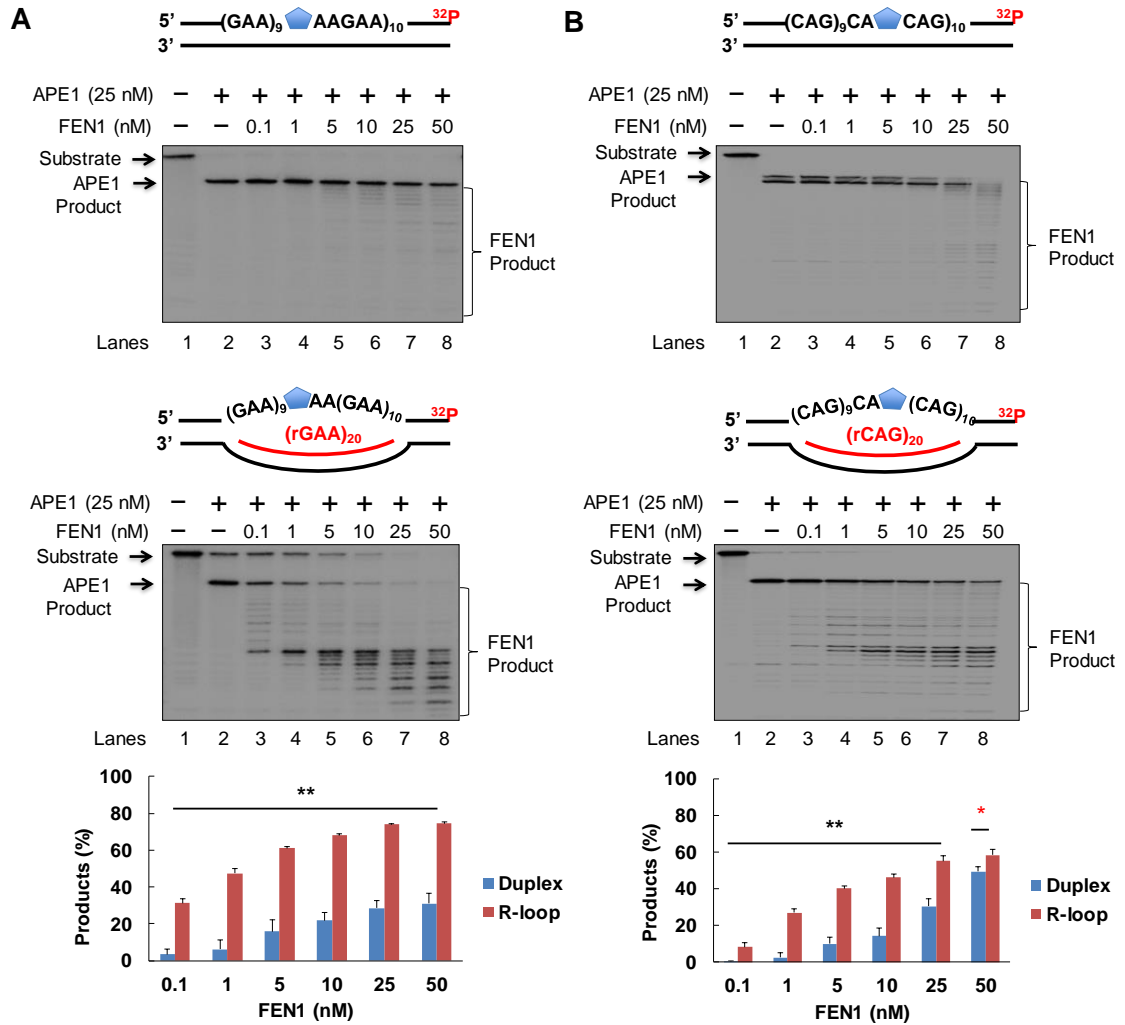


**Figure 2.5 Removal of the RNA strand in TNR R-loops by bacterial RNase HI restores the efficient synthesis of TNRs by pol β.** The effects of the RNA strand in TNR R-loops on the DNA synthesis activity of pol β was tested by determining pol β DNA synthesis activity in the presence of bacterial RNase HI, which cleaves the RNA strand in TNR R-loops. Nicked R-loop substrates containing the (GAA)<sub>20</sub> (A) or (CAG)<sub>20</sub> repeats (B) were incubated with 5 nM pol β and various concentrations of RNase HI (0.1 nM-5 nM) at 37°C for 30 min. Lane 1 indicates the substrate alone. Lane 2 indicates the reaction with 5 nM pol β. Lanes 3-6 indicate the reactions with 5 nM pol β and RNase HI at 0.1 nM-5 nM. Substrates were <sup>32</sup>P-labeled at the 5'-end of the upstream strand of the substrates and illustrated above the gels. Substrates and products were separated in 15% urea-denaturing polyacrylamide gel and detected by phosphorimager.

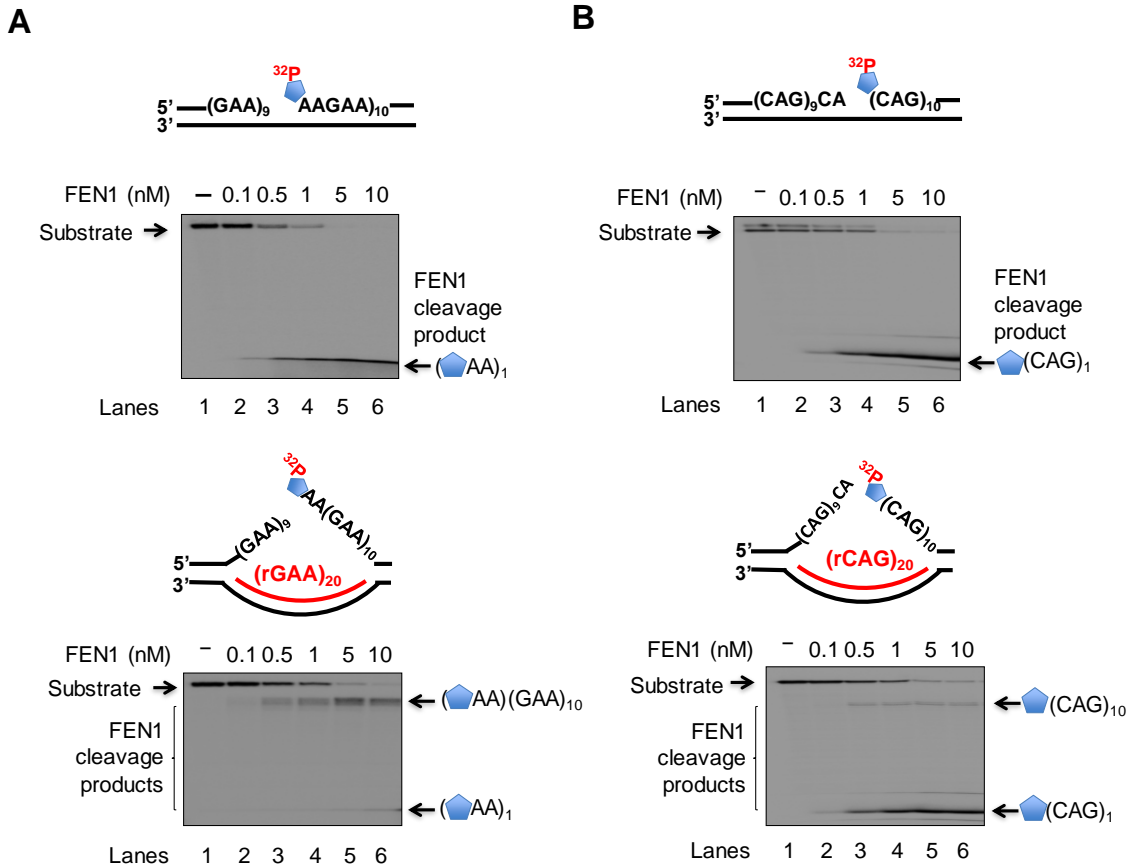
## **FEN1 cleavage of TNRs is stimulated during BER in TNR R-loops**

FEN1 flap cleavage activity can be inhibited by the secondary structures formed by TNRs during DNA replication [162-165] and BER [4, 90] resulting in TNR repeat expansion. To test whether FEN1 flap cleavage activity can also be altered during BER in TNR R-loops, we measured the FEN1 cleavage activity on the duplex DNA substrates containing (GAA)<sub>20</sub> or (CAG)<sub>20</sub> repeats with an abasic site in the middle of the repeats and (GAA)<sub>20</sub> and (CAG)<sub>20</sub> R-loop substrates containing an abasic site in the middle of the non-template strand (Fig 2.6). We found that increasing concentrations of FEN1 at 0.1-50 nM exhibited weak flap cleavage activity generating 1% to 20% cleavage products from the (GAA)<sub>20</sub> repeat duplex substrate and 1% to 50% products from the (CAG)<sub>20</sub> repeat duplex substrate (Fig. 2.6A and 2.6B, lanes 3-8 and the bottom panel). Surprisingly, we found that FEN1 at 0.1 nM cleaved ~30% of the (GAA)<sub>10</sub> flap and 10% of the (CAG)<sub>10</sub> flap from the R-loop substrates (Fig. 2.6A and 2.6B, lane 3 of the middle panel). This indicates that a (GAA)<sub>10</sub> or (CAG)<sub>10</sub> 5'-flap was generated more frequently or more stably in the presence of the RNA:DNA hybrid in the R-loop substrates, thereby stimulating FEN1 flap cleavage. To characterize the FEN1 cleavage activity on the TNR R-loop substrates, we examined the FEN1 cleavage on the nicked (GAA)<sub>20</sub> or (CAG)<sub>20</sub> duplex substrate and (GAA)<sub>20</sub> or (CAG)<sub>20</sub> nicked-R-loop substrate containing a 3'-(GAA)<sub>9</sub> or (CAG)<sub>9</sub> flap and 5'-(GAA)<sub>10</sub> or (CAG)<sub>10</sub> flap with a (rGAA)<sub>20</sub>/(TTC)<sub>20</sub> or (rCAG)<sub>20</sub>/(CTG)<sub>20</sub> RNA: DNA hybrid (Fig. 2.7). This time, we included the <sup>32</sup>P label on the 5' end of the flap, so that we could follow the fate of the cleaved off nucleotides. Our results showed that FEN1 mainly

cleaved one repeat from the 5'-end of the downstream strand of the (GAA)<sub>20</sub> or (CAG)<sub>20</sub> nicked-duplex substrate (Figure 2.7A and 2.7B, lanes 2-6 of the top panel), indicating that a small flap of one repeat is usually displaced. However, FEN1 cleavage on the nicked (GAA)<sub>20</sub> R-loop substrate predominantly generated the (GAA)<sub>11</sub> flap cleavage product along with the product containing one repeat with the THF (Figure 2.7A, lanes 2-6 of the bottom panel). Similarly, for the nicked (CAG)<sub>20</sub> R-loop substrate, FEN1 at all concentrations resulted in the products containing (CAG)<sub>10</sub> repeats and (CAG)<sub>1</sub> with the THF (Figure 2.7B, lanes 2-6 of the bottom panel). The results indicate that FEN1 flap cleavage activity was significantly stimulated in the presence of the RNA:DNA hybrid during BER in the TNR R-loops. Also, the size of the cleaved off flap is increased substantially, which is determined by the length of the RNA:DNA hybrid. We found that the removal of the RNA strand of the TNR R-loop substrates by bacterial RNase HI resulted in the FEN1 cleavage products containing one repeat with the THF (Fig. 2.8) indicating that the presence of the RNA strand was required for the stimulation of FEN1 cleavage of long TNR flaps.

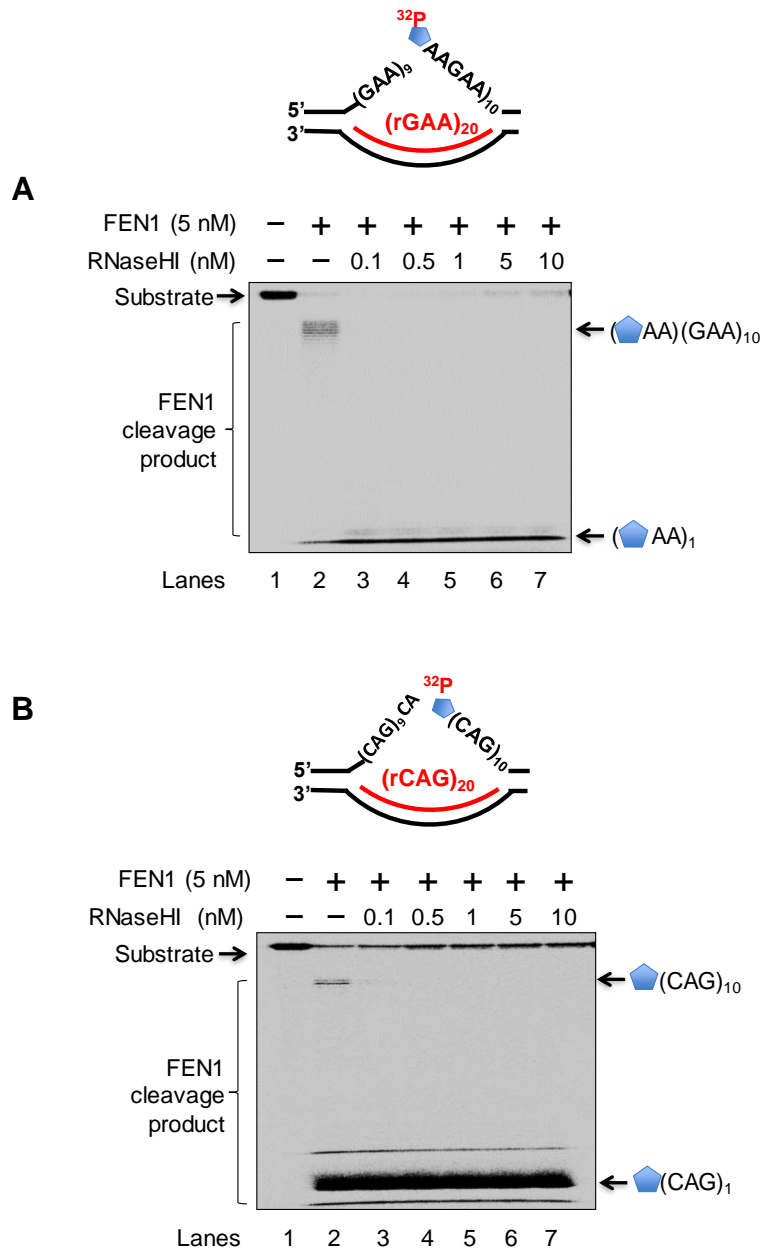


**Figure 2.6 FEN1 cleavage activity of TNRs on R-loops.** FEN1 cleavage of TNRs during BER in duplex TNRs or R-loops was determined by incubating the (GAA)<sub>20</sub> or (CAG)<sub>20</sub> repeat duplex and R-loop substrate with FEN1 (0.1 nM-25 nM) at 37°C for 30 min. Substrates were <sup>32</sup>P-labeled at the 3'-end of the strand containing an abasic site and are illustrated above the gels. Substrates and products were separated using a 15% urea-denaturing polyacrylamide gel and detected by phosphorimager. The quantification of FEN1 cleavage products is shown in the bar chart below the gels. (A) FEN1 cleavage activity on duplex or R-loop substrate containing (GAA)<sub>20</sub> repeats with an abasic site in the middle of the repeats. (B) FEN1 cleavage on the duplex or R-loop substrate containing (CAG)<sub>20</sub> repeats with an abasic site in the repeats. Lane 1 indicates the substrate only. Lane 2 indicates the reaction with 25 nM APE1. Lane 3-8 indicate the reactions with 0.1 nM-50 nM of FEN1 in the presence of 25 nM APE1. “\*\*” indicates a significant difference in the FEN1 cleavage products between the duplex DNA substrate and R-loop substrate with  $P < 0.05$ . “\*\*\*” indicates a significance with  $P < 0.01$ .



**Figure 2.7 FEN1 cleavage activity on the nicked TNR R-loops.** FEN1 cleavage of TNRs on the nicked duplex DNA and R-loop was examined by incubating the nicked duplex or R-loop substrates (10 nM) containing (GAA)<sub>20</sub> or (CAG)<sub>20</sub> repeats with FEN1 (0.1 nM-25 nM) at 37°C for 30 min. Substrates were <sup>32</sup>P-labeled at the 5'-end of the downstream strand and illustrated above the gels. Substrates and products were separated using a 15% urea-denaturing polyacrylamide gel and detected by phosphorimager. FEN1 cleavage products are indicated by arrows. (A) FEN1 cleavage activity on the duplex substrate containing (GAA)<sub>10</sub> repeats with an abasic site in the middle of the repeats (the gel on the top) and double-flap substrate resulting from a (GAA)<sub>20</sub> repeat R-loop containing an abasic site in the middle of the repeats (the gel on the bottom). (B) FEN1 cleavage of CAG repeats on the duplex substrate containing (CAG)<sub>20</sub> repeats with an abasic site in the middle of the repeats (the gel on the top) and R-loop substrate containing (CAG)<sub>20</sub> with an abasic site in the repeats (the gel at the bottom). Lane 1 indicates the substrate only. Lanes 2-7 indicate the reactions with different concentrations of FEN1 (0.1 nM-25 nM).

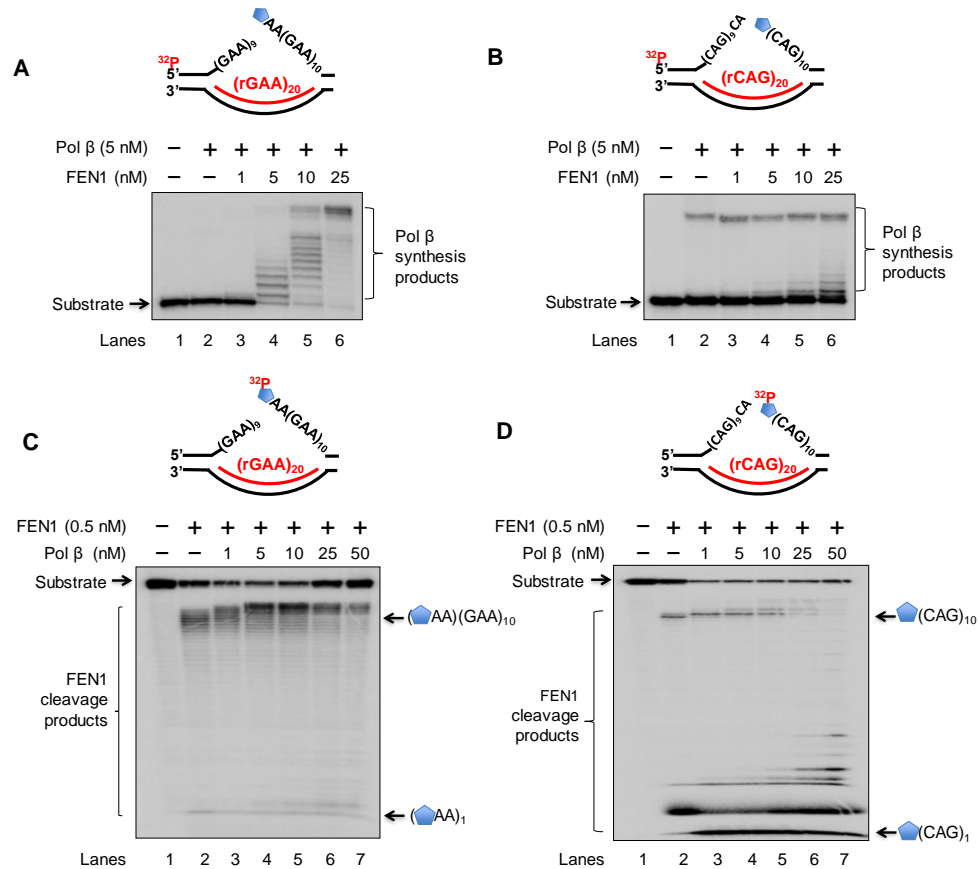




**Figure 2. 8 Removal of the RNA strand in TNR R-loops prevents the cleavage of a long repeat flap by FEN1.** The effects of the RNA strand in R-loops on FEN1 cleavage of TNRs was examined by detecting the size of the FEN1 cleavage products on (GAA)<sub>20</sub> (A) or (CAG)<sub>20</sub> (B) repeat R-loops in the absence or presence of bacterial RNase HI at the concentrations of 0.1 nM-10 nM. Lane 1 represents the substrate only. Lane 2 indicates the reaction with the substrate with 5 nM FEN1. Lanes 3-7 indicate the reactions with 5 nM FEN1 and increasing concentrations of RNase HI at 0.1 nM-10 nM. Substrates were <sup>32</sup>P-labeled at the 5'-end of the downstream primer and schematically illustrated above the gels. FEN1 cleavage products were indicated by arrows and separated from the substrates in 15% urea-denaturing polyacrylamide gel. Substrates and products were detected by phosphorimager.

## **Pol $\beta$ DNA synthesis is stimulated in the presence of FEN1 during BER in TNR R-loops**

Previous studies have shown that pol  $\beta$  coordinates with FEN1 alternate flap cleavage to promote TNR expansion during BER [4, 90]. To further determine the coordination between pol  $\beta$  and FEN1 in a TNR R-loop, we examined if pol  $\beta$  and FEN1 can influence each other during BER in the (GAA)<sub>20</sub> and (CAG)<sub>20</sub> R-loop substrates by testing their enzymatic activity in the presence of the other (Fig. 2.9). The results showed that for the nicked (GAA)<sub>20</sub> repeat R-loop substrate, pol  $\beta$  did not perform DNA synthesis in the absence of FEN1 or in the presence of a low concentration of FEN1 at 1 nM (Fig. 2.9A, lanes 2-3). However, with increasing concentrations of FEN1 from 5 nM to 25 nM, pol  $\beta$  was able to synthesize GAA repeats, and the synthesis products were significantly increased (Fig. 2.9A, lanes 4-6). Similarly, for the nicked (CAG)<sub>20</sub> repeat R-loop substrate, pol  $\beta$  DNA synthesis products were detected in the presence of 10 nM and 25 nM FEN1 but not 1 nM and 5 nM FEN1 (Fig. 2.9B, compare lanes 5-6 with lanes 3-4). The results indicate that the presence of FEN1 flap cleavage stimulated the pol  $\beta$  synthesis of the repeats. We then tested if pol  $\beta$  DNA synthesis affected the FEN1 flap cleavage activity. The results showed that pol  $\beta$  at 1 nM-10 nM stimulated the

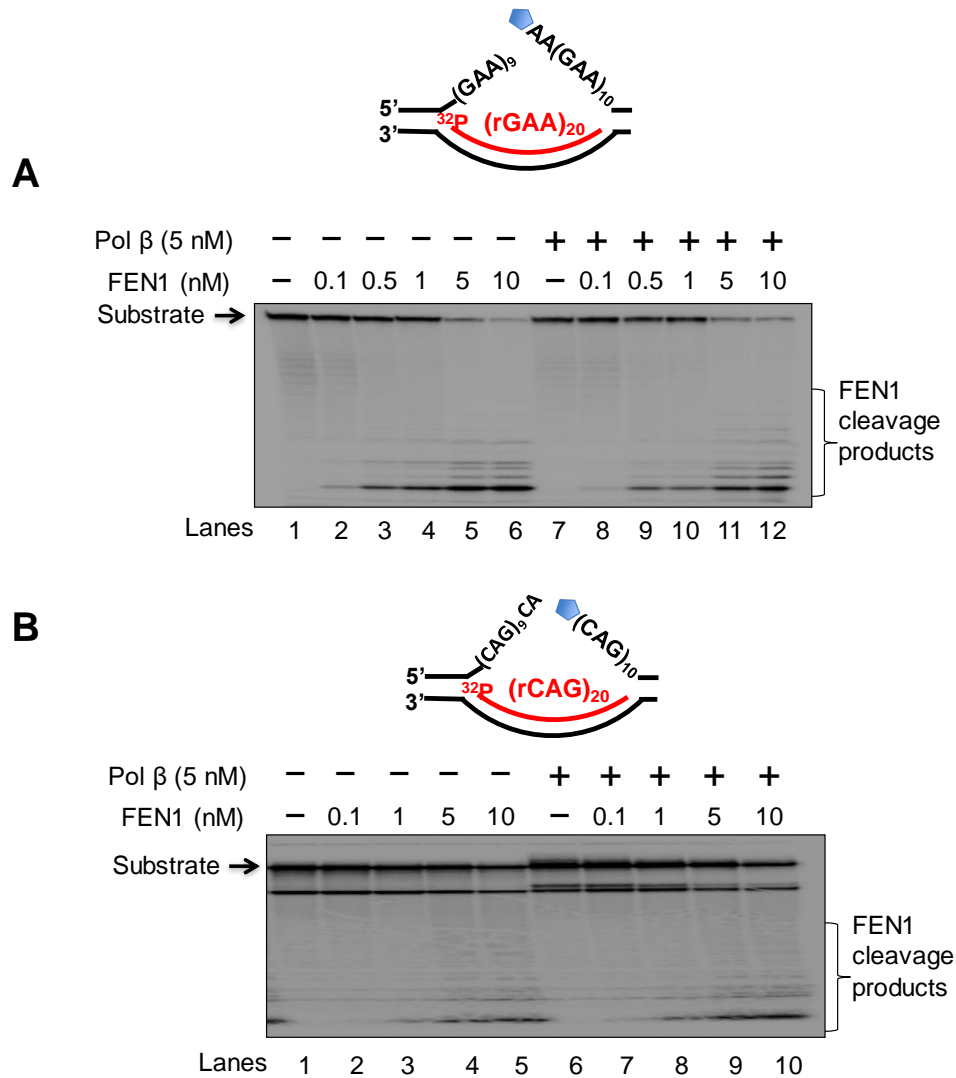


**Figure 2.9 Coordination of FEN1 flap cleavage and pol β DNA synthesis in TNR R-loops during BER.** The coordination of FEN1 cleavage of TNRs and pol β synthesis of the repeats in R-loops was determined by testing pol β DNA synthesis in the presence of various concentrations of FEN1 or by examining FEN1 cleavage activity with the presence of different concentrations of pol β. Substrates (10 nM) were incubated with 5 nM pol β and different concentrations of FEN1 (1 nM-25 nM) (A) and (B) or 0.5 nM FEN1 and increasing concentrations of pol β (1 nM-50 nM) (C) and (D) at 37°C for 30 min. (A) and (B) showed the pol β DNA synthesis activity in the presence of FEN1 at the concentrations of 1 nM-25 nM with the nicked  $(GAA)_{20}$  and  $(CAG)_{20}$  repeat R-loop substrates containing an abasic site in the repeats. Substrates were  $^{32}P$ -labeled at the 5'-end of the strand containing the nick. Lane 1 represents the substrate alone. Lane 2 indicates the reaction with 5 nM pol β. Lanes 3-6 represent the reactions with 5 nM pol β in the presence of various concentrations of FEN1 at 1 nM-25 nM. (C) and (D) showed FEN1 cleavage of TNRs in the presence of various concentrations of pol β (1 nM-50 nM) with the nicked  $(GAA)_{20}$  and  $(CAG)_{20}$  repeat R-loop substrates containing an abasic site in the repeats. Lane 1 indicates the substrate alone. Lane 2 indicates the reaction with 0.5 nM FEN1. Lanes 3-6 represent the reactions with 0.5 nM FEN1 in the presence of different concentrations of pol β (1 nM-50 nM). Substrates were  $^{32}P$ -labeled at the 5'-end of the downstream strand and illustrated above the gels. Substrates and products were separated in a 15% urea-denaturing polyacrylamide gel and detected by phosphorimager. Pol β DNA synthesis products and FEN1 cleavage products are indicated by arrows.

FEN1 cleavage of the downstream long and short 5'-GAA and -CAG repeat flaps (Fig. 2.9C and 6D, lanes 3-5) indicating that pol  $\beta$  DNA synthesis performed the strand displacement synthesis to facilitate the FEN1 cleavage of a long repeat flap. However, pol  $\beta$  at 25 nM and 50 nM inhibited FEN1 cleavage of a long repeat flap and promoted FEN1 alternate flap cleavage of a short repeat flap (Fig. 2.9C and 6D, lanes 6-7). This suggests that the excessive pol  $\beta$  DNA synthesis at the high concentrations displaced the RNA strand of the TNR R-loops. This then prevented FEN1 from binding to the flaps leading to the formation of a long TNR flap. Subsequently, the long TNR flaps form small repeat bubbles and hairpins that inhibited FEN1 cleavage of the repeats leading to its inefficient alternate flap cleavage of short TNR repeats.

### **FEN1 cleaves the RNA strand in TNR R-loops during BER**

To exclude the possibility that FEN1 flap cleavage may affect the stability of the RNA:DNA hybrid during BER in TNR R-loops, we further tested if FEN1 can cleave the RNA from the RNA:DNA hybrid in the TNR R-loop substrates by examining the cleavage activity of FEN1 on the RNA strand in the nicked TNR R-loop substrates (Fig. 2.10). Surprisingly, we found that FEN1 at 0.1 nM-10 nM efficiently cleaved the RNA strand from the RNA:DNA hybrid in the nicked (GAA)<sub>20</sub> and (CAG)<sub>20</sub> repeat R-loop substrates (Fig. 2.10A, lanes 2-6 and lanes 8-12, 2.10B, lanes 2-5 and lanes 7-10). We further demonstrated that the FEN1 cleavage pattern and activity on the RNA strand was not altered by pol  $\beta$  DNA synthesis (Fig. 2.10A, lanes 8-12, 2.10B, lanes 7-10). The results showed that

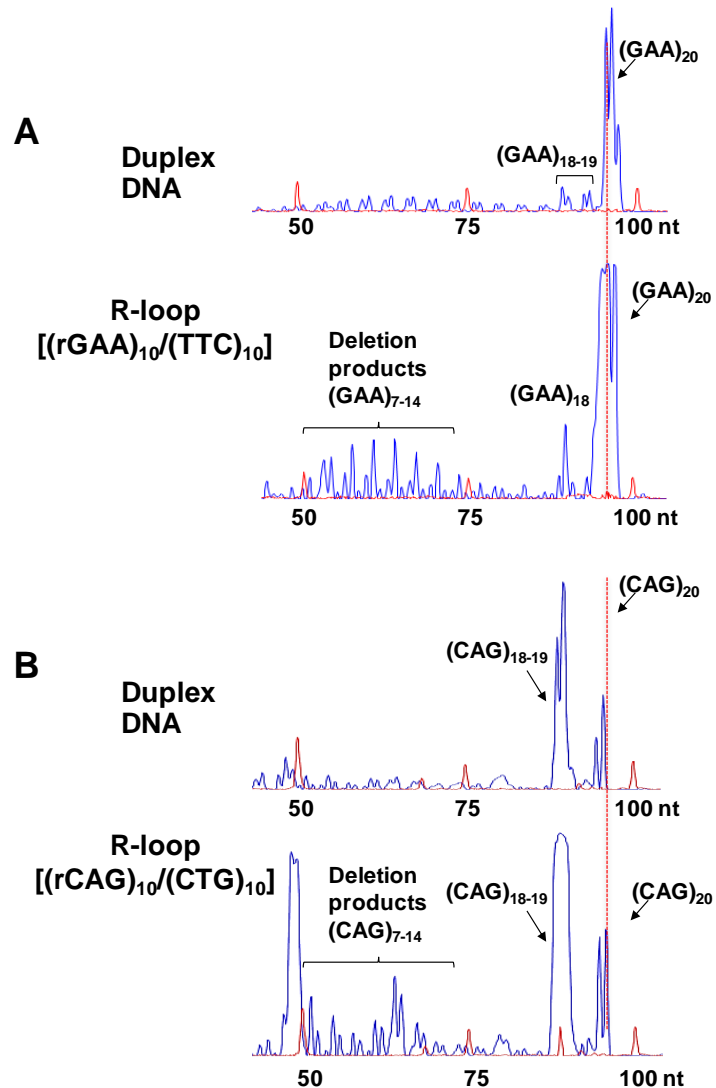


**Figure 2.10 FEN1 cleaves the RNA strand of TNR R-loops.** The cleavage of the RNA strand of TNR R-loops by FEN1 was determined using the nicked-R-loop substrates containing an (rGAA)<sub>20</sub>/(TTC)<sub>20</sub> (A) or (rCAG)<sub>20</sub>/(CTG)<sub>20</sub> (B) RNA:DNA hybrid. Substrates were <sup>32</sup>P-labeled at the 5'-end of the RNA strand in the R-loop substrates were schematically illustrated above the gels. Substrates (10 nM) were incubated with FEN1 in the absence or presence of pol β at 37°C for 30 min. Substrates and products were separated in a 15% urea-denaturing polyacrylamide gel and detected by phosphorimager. (A) FEN1 cleavage of the (rGAA)<sub>20</sub> repeats in the GAA repeat nicked-R-loop substrate in the absence or presence of 5 nM pol β. Lane 1 represents the substrate only. Lanes 2-6 correspond to the reactions with various concentrations of FEN1 at 0.1 nM-10 nM without pol β. Lanes 7-12 represent the reactions with FEN1 at 0.1 nM-10 nM in the presence of 5 nM pol β. (B) FEN1 cleavage of the (rCAG)<sub>20</sub> repeats in the CAG repeat nicked-R-loop substrate without or with 5 nM pol β. Lane 1 represents the substrate only. Lanes 2-5 correspond to the reactions with various concentrations of FEN1 at 0.1 nM-10 nM without pol β. Lanes 7-10 represent the reactions with FEN1 at 0.1 nM-10 nM in the presence of 5 nM pol β. All experiments were done in triplicates.

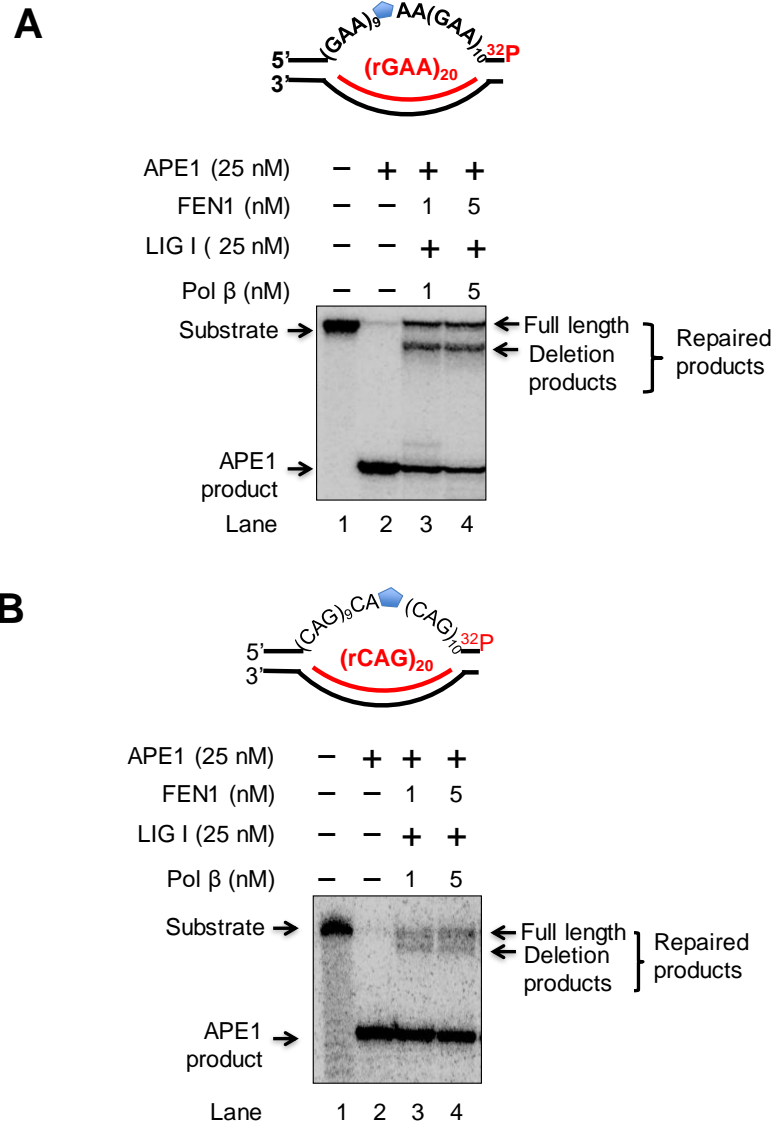
FEN1 cleaved the RNA strand in the TNR R-loop independent of pol  $\beta$  DNA synthesis during BER. The results indicate that FEN1 simultaneously removed the downstream 5'-repeat flap and RNA strand from the RNA:DNA hybrid of a TNR-R-loop, allowing the reannealing of the upstream strand to facilitate pol  $\beta$  DNA synthesis during BER.

### **BER on TNR R-loops promotes repeat deletion**

The Su and Freudenreich study demonstrated that BER of uracils generated from cytosine deamination in a CAG repeat R-loop, which are converted to abasic sites by Ung1, causes repeat deletion [120]. Our previous studies have shown that BER can induce TNR deletion during the repair of an abasic site located in the middle of duplex TNR tracts via weak pol  $\beta$  DNA synthesis [82, 92]. To further examine if the weak synthesis of TNRs by pol  $\beta$  and efficient removal of the repeats by FEN1 during BER in a TNR R-loop can result in repeat deletion, we determined TNR instability resulting from BER in a TNR R-loop (Fig. 2.11). We found that BER in the (GAA)<sub>20</sub> and (CAG)<sub>20</sub> repeat R-loop substrates resulted in repaired products containing both full length and deletion products (Fig. 2.12). Further analysis of the size of the repaired products showed that BER led to the products with both full-length and shorter repeats (Fig. 2.11). For the GAA repeat R-loop substrate, BER resulted in a large amount of the deletion products containing (GAA)<sub>18</sub> and (GAA)<sub>7-14</sub> repeats (Figure 2.11A). Similarly, for the CAG repeat R-loop substrate, BER resulted in deletion products containing (CAG)<sub>18-19</sub> and (CAG)<sub>7-14</sub> (Fig.2.11B). However, BER of an abasic site in the middle of duplex GAA and CAG repeats



**Figure 2.11 BER in R-loops promotes TNR deletion.** To test the effects of R-loops on TNR instability via BER, BER in TNR R-loops was reconstituted by incubating purified BER enzymes and the  $(GAA)_{20}$  and  $(CAG)_{20}$  repeat duplex or R-loop substrates containing an abasic site in the middle of the repeats at  $37^{\circ}\text{C}$  for 30 min. The repaired products were isolated and amplified by PCR. The PCR-amplified repaired products were further analyzed by capillary electrophoresis. The size of TNRs in the repaired products was determined by DNA fragment analysis. (A) BER in the  $(GAA)_{20}$  duplex DNA substrate with an abasic site in the middle of the repeats led to the production of a small amount of repeat deletion products containing  $(GAA)_{18-19}$  repeats. BER in the  $(GAA)_{20}$  R-loop substrate containing an abasic site in the middle of the repeats resulted in the production of repeat deletion products of  $(GAA)_{7-14}$  and  $(GAA)_{18}$ . (B) BER was reconstituted with the  $(CAG)_{20}$  repeat duplex or R-loop substrate with an abasic site in the middle of the repeats. The repaired products were isolated and amplified by PCR. The sizes of the repeats in the repaired products were analyzed and determined by capillary electrophoresis and DNA fragment analysis. All experiments were done in triplicates.



**Figure 2.12 BER in TNR R-loops lead to repaired products.** The repaired products resulting from BER in TNR R-loops was determined by reconstituting BER reactions with purified APE1 (25 nM), LIG I (25 nM), pol β (1 nM-5 nM), and FEN1 (1 nM-5 nM) and 10 nM (GAA)<sub>20</sub> (A) or (CAG)<sub>20</sub> (B) repeat R-loop substrate. Substrates were <sup>32</sup>P-labeled at the 3'-end and incubated with the BER enzymes at 37°C for 30 min. The repaired products were separated from the substrates and repair intermediates in a 15% urea-denaturing polyacrylamide gel. Substrates and products were detected by phosphorimager. Lane 1 represents the substrate alone. Lane 2 represents the reaction with 25 nM APE1. Lanes 3-4 represent the reactions reconstituted with purified BER enzymes. The repaired products and intermediates are indicated by arrows. Substrates were illustrated above the gels. All experiments were done in triplicates.

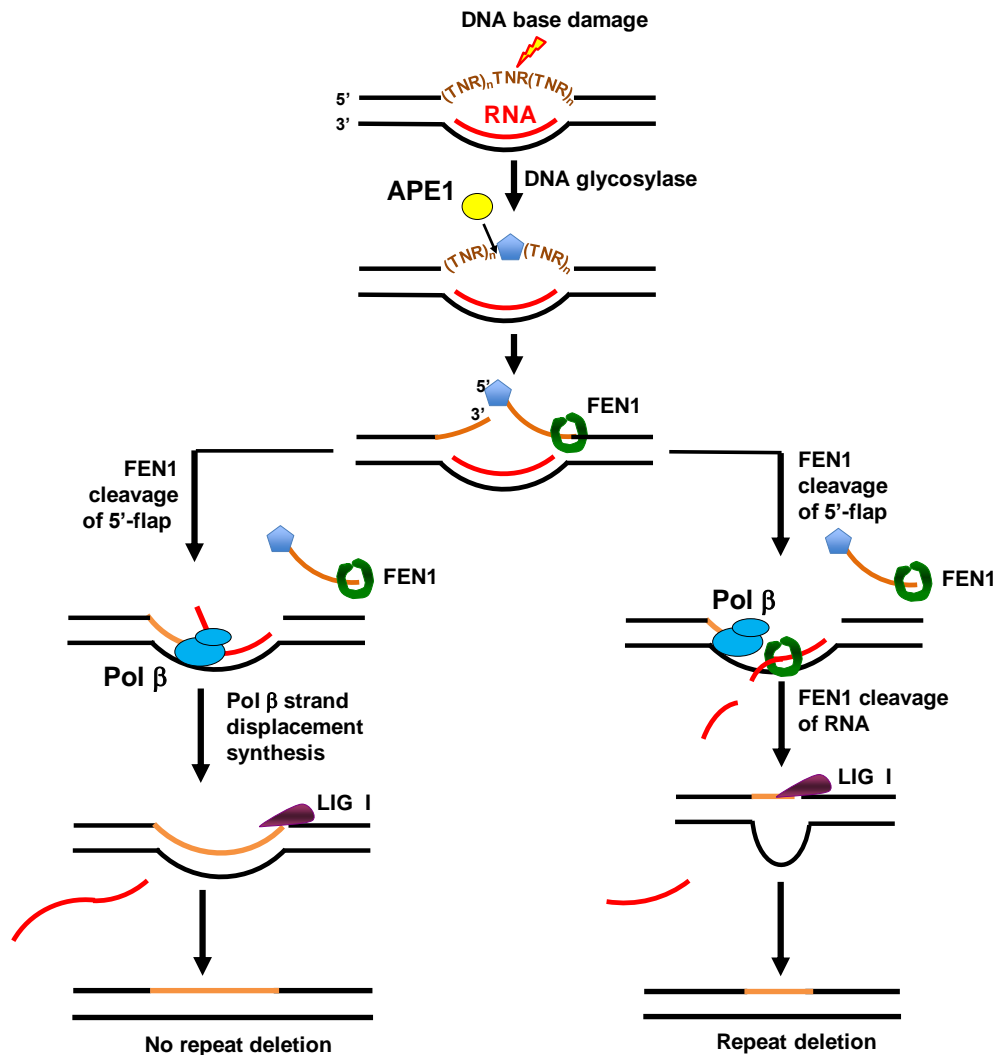


only resulted in a small number of deletion products containing (GAA)<sub>18-19</sub> and (CAG)<sub>18-19</sub> repeats, indicating loss of only 1-2 repeat units (Fig. 2.11A and 2.11B). These results suggest that BER in TNR R-loops promotes more and larger TNR deletions.

## DISCUSSION

It has been proposed that the single-strand non-template strand of R-loops [5, 119, 176] is susceptible to DNA damage, including DNA base damage. Thus, it is possible that BER, acting on the damage in TNR R-loops, is involved in mediating repeat instability. This is supported by a recent study from the Freudenreich group showing that cytosine deamination in expanded CAG repeats engaged in R-loops results in repeat deletions [120]. This study showed that deletions in the CAG tract were suppressed when BER pathway members UNG1 and Ape1/APE1 were deleted, suggesting that the BER of abasic sites generated by processing deaminated cytosines in the non-template strand of the CAG repeat R-loops underlies the repeat instability. However, the molecular basis of R-loop-mediated TNR deletion remained unknown. In this study, we explored the underlying mechanisms of R-loop-induced repeat instability by determining the activities of BER enzymes and their coordination during BER in TNR R-loops. We demonstrated that an abasic site in the non-template strand of (CAG)<sub>20</sub> and (GAA)<sub>20</sub> repeat R-loops was incised by APE1 resulting in a double-flap intermediate containing an RNA:DNA hybrid with a 5'-flap and 3'-flap (Fig. 2.1). Furthermore, pol  $\beta$  synthesis of repeats engaged in RNA:DNA hybrids were

significantly inhibited, with rGAA/dTTC being even more inhibitory than rCAG/dCTG (Fig. 2.2 and 2.3). In contrast, FEN1 flap cleavage of the 5'-flap was significantly stimulated and biased more toward cleavage at the end of the hybrid, resulting in the release of a large flap (Fig. 2.4 and 2.5). We found that the presence of FEN1 facilitated pol  $\beta$  synthesis of TNRs (Fig. 2.9A and 2.9B). However, the presence of pol  $\beta$  suppressed FEN1 cleavage of the repeats (Fig. 2.9C and 2.9D). We further demonstrated that FEN1 efficiently cleaved the 5'-end of the RNA strand of the RNA:DNA hybrid of the TNR R-loops during BER (Fig. 2.10). Finally, BER in the R-loops led to an increase in TNR deletions and a shift to larger deletions compared to repair in a duplex substrate (Fig. 2.11). All the results support a hypothetical model for the repair of abasic lesions generated from a damaged DNA base in the non-template strand of a TNR R-loop (Fig. 2.13). APE1 incises the abasic site resulting in a single-strand break and the formation of a double-flap intermediate with an RNA:DNA hybrid on the template strand, along with an upstream 3'-flap and downstream 5'-flap. In the scenario where the upstream 3'-flap reanneals back to the template, this creates a 5'-RNA flap. Pol  $\beta$  will then perform strand displacement synthesis to displace the RNA strand, and the nick will be sealed by LIG I (the subpathway on the left). Since the template strand is relatively protected in this pathway and doesn't have a chance to form a DNA structure, this pathway will generally not lead to the deletion of TNRs. In the scenario where pol  $\beta$  synthesis of TNRs is inhibited by the RNA:DNA hybrid, FEN1 still cleaves the downstream 5'-flap efficiently. Subsequently, FEN1 could cleave the RNA strand in the TNR R-loops leading to the dissociation of the RNA from the



**Figure 2.13 A hypothetical model of BER in R-loops leads to TNR deletion.** DNA base damage that is induced in the non-template strand of TNR R-loops is removed by a DNA glycosylase leaving an abasic site that is incised by APE1 at the 5'-end. Subsequently, this results in a nick and the formation of a double-flap intermediate with an upstream 3'-flap and 5'-downstream 5'-flap stabilized by the RNA:DNA hybrid in the R-loop. FEN1 efficiently cleaves the 5'-flap, whereas pol  $\beta$  DNA synthesis is inhibited by the 3'-flap. In a scenario where 3'-flap reannealed to the template, it displaces the RNA strand resulting in the formation of a 5'-RNA flap. Pol  $\beta$  then performs a strand-displacement DNA synthesis to dislodge the RNA strand from the template strand leaving a nick for the ligation by LIG I. This results in no-repeat deletion (The subpathway on the left). In a scenario where FEN1 cleaves the RNA strand leaving a short segment of RNA that dissociates from the template. This results in the formation of a loop structure in the template strand. Subsequently, pol  $\beta$  skips over the template loop generating a ligable nick that is sealed by LIG I. Consequently, this results in more repeats removed by FEN1 than those synthesized by pol  $\beta$ , thereby leading to repeat deletion (the subpathway on the right).

R-loops and the formation of a loop structure in the template strand. Pol  $\beta$  then skips over the loop to perform DNA synthesis creating a deletion. This results in more repeats removed than synthesized during BER in R-loops, leading to large repeat deletions (Fig. 2.11).

The role of R-loops in causing TNR instability has been implicated by several studies [121, 178]. Loomis et al found that hairpin structures form in the single-strand non-template strand of a CGG repeat R-loop, suggesting that these secondary structures may underlie repeat expansions [180]. A recent study from Su and Freudenreich has demonstrated that R-loops formed in CAG repeats result in repeat deletions through DNA base damage and the BER pathway [120]. In this study, we further explored how TNR deletion can be mediated by BER in TNR R-loops. We showed that BER of DNA base lesions in the non-template strand of a TNR R-loop disrupted the coordination between pol  $\beta$  and FEN1 and the balance between the synthesis of TNRs by pol  $\beta$  and their removal by FEN1 via the inhibition of pol  $\beta$  DNA synthesis activity (Fig. 2.3) and stimulation of FEN1 flap cleavage activity (Fig. 2.6), respectively. We demonstrated that the RNA strand of the R-loops is playing a crucial role in modulating the activities of these BER enzymes (Fig. 2.5 and 2.8), presumably by displacing the up- and downstream strands to create the 3'- and 5'-flaps that subsequently inhibit pol  $\beta$  DNA synthesis and stimulate FEN1 flap cleavage during BER. We propose that this situation results in more TNR deletions either by removing the option of expansions by incorporation of unprocessed flaps or by increasing the possibility of template hairpins during pol  $\beta$  synthesis, or both.

Here, we identified a unique mechanism that leads to TNR deletion via BER in R-loops through the stimulation of FEN1 cleavage of TNR flaps that occur due to BER processing of R-loops (Fig. 2.13). However, we cannot exclude the possibility that other BER cofactors not studied here also play a role. For example, proliferating cell nuclear antigen (PCNA) that can stimulate FEN1 cleavage on TNR flaps [95] may also facilitate FEN1 cleavage of TNR flaps in R-loops, thereby further promoting repeat deletion through BER. Moreover, it is possible that the others 5' endonucleases, Exonuclease I (Exo I) [183] and Fanconi anemia associated nuclease 1 (FAN1) [155] may coordinate with FEN1 to remove the 5'-TNR flap formed on the non-template strand of TNR R-loops during BER. This is supported by the fact that both Exo 1 and FAN1 can prevent TNR expansions in mice [184, 185], and FAN1 human polymorphic variations are associated with a late age of onset of several CAG repeat expansion neurodegenerative diseases [149, 150, 186]. The synergistic effects of the 5'-endo and exonucleases on the processing of TNR R-loops need to be elucidated in the future.

Our study also suggests that BER-mediated TNR deletion via R-loops can serve as a new pathway that resolves R-loops formed in TNRs, attenuates TNR expansion and prevents R-loop-induced gene silencing in the expanded TNRs. It has been shown that R-loops are generated during gene transcription in open chromatin and can inhibit DNA methyltransferase I activity ensuring the sustainment of an open conformation of chromatin during gene transcription [187]. However, it is also found that R-loops formed on the expanded TNRs promote heterochromatinization via the recruitment of G9a methyltransferase and

increased H3K9me2 on the repeats inducing gene silencing [119]. This further indicates that R-loops on TNRs, if not resolved can ultimately lead to heterochromatinization and gene silencing. Thus, BER-mediated resolution of TNR R-loops and associated TNR deletions may be essential protection against the development of TNR diseases. This pathway could potentially be exploited as a new therapy for TNR expansion diseases by targeting expanded TNRs and their associated heterochromatinization and gene silencing.

## **CHAPTER 3. FLAP ENDONUCLEASE 1 CLEAVES RNAs TO RESOLVE R- LOOPS THROUGH DNA BASE EXCISION REPAIR**

### **ABSTRACT**

Flap endonuclease 1 (FEN1) is an essential enzyme that removes RNA primers and DNA base lesions during DNA lagging strand processing and long-patch base excision repair (BER). FEN1 plays a crucial role in maintaining genome stability and integrity. Several studies have also pointed out the role of FEN1 in RNA processing and RNA biogenesis. However, it remains unknown how FEN1 RNA processing can facilitate genome stability and integrity. In this study, we initially characterized the cleavage activity of FEN1 on the RNA intermediates formed during DNA lagging strand maturation and BER in an R-loop. We found that both human and yeast FEN1 efficiently cleaved RNA endonucleolytically. We demonstrated that FEN1 used its tracking mechanism to track down an RNA flap to the DNA to make a cleavage removing an RNA flap. Furthermore, we found that FEN1 efficiently cleaved RNA during BER in an R-loop. This resolved the R-loop via BER. Our study provides the first evidence that FEN1 endonucleolytic cleavage of RNA promotes the resolution of R-loops via the BER pathway, thereby maintaining genome integrity.

## INTRODUCTION

Flap endonuclease 1 (FEN1) is a central component of DNA metabolism and plays an essential role in removing RNA primers and DNA base damage during DNA lagging strand maturation and long-patch base excision repair (BER) [165, 166]. It belongs to the RAD2 structure-specific nuclease superfamily and has both endo- and exonucleolytic cleavage activity and gap endonuclease activity [53, 188-190]. During DNA replication, the DNA lagging strand is initially processed by RNase H1 that removes the RNA primers, leaving a gap with the last ribonucleotide attached to a downstream DNA strand. Subsequently, DNA polymerase  $\delta$  (pol  $\delta$ ) synthesizes nucleotides to fill in the gap and strand-displace the downstream DNA generating a 5'-flap with the ribonucleotide. FEN1 endonucleolytically cleaves the flap and generates a nick that is sealed by DNA ligase I (LIG I) [165, 166]. FEN1 can cleave a flap using a "tracking mechanism" [191] during which FEN1 is loaded from the 5'-end of the flap, tracks along the flap to the junction of between the flap and annealed region and cleaves the flap. It can also employ a "flap threading mechanism" by binding to the junction of the flap first, then threads the flap through itself and cleaves the flap [167]. The unique mechanisms of flap cleavage of FEN1 make it a critical enzyme in the maintenance of genome stability and integrity. Studies in the past decades have demonstrated that FEN1 is also the major player that prevents sequence duplication, repeat sequence expansion in mammals and telomere instability and fragility [165, 166, 192, 193]. Such roles of FEN1 have been further demonstrated by the fact that the insufficiency of FEN1 function resulting from mutations is associated with lung and



gastrointestinal cancers [194, 195] and induce a variety of mutations in multiple cancers [196, 197].

Besides its importance in DNA replication, BER, and genome maintenance, FEN1 is involved in other cellular functions by processing DNA and modulating epigenetic factors. It is implicated that FEN1 can induce apoptosis by interacting with the proteins of DNA degradosome [198]. The enzyme is also involved in the formation of covalently closed circular (cccDNA) in the hepatitis B virus [199]. FEN1 also can modify the epigenetic features in cancer cells. It has been shown that FEN1 can induce the upregulation of DNA methyltransferase 1 and 3a (DNMT1 and DNMT3a) and interact with DNMT3a through proliferating cell nuclear antigen (PCNA) [200]. This facilitates the DNA hypermethylation at the promoter region of microRNA-200a (miRNA-200a) upregulating the expression of hepatocyte growth factor (MET) and epidermal growth factor receptor (EGFR) promoting cancer progression [200].

Interestingly, it has been implicated that FEN1 also processes RNA. A recent study with FEN1 gene knockdown has suggested that FEN1 is involved in RNA processing, ribosome biogenesis, morphogenesis, cell development, and protein binding [201]. An early study from the Bambara group has shown that calf thymus FEN1 can endonucleolytically cleave an RNA flap during DNA lagging strand processing [202]. Biochemical characterization of human FEN1 cleavage activity also shows that the enzyme can make a cleavage on mRNA and rRNA molecules with lower activity than that on a DNA flap [203]. These findings support the notion that FEN1 endonuclease activity can process RNA molecules. However,

the biological significance of FEN1 RNA processing remains unknown. A recent study has suggested that FEN1 is involved in processing the RNA in R-loops formed on the telomeres of the leading strand DNA alleviating telomere fragility induced by R-loops [193]. This further suggests that FEN1 can cleave an RNA flap formed on an R-loop during DNA replication and repair.

R-loops are bulky non-B form DNA structures with an RNA:DNA hybrid and a non-template single-strand DNA (ssDNA) that are frequently generated during gene transcription [176]. R-loops can be physiological and pathological. They play an essential role in mediating immunoglobulin class switch recombination [204], CRISPR-Cas9 [205], mitochondrial DNA replication [5], and prevention of gene silencing by suppressing DNA methylation at CpG islands [187]. However, non-scheduled spontaneous accumulation of R-loops in the genome can result in a series of pathological consequences. These include replication fork stalling, blockage of mRNA synthesis by RNA polymerases, DNA damage, transcription-associated DNA recombination, and DNA repeat sequence instability, which can ultimately lead to cancer and neurodegeneration [176, 206]. To combat the adverse effects, cells have evolved the multiple mechanisms to remove unscheduled R-loops through RNase H, RNA-dependent helicases/ATPases such as senataxin (SETX) and Fanconi anemia complementation group M (FANCM), DNA topoisomerases, and DNA repair [176]. However, it remains to be elucidated how the multiple mechanisms and their coordination can lead to the removal of R-loops. In particular, little knowledge about how DNA repair is involved in the removal of R-loops. It has been proposed that the non-template ssDNAs on R-

loops are susceptible to nucleases and reactive oxygen species (ROS) [176], thus form hotspots of ssDNA breaks that can be repaired by BER. Since FEN1 can process RNAs, we further hypothesize DNA base damage that occurs on the non-template ssDNA on R-loops induces BER during which FEN1 removes the RNA and DNA flap in the R-loops through the coordination with DNA polymerase  $\beta$  (pol  $\beta$ ), thereby leading to the removal of R-loops. To test this hypothesis, we initially characterized FEN1 flap cleavage activity on RNAs. We then determined the FEN1 cleavage activity of RNA and DNA on an R-loop and its functional coordination with pol  $\beta$  during BER. We found both yeast and human FEN1 cleavage RNA flaps endonucleolytically. FEN1 also used its tracking mechanism to track down to the annealed DNA region to cleave a flap with RNA. We further demonstrated that FEN1 was recruited to R-loops in human cells upon oxidative DNA damage. We found that AP endonuclease 1 (APE1) incised the 5'-side of an abasic site on the R-loop generating a double-flap intermediate with an RNA:DNA hybrid, which inhibited DNA synthesis of pol  $\beta$ . FEN1 made cleavage on both 5'-DNA flap and a RNA flap formed on the R-loop. The coordination between FEN1 and pol  $\beta$  led to the repaired products and removal of the R-loop.

## **MATERIALS AND METHODS**

### **Materials**

Primary normal human fibroblasts and ataxia with oculomotor apraxia type 2 (AOA2) patient fibroblasts were generously provided by Dr. Kenneth Fischbeck at the National Institute of Neurological Disorders and Stroke/ National Institutes of Health. Fetal Bovine Serum (FBS) and Dulbecco's Modified Medium (DMEM) high glucose cell culture medium were from Thermo Fischer Scientific (Waltham, MA).

RNA and DNA oligonucleotides were synthesized by Integrated DNA Technologies (IDT, Coralville, IA) or Eurofins Genomics (Louisville, KY). Radionucleotides,  $\alpha$ -<sup>32</sup>P-cordycepin triphosphate (5000Ci/mmol) and  $\gamma$ -<sup>32</sup>P-adenosine triphosphate (3000 Ci/mmol) were from PerkinElmer Inc. (Boston, MA). T4 polynucleotide kinase was from New England BioLabs Inc. (Ipswich, MA). Terminal deoxynucleotidyl transferase was from Thermo Fisher Scientific (Waltham, MA). Diethyl pyrocarbonate (DEPC) was from MP Biomedicals (Santa Ana, CA). Deoxynucleosides triphosphate was from Sigma-Aldrich (St. Louis, MO). All other chemicals were from Thermo Fisher Scientific (Waltham, MA) and Sigma-Aldrich (St. Louis, MO).

### **Purification of recombinant BER enzymes**

Recombinant human APE1, pol  $\beta$ , FEN1, and DNA ligase I (LIG I) were expressed in *E. Coli* BL21(DE3) and purified using FPLC according to the procedures described [94, 95, 207]. For all recombinant BER enzymes, two liters of bacterial cell culture was induced at OD of 0.6 at 37°C or 16°C (LIG I) for 3.5 h

or 16 h (LIG I). Bacterial cells were harvested and lysed by a French Press (GlenMills, Clifton, NJ). Soluble proteins were harvested through centrifugation at 12,000 rpm for 30 min. The supernatant was loaded on the first column. The proteins were separated through chromatography through affinity or ion-exchange. After a series of sequential chromatographic purification, the peak fractions of BER proteins were pooled and dialyzed into the storage buffer, 30 mM HEPES, pH 7.5, 50 mM KCl, 20% glycerol, and 1 mM PMSF. All purified recombinant proteins were stored at -80°C.

### **Oligonucleotide substrates**

Duplex oligonucleotide substrates containing a 19 nt RNA strand were constructed by annealing a downstream strand with the RNA strand with the template strand along with an upstream strand with various lengths (29 nt-48 nt) at a molar ratio of 1:3:5. The R-loop substrate containing an abasic site (tetrahydrofuran, THF an analog of an abasic site) 36 nt RNA:DNA hybrid and an abasic site were created by annealing the DNA strand containing a THF and a 36 nt RNA strand with the template strand at a molar ratio of 4:1:3. The substrate containing a double DNA flap with a 36 nt RNA:DNA hybrid was constructed by annealing a 34 nt upstream oligonucleotide strand, the RNA strand, a 33 nt downstream DNA strand with a THF with the template strand at a ratio of 4:1:4:3. Substrates were constructed in the annealing buffer (50 mM NaCl, 1 mM EDTA, 10 mM Tris, pH 7.5) that was prepared with 0.1 DEPC-treated water and generated by denaturing at 96°C for 5 min and cool down to 25 °C. The sequences of the

substrates were listed in Table 3. Substrates were radiolabeled at the 5'-end of the RNA strand or the 5'-end of the upstream strand or the 5' or 3'-end of the downstream strand.

**Table 3. Oligonucleotide sequence**

Oligonucleotides	nt	Sequence
<b>Downstream/ Damaged strand</b>		
D3.1	38	5'-pFGAT GAC GTA AAA GGA AAG AGA CGG AAG AGG AAG AAT TC-3'
D3.2	75	5'- CTC TCG GGG CTC TGG ATT GGC CAC CCA GTC TGC CCC CFG ATG ACG TAA AAG GAA AGA GAC GGA AGA GGA AGA ATT C-3'
R3.1	39	5'-rCrGrU rArCrG rCrGrG rArArU rArCrU rUrCrG rA TA CGT AGA CTT ACT CAT TGC- 3'
R3.2	36	5'-rCrCrA rCrCrC rArGrU rCrUrG rCrCrC rCrCrG rGrArU rGrArC rGrUrA rArArA rGrGrA rArArG-3'
<b>Template strand</b>		
T3.1	60	5'-GGC AAT GAG TAA GTC TAC GTA TCG AAG TAT TCC GCG TAC GTA CGG ATG CTA GAT GAC TCG-3'
T1.2	77	5'-GGAA TTC TTC CTC TTC CGT CTC TTT CCT TTT ACG TCA TCC GGG GGC AGA CTG GGT GGC CAA TCC AGA GCC CCG AGA G - 3'
<b>Upstream strand</b>		
U1	39	5'-CGA GTC ATC TAG CAT CCG TAC GTA CGC GGA ATA CTT CGA-3'
U2	29	5'-CGA GTC ATC TAG CAT CCG TAC GTA CGC GG-3'
U3	20	5'-CGA GTC ATC TAG CAT CCG TA-3'
U4	37	5'-CTC TCG GGG CTC TGG ATT GGC CAC CCA GTC TGC CCC C-3'

**F:** Tetrahydrofuran, THF

**p:** Phosphate

**r:** RNA

## **Cleavage of RNA and BER enzymatic reactions**

Cleavage of RNA by FEN1 and APE1 and BER enzymatic activity was determined by incubating 25 nM substrates with different concentrations of FEN1 and APE1 in the presence of pol  $\beta$  and LIG I in BER reaction buffer (30 mM HEPES, pH 7.5, 50 mM KCl, 0.1 mg/ml BSA, 0.1 mM EDTA, and 0.01% NP-40). Reconstituted BER reactions were performed by incubating 25 nM substrate with various concentrations of FEN1 in the presence of 50 nM APE1, 5 nM pol  $\beta$ , and 10 nM LIG I or with various concentrations of APE1 in the presence of 25 nM FEN1, 5 nM pol  $\beta$ , and 10 nM LIG I. Reactions (20  $\mu$ l) were assembled in BER reaction buffer containing 50  $\mu$ M dNTPs, 5 mM Mg<sup>2+</sup>, and 2 mM ATP at 37 °C for 30 min. Reactions were stopped with 2x stopping buffer (95% deionized formamide and 10 mM EDTA).

## **Detection of FEN1 recruitment to R-loops upon oxidative DNA damage by immunofluorescence**

Primary normal and AOA2 fibroblasts ( $1 \times 10^5$ ) were seeded in 200  $\mu$ l DMEM medium containing 15% FBS in an 8-well chamber slide (IbiTreat, ibidi GmbH, Martinsried, Germany). Cells were treated with 10 mM potassium bromate (KBrO<sub>3</sub>) for 2 h. Untreated normal and AOA2 fibroblasts were used as control. At the end of the treatment, the fibroblasts were fixed by 4% paraformaldehyde in PBS for 30 min at 25 °C. Cells were washed with PBS three times and permeabilized with 0.1 Triton X-100 in PBS. Cells were then incubated with blocking buffer (1% BSA, 22.52 mg/ml glycine in PBST) for 30 min and subsequently with anti-RNA:DNA

hybrid antibody (1:100, Millipore-Sigma, MABE1095, Burlington, MA) and anti-FEN1 antibody (1:100, AbCam, ab17994, Cambridge, MA) in PBST containing 1% BSA at 4°C overnight. Cells were then washed with PBS three times and incubated with an anti-mouse and anti-rabbit secondary antibody conjugated with Alexa-Fluor-594 (1:1000, AbCam, ab150116) or Alexa-Fluor-488 (1:1000, AbCam, ab150077) in PBS with 1% BSA for 1 h. Cells were washed three times with PBS in the dark and incubated with 5 µg/µl DAPI in PBS for 1 min. Cell images were acquired using a Nikon microscope with a 60X objective.

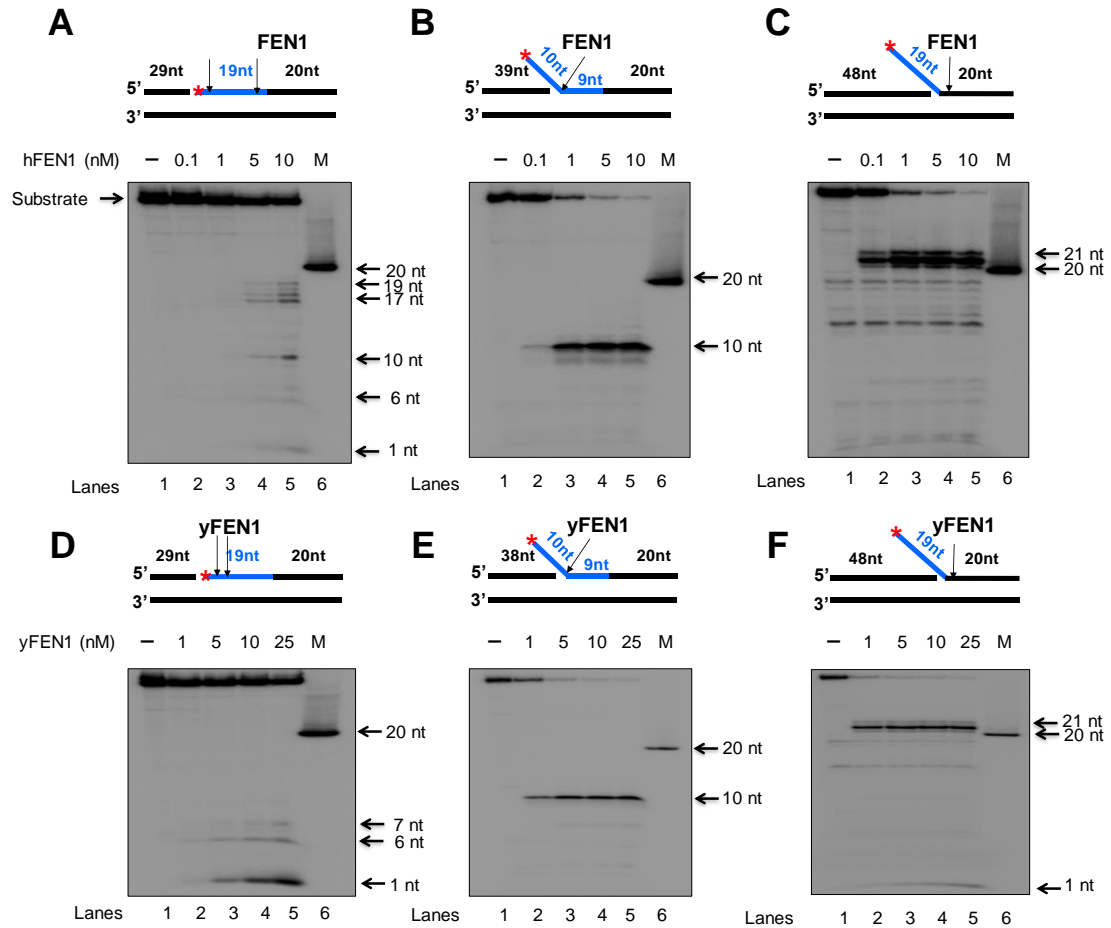
## **RESULTS**

### **FEN1 endonucleolytically cleaves RNA in the intermediates formed during DNA lagging strand processing.**

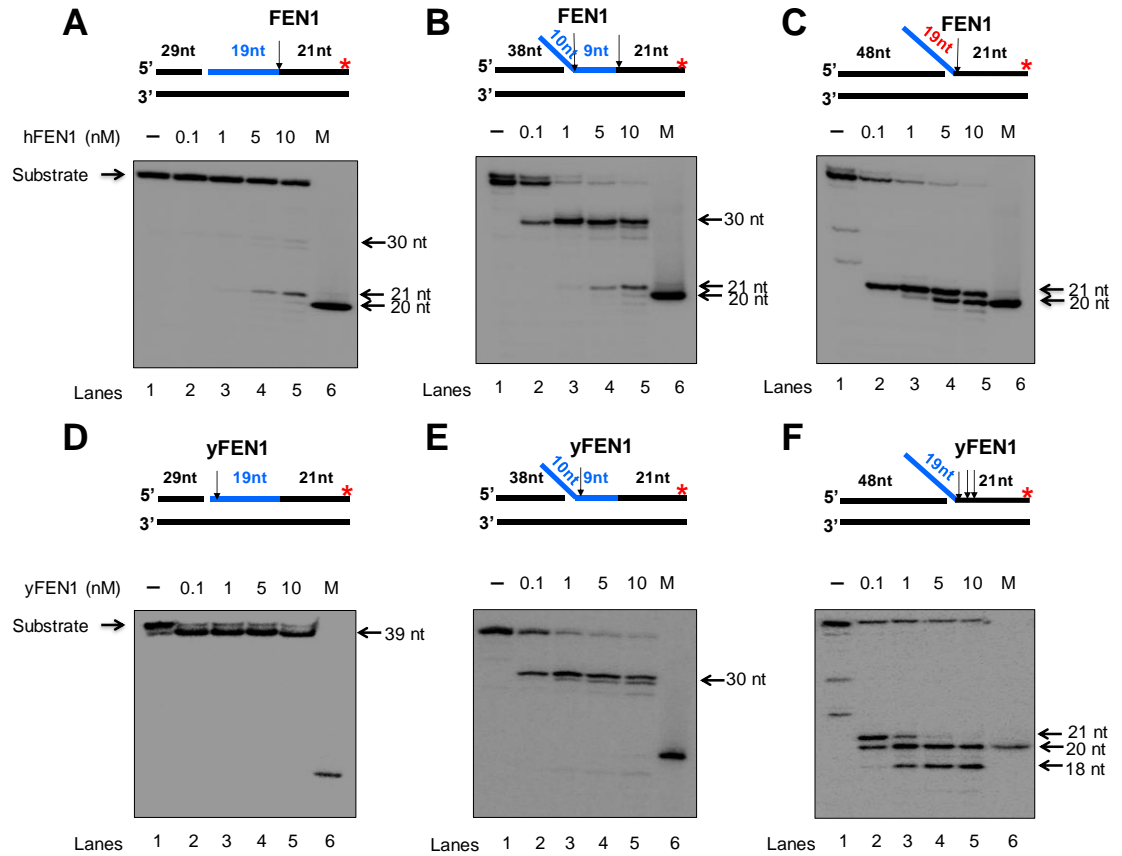
Since FEN1 can cleave an RNA flap generated on a DNA lagging strand [202], we initially characterized the endo/exonucleolytic activity of human and yeast FEN1 on a nicked RNA, a RNA flap generated within a RNA strand, and a RNA flap with a junction in between the RNA and DNA strand (Fig 3.1 and Fig 3.2). We found that both human and yeast FEN1 mainly employed its endonucleolytic activity to cleave the nicked RNA, the RNA flap within an RNA fragment, and the RNA flap with an RNA-DNA junction (Fig 3.1A-1F). Human FEN1 at 1 nM-10 nM made cleavage on the nicked RNA at multiple sites generating the products with 1 nt, 6 nt, 10 nt, and 17 nt-19 nt (Fig. 3.1A, lanes 3-5), whereas yeast FEN1 cleavage at 1 nM-25 nM resulted in the products with 1 nt, 6 nt, and 7 nt (Fig. 3.1D, lanes 2-5) indicating that FEN1 cleaved RNA. For a 10 nt-RNA flap within the RNA strand,



human and yeast FEN1 at 0.1 nM-25 nM resulted in 10 nt cleavage products (Fig. 3.1B and 1E, lanes 2-5) indicating that FEN1 made a cleavage at the junction of the RNA flap. FEN1 cleavage of a 19 nt-RNA flap attached to a DNA strand at 0.1 nM-25 nM resulted in a product with 21 nt (Fig. 3.1C and 1F, lanes 2-5). This indicates that FEN1 efficiently cleaved the RNA flap by making a cleavage within the DNA strand at the site of 2 nt from the RNA-DNA junction. We then examined if FEN1 5'-3' exonuclease also cleaved RNA during DNA lagging strand processing using the nick and flap RNA substrates radiolabeled at the 3'-end of the downstream strand (Fig. 3.2). Human FEN1 cleavage at 1 nM-10 nM on the nick RNA mainly generated the 21 nt-product along with a small amount of 30 nt-product (Fig. 3.1A, lanes 3-5) indicating that FEN1 also made the cleavage at the RNA-DNA junction. However, yeast FEN1 at 0.1 nM-10 nM only resulted in a product with 39 nt (Fig. 3.2D, lanes 2-5) indicating that the enzyme cleaved one nucleotide at the 5'-end of the RNA strand. For the 10 nt-RNA flap, human FEN1 at 0.1 nM-10 nM resulted in the products with 30 nt and 21 nt, whereas yeast FEN1 at the same concentrations mainly generated the 30 nt-cleavage products (Fig. 3.2B and 2E, lanes 2-5). The results indicate that human and yeast FEN1 efficiently cleaved the 10 nt-RNA within the RNA. Human FEN1 also made the endonucleolytic cleavage at the DNA (Fig. 3.2B, lanes 3-5) For the 19 nt RNA flap attached to the DNA strand, human FEN1 cleavage only produced 21 nt (Fig. 3.2C,



**Figure 3.1 FEN1 endo-cleavage of RNA/DNA hybrids.** To test FEN1 RNA endonuclease cleavage activity in the context of RNA/DNA hybrids, we performed FEN1 RNA cleavage assay using Okazaki fragment intermediate substrates radioactively labeled at the 5' end of the downstream strand containing a RNA and DNA on the same strand (RNA blue segments). (A) Human FEN1 RNA cleavage on nicked RNA:DNA hybrid Okazaki fragment intermediate substrate. (B) Human FEN1 RNA cleavage on RNA/DNA hybrid Okazaki fragment intermediate containing a 9nt RNA flap. (C) Human FEN1 RNA cleavage on duplex DNA containing a 19nt RNA flap. (D) Yeast FEN1 RNA cleavage on nicked RNA:DNA hybrid Okazaki fragment intermediate substrate. (E) Yeast FEN1 RNA cleavage on RNA/DNA hybrid Okazaki fragment intermediate containing a 9nt RNA flap. (F) Yeast FEN1 RNA cleavage on duplex DNA containing a 19nt RNA flap. Lane 1 represents substrate only. Lanes 2-6 represent increasing concentrations of FEN1. Substrates and products were separated in a 15% urea-denaturing polyacrylamide gel and detected by phosphorimager. All experiments were performed in triplicates.



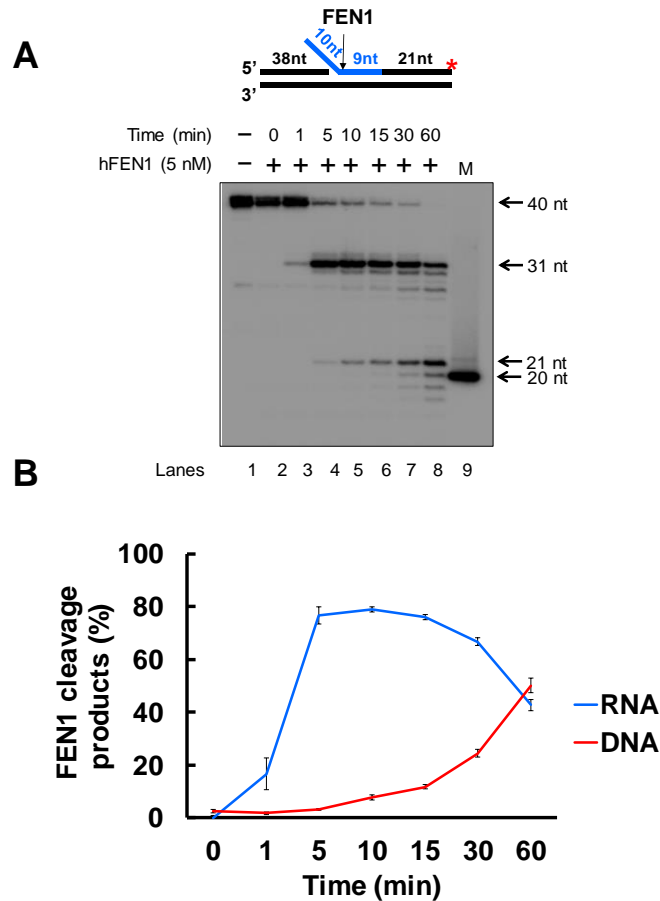
**Figure 3.2 FEN1** **exo-cleavage of RNA/DNA hybrids.** To test FEN1 RNA exonuclease cleavage activity in the context of RNA/DNA hybrids, we performed FEN1 RNA cleavage assay using Okazaki fragment intermediate substrates radioactively P32 labeled at the 3' end of downstream strand containing a RNA and DNA on the same strand (RNA blue segments). (A) Human FEN1 RNA cleavage on nicked RNA:DNA hybrid Okazaki fragment intermediate substrate. (B) Human FEN1 RNA cleavage on RNA/DNA hybrid Okazaki fragment intermediate containing a 9nt RNA flap. (C) Human FEN1 RNA cleavage on duplex DNA containing a 19nt RNA flap. (D) Yeast FEN1 RNA cleavage on nicked RNA:DNA hybrid Okazaki fragment intermediate substrate. (E) Yeast FEN1 RNA cleavage on RNA/DNA hybrid Okazaki fragment intermediate containing a 9nt RNA flap. (F) Yeast FEN1 RNA cleavage on duplex DNA containing a 19nt RNA flap. Lane 1 represents substrate only. Lanes 2-6 represent increasing concentrations of FEN1. Substrates and products were separated in a 15% urea-denaturing polyacrylamide gel and detected by phosphorimager. All experiments were done in triplicates.

lanes 2-5) indicating that the enzyme made a cleavage at the junction between the 19 nt flap. However, yeast FEN1 cleavage on the flap resulted in the products with 21 nt, 20 nt, and 18 nt (Figure 3.2F, lanes 2-5). The results indicate that human and yeast FEN1 exerted a different cleavage pattern by making cleavages at distinct sites in the RNA strand.

### **FEN1 can track down to the DNA region to remove an RNA primer**

Since FEN1 also exhibits a 5'-3'-exonuclease activity, we then asked if FEN1 also used its 5'-3'-exonuclease activity to continue to remove the RNA strand after the enzyme cleaved an RNA flap. We tested this possibility by examining human FEN1 cleavage activity on the substrate containing the 10 nt-RNA flaps with a 9 nt-RNA annealed on the template strand at various time intervals (Fig. 3.3). The substrate was radiolabeled at the 3'-end of the downstream strand for detecting the FEN1 exonucleolytic cleavage products. The results showed that FEN1 endonucleolytically cleaved the 10 nt-RNA flaps leaving a 31 nt-cleavage product at the time of intervals of 1-60 min (Fig. 3.3, lanes 1-8). Starting from 5 min, a 21 nt-cleavage product was generated (Fig. 3.3, lanes 4-8) indicating that FEN1 made the cleavage at the junction between the RNA and DNA strand after FEN1 removed the RNA flap. Moreover, we found that FEN1 exhibited its 5'-3' exonuclease activity within the annealed DNA strand further demonstrating that FEN1 used its endonucleolytic activity to remove the RNA. Quantification of the FEN1 cleavage products showed that FEN1 endonucleolytic cleavage activity removed the 10 nt-RNA flaps within 15 min along with a significant increase of the

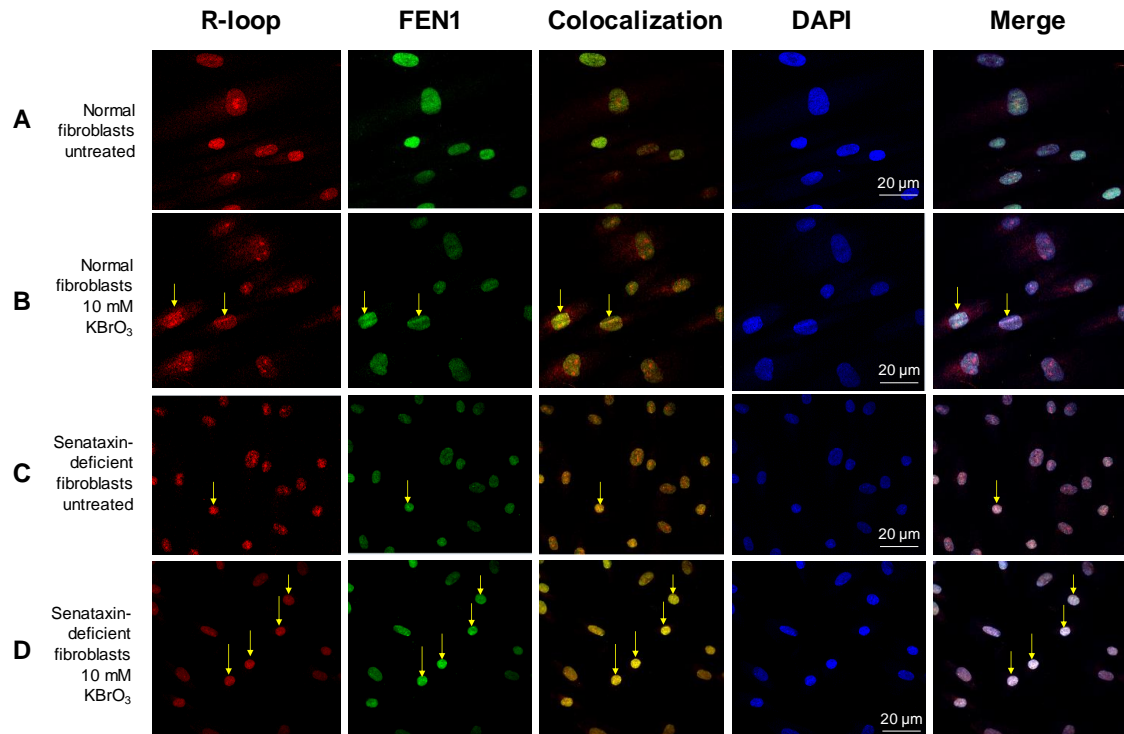
endonucleolytic cleavage at the junction between RNA and DNA (Fig 3.3B). This suggests that FEN1 removed the RNA flap leaving a nicked RNA for the enzyme to track down to the junction between RNA and DNA and made the endonucleolytic cleavage and removed the RNA strand.



**Figure 3.3 FEN1 cleavage of RNA and DNA tracking in an RNA/DNA hybrid.** The rate of FEN1 RNA exonuclease cleavage activity in the context of RNA/DNA hybrids was performed by a FEN1 RNA cleavage assay using Okazaki fragment intermediate at different time intervals. (A) shows 5 nM of FEN1 cleavage at increasing time points were performed by using 25nM of  $^{32}\text{P}$  labeled substrate at the 3' end of the downstream strand containing an RNA and DNA on the same strand (RNA blue segments). Lane 1 represents substrate only. Lane 2 represents FEN1 at 0 minutes and lanes 2 to 8 represents increasing incubation time. Lane 9 is a 20 nt size marker. (B) Quantification results from A. Blue line indicates FEN1 RNA cleavage segment and red line indicates FEN1 cleavage segments. Substrates and products were separated in a 15% urea-denaturing polyacrylamide gel and detected by phosphorimager. All experiments were performed in triplicates.

## **FEN1 is recruited to R-loops in human cells upon oxidative DNA damage and cleaves the RNA in an R-loop during BER**

Since a recent study has implicated a role of FEN1 in processing R-loops formed in telomeres [193], and the non-template strand of an R-loop is susceptible to DNA base damage, we asked if FEN1 can cleave the RNA in R-loops formed during gene transcription facilitating the resolution of R-loops during BER. We initially examined the recruitment of FEN1 to R-loops and its dependency on oxidative DNA damage in normal human fibroblasts and senataxin gene-deficient, Ataxia oculomotor apraxia type 2 (AOA2) fibroblasts using immunofluorescence (Fig. 3.4). The results showed that in untreated normal fibroblasts, R-loops (red) and FEN1 (green) were detected in the nucleus of the fibroblasts. However, a little signal of the co-localization of R-loop and FEN1 was detected (Fig. 3.4A). Upon the treatment of 10 mM KBrO<sub>3</sub> that specifically induces 8-oxoguanines (8-oxoGs) [208], the colocalization of R-loop and FEN1 was detected in a few numbers of the nucleus (Fig 3.4B). For the untreated AOA2 fibroblasts, which usually exhibit the accumulation of R-loops because of the deficiency of the senataxin helicase activity [181], the R-loop-FEN1 colocalized was detected in a few nuclei (Fig. 3.4C). Upon the treatment of KBrO<sub>3</sub>, the signal of the R-loop-FEN1 colocalization was significantly increased in AOA2 fibroblasts (Fig. 3.4D). The results indicate that FEN1 was recruited to R-loops upon oxidative DNA damage suggesting that FEN1 processed the RNA in an R-loops during BER of oxidative DNA damage in AOA2 fibroblasts. We then examined FEN1 cleavage activity on the RNA on an R-loop containing a 36 nt RNA:DNA hybrid and an abasic site in the middle of the

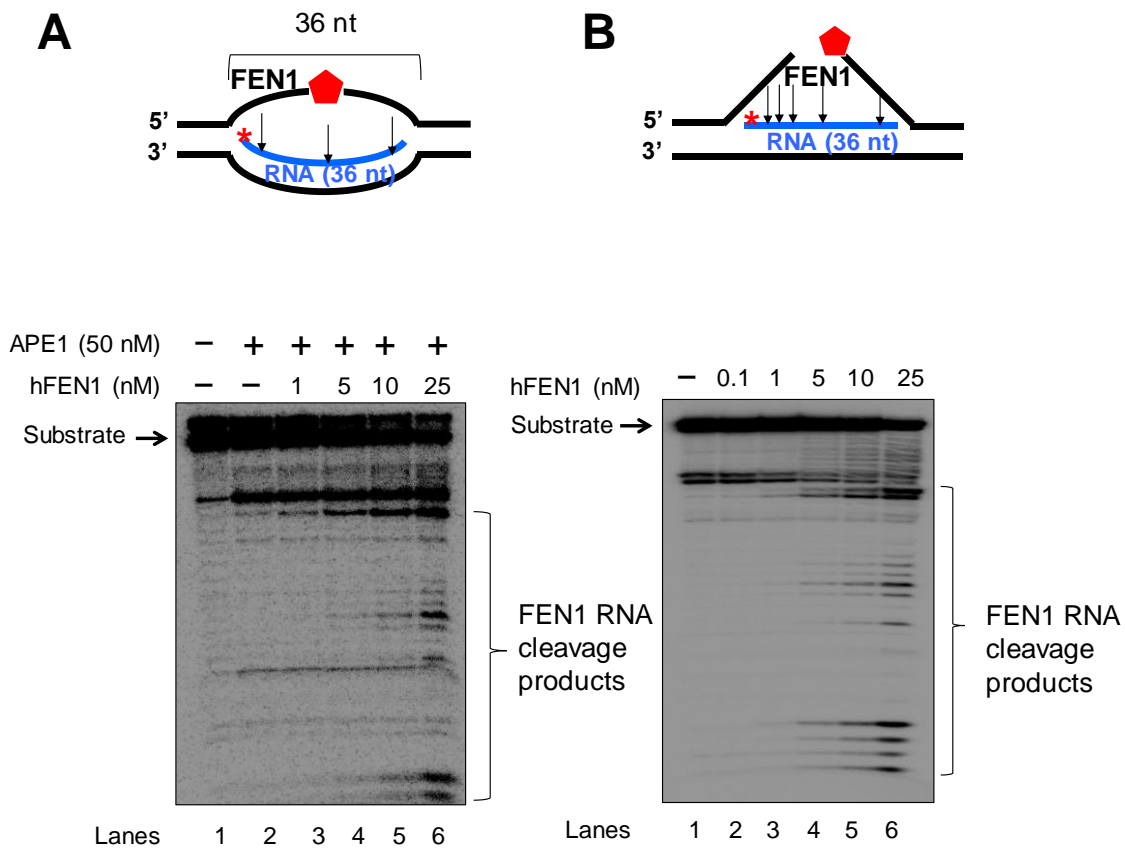


**Figure 3. 4 FEN1 recruitment to R-loops upon oxidative DNA damage.** Primary normal fibroblast cells were either left untreated as control (A) or treated with 10 mM  $\text{KBrO}_3$  for 2 hours. (B). AOA2 fibroblast cells were either left untreated as control (C) or treated with 10 mM  $\text{KBrO}_3$  for 2 hours. (D). Anti-RNA:DNA hybrid antibody (S9.6) were used to detect R-loop accumulation on human cells. Anti-FEN1 antibody was used to detect the colocalization of FEN1 to the R-loop. Anti-rabbit secondary antibody conjugated with Alexa-Fluor-594 were used for imaging detection. Cell images were acquired using a Nikon microscope with a 60X objective. The image exposures were adjusted to show the presence of foci.

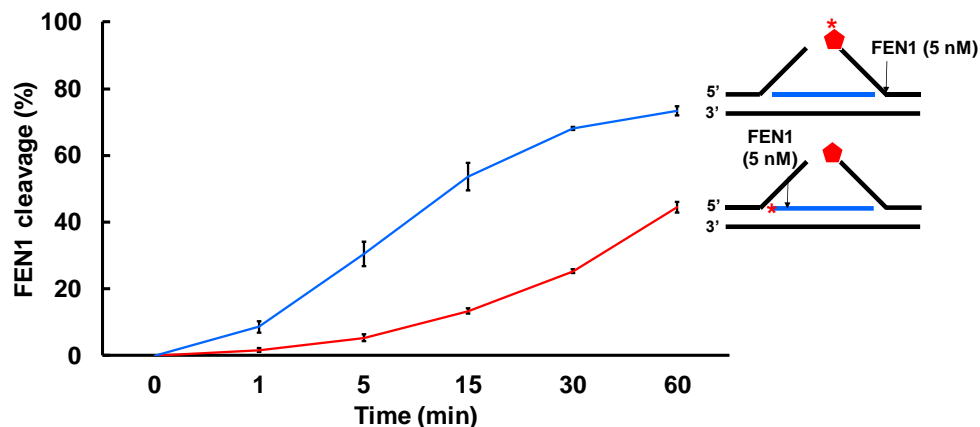
non-template (Fig. 3.5). FEN1 cleavage of the RNA was determined with the substrate with an intact abasic site (THF) or the substrate that mimics the intermediate that is pre-cut by APE1 containing a 3'- and 5'-DNA flaps with an RNA:DNA hybrid (Fig. 3.5). We found that FEN1 at 1 nM-25 nM endonucleolytically cleaved the RNA strand from the R-loop substrate with an abasic site incised by 50 nM APE1 (Fig. 3.5A, lanes 3-6). Similarly, FEN1 at the same concentrations endonucleolytically cleaved the RNA strand on the double-flap intermediate (Fig.

3.5B, lanes 3-6). The enzyme exhibited a unique flap cleavage pattern by initially cleaving 2 to 4 ribonucleotides and subsequently removing a series of larger RNA oligonucleotides (Fig. 3.5A and 5B, lanes 5-6). This further suggests that FEN1 also tracked along with the RNA strand and made an endonucleolytic cleavage of the RNA in an R-loop during BER. We then compared the activity for FEN1 to remove the downstream 5'-DNA flap with its cleavage activity on the RNA in the double-flap substrate with an RNA:DNA hybrid (Fig. 3.6). We found that FEN1 cleavage of the 5'-flap is much faster than its cleavage (Fig. 3.5, compare the blue line with the red line). For example, at 1 min, FEN1 at 5 nM resulted in 10% of DNA flap cleavage products but little RNA cleavage product. At 5 min, FEN1 generated 30% of DNA flap cleavage products but only 5% of RNA cleavage products. At 15 min, FEN1 produced 50% of DNA flap cleavage products and 10% of RNA cleavage products (Fig. 3.6). The results indicate that FEN1 cleaved the 5'-DNA flaps much faster than the RNA suggesting that FEN1 cleaved the 5'-DNA flap first and then removed the RNA during BER in the R-loop.





**Figure 3.5 FEN1 cleavage of RNA on R-loops through BER.** The ability of FEN1 cleavage on RNA on R-loops was examined by titrating increasing concentrations of FEN1 on R-loop and R-loop intermediate substrate. 25 nM of 5' <sup>32</sup>P label RNA substrate was incubated at 37 °C for 15 min. (A) R-loop substrate was cleaved by using 50 nM APE1 to generate the strand break. The activity of FEN1 on RNA was measured by using increasing concentration of FEN1 (1 nM-25 nM). Lane 1 represents substrate only. Lane 2 represents APE1 cleavage of R-loops. Lanes 3-6 represent increasing concentration of FEN1. (B) R-loop intermediate substrate was used to measure the cleavage of RNA by FEN1 (0.1-25 nM). Lane 1 represents substrate only. Lanes 2-6 represent increasing concentration of FEN1. Substrates and products were separated in a 15% urea-denaturing polyacrylamide gel and detected by phosphorimager. All experiments were done in triplicates.

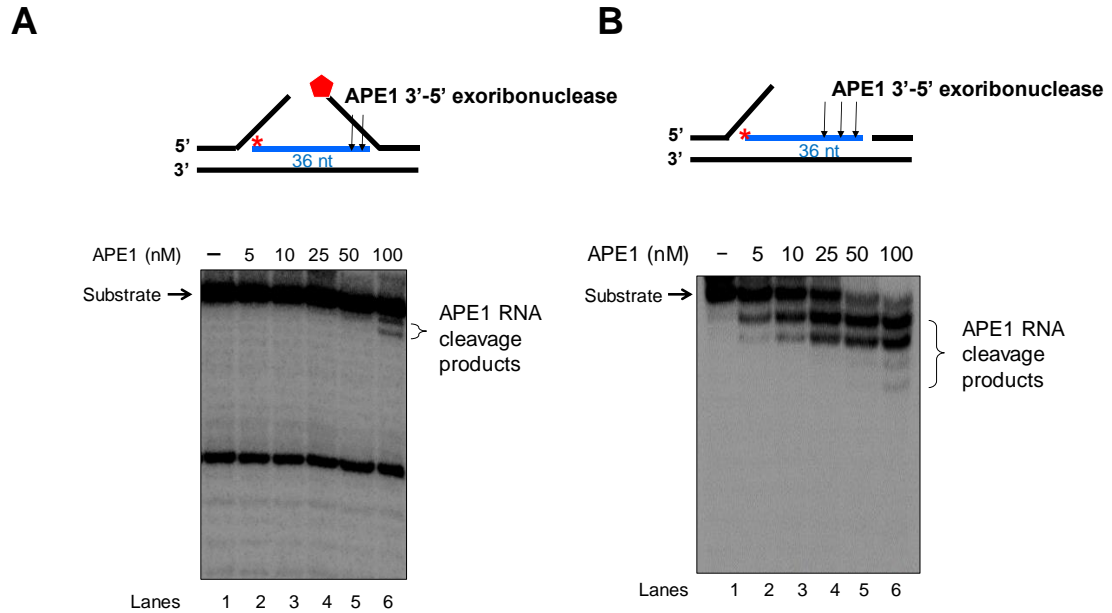


**Figure 3.6 FEN1 rate of cleavage between flap and RNA R-loops intermediate.** The comparison between the rate of FEN1 in cleaving 5' DNA flaps and RNA was examined by labeling the upstream DNA flap and the RNA on the R-loop intermediate and incubating them with 5 nM FEN1 with increasing incubation time at 37 °C. Blue line represents the rate of FEN1 cleavage on the 5' DNA flap of the R-loop intermediate. Red line represents the rate of FEN1 cleavage of the RNA on the R-loop intermediate. All experiments were performed in triplicates.

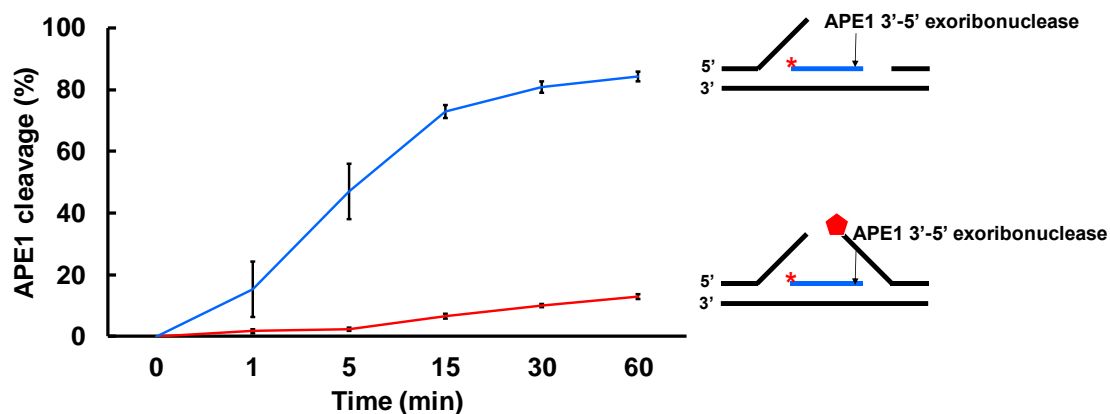
### **APE1 3'-5' exonuclease cleaves RNA during BER in an R-loop**

As one of the key BER enzymes, APE1 can incise the 5'-end of an abasic site using its endonucleolytic cleavage activity creating a 1 nt-gap as the substrate for pol  $\beta$  to fill in [73]. APE1 also possesses a 3'-5' exonuclease activity [73] that can remove a mismatched nucleotide at the 3'-end of the upstream strand of a one-nucleotide gap [209]. APE1 can also make endo- and exonucleolytic cleavage on RNA [210-212]. Thus, we reasoned that APE1 can also cleave the RNA in an R-loop during BER. To test this, we examined the APE1 cleavage activity on the RNA of the double-flap substrate with a 36 nt-RNA and the 3'-flap substrates with the RNA that mimics an intermediate with its 5'-flap precleaved by FEN1 (Fig. 3.7).

We found that APE1 at 100 nM used its 3'-5' exonuclease removed 1 and 2 ribonucleotides from the 3'-end of the RNA strand on the double-flap substrate (Fig. 7A, lane 6). APE1 at the concentrations lower than 100 nM failed to make any cleavage on the RNA (Fig. 3.7A, lanes 2-5). However, we found that APE1 exonuclease made an efficient cleavage on ribonucleotides from the 3'-end of the RNA on the substrate containing a 3'-flap (Fig. 3.7B). The results indicate that APE1 3'-5'-exonuclease progressively cleaved the ribonucleotides from the 3'-end of the RNA in an R-loop during BER, and APE1 cleavage of the RNA on the R-loop was stimulated by the removal of the downstream 5'-DNA flap. This suggests that APE1 cleaved the RNA after FEN1 removes the 5'-DNA flaps. To further test this possibility, we compared the APE1 exoribonuclease activity on the RNA in the double-flap substrate with its exoribonuclease activity on the RNA in the substrate with a 3'-flap at various time intervals (Fig. 3.8). We found that APE1 exhibited inefficient 3'-5' exoribonuclease activity on the RNA strand on the double flap substrate and only produced 10% cleavage products in 60 min (Fig. 3.8, red line). APE1 exoribonuclease efficiently cleaved the RNA in the substrate with the 5'-flap removed (Fig. 3.8, blue line). This further indicates that APE1 3'-5' exoribonuclease cleavage of the RNA on an R-loop is dependent on the removal of the 5'-DNA flaps by FEN1 suggesting the coordination between FEN1 DNA flap cleavage activity and APE1 exoribonuclease in processing the RNA during BER in an R-loop.



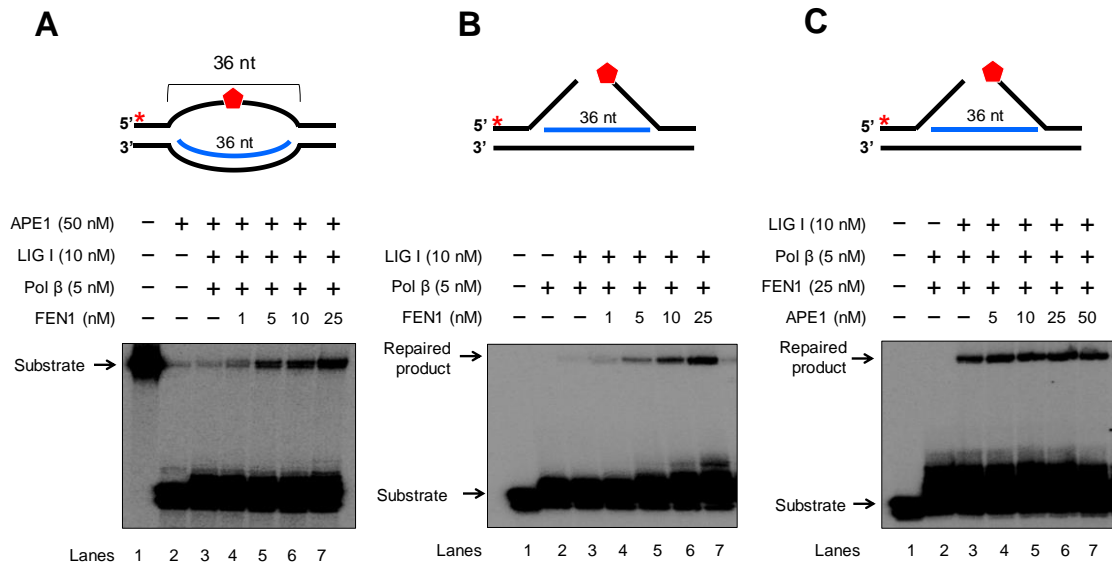
**Figure 3.7 APE1 cleavage of RNA on R-loops intermediates through BER.** The ability of APE1 cleavage on RNA on R-loops was examined by titrating increasing concentrations of APE1 on R-loop intermediates and R-loop intermediate without a 5' flap substrate. 25 nM of 5' <sup>32</sup>P label RNA substrate was incubated at 37 °C for 15 min. (A) The activity of APE1 cleavage of RNA was measured by using increasing concentration of APE1 (5 nM-100 nM) on the R-loop intermediate. Lane 1 represents substrate only. Lane 2-6 represents increasing concentration of APE1. (B) The activity of APE1 cleavage of RNA was measured by using increasing concentration of APE1 (5 nM-100 nM) on the R-loop intermediate without a 5' flap. Lane 1 represents substrate only. Lane 2-6 represents increasing concentration of APE1. Substrates and products were separated in a 15% urea-denaturing polyacrylamide gel and detected by phosphorimager. The results show that APE1 can efficiently cleave RNA when the upstream strand has been removed. All experiments were done in triplicates.



**Figure 3.8 APE1 rate of cleavage of RNA on R-loops.** The comparison between the rate of APE1 to cleave on RNA on R-loop intermediate with and without a 5' flap was examined by labeling the RNA with  $^{32}\text{P}$  and incubating them with 25 nM APE1 with increasing incubation time at 37 °C. Blue line represents the rate of APE1 cleavage on the R-loop intermediate substrate without the 5' flap. Red line represents the rate of APE1 cleavage of the RNA on the R-loop intermediate containing a 5' flap. The results show that the upstream flap on R-loop intermediate block APE1 cleavage on RNA. All experiments were performed in triplicates.

### FEN1 and APE1 cleavage of RNA facilitates BER in an R-loop

To further determine if FEN1 and APE1 cleavage on RNA can facilitate BER leading to the resolution of an R-loop, we reconstituted BER on the R-loop substrates in the increasing concentrations of FEN1 or APE1 (Fig. 3.9). We found that increasing concentrations of FEN1 (1 nM-25 nM) led to the increasing amount of BER repaired product with the presence of 5 nM pol  $\beta$ , 10 nM LIG I with or without 50 nM APE1 (Fig. 3.9A and 3.9B, lanes 4-7). Increasing concentrations of APE1 (5 nM-50 nM) in the presence of 25 nM FEN1 led to a small increase of BER repaired product (Fig. 3.9C, lanes 4-7). The results indicate that FEN1 and APE1 cleavage of the RNA in an R-loop facilitated BER leading to the resolution of the R-loop.

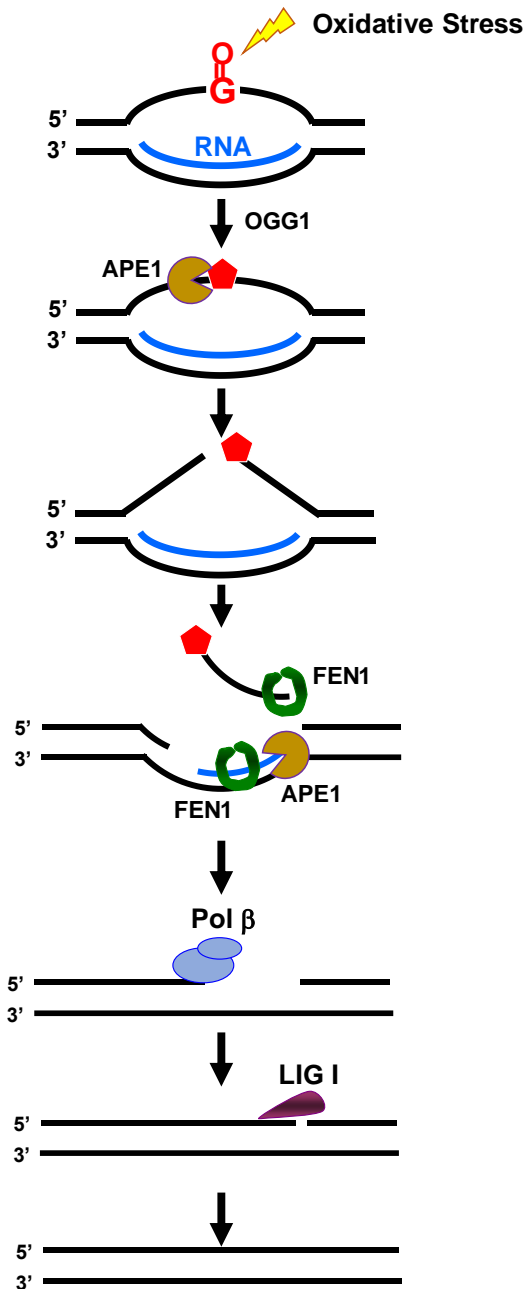


**Figure 3.9 FEN1 and APE1 promotes the resolution of R-loops during BER.** The role of FEN1 and APE1 in resolving R-loop accumulation through BER was determined by reconstituting BER on 25 nM of R-loop and R-loop intermediate incubated at 37 °C for 30 min. (A) BER reconstitution on R-loop containing a THF as a damage. Lane 1 represents substrate only. Lane 2 represents APE1 cleavage on R-loop. Lane 3-7 shows the effect of FEN1 at increased concentrations to promote the resolution of the R-loop and repair of the lesion. (B) BER reconstitution in the absence of APE1 on an R-loop intermediate. Lane 1 represents substrate only. Lane 2 represents pol β synthesis. Lanes 3-7 shows the effect of FEN1 at increasing concentrations to promote the repair of the damage and R-loop removal. (C) BER reconstitution on R-loop intermediate with increasing concentration of APE1. Lane 1 represents substrate only. Lane 2 shows pol β synthesis in the presence of FEN1. Lane 3 shows repaired products of FEN1, pol β, and LIG I. Lanes 4-7 shows repaired products during BER in the presence of increasing concentration of APE1. Substrates and products were separated in a 15% urea-denaturing polyacrylamide gel and detected by phosphorimager. All experiments were performed in triplicates.

## DISCUSSIONS

In this study, we characterized FEN1 flap cleavage on the RNA produced in the DNA lagging strand and R-loop. We further determined the cleavage of RNA by both FEN1 and APE1 during BER in an R-loop and its effect on the resolution of the R-loop. We found that human and yeast FEN1 endonucleolytically cleaved a nicked RNA and an RNA flap attached to an RNA or DNA strand efficiently (Fig.

3.1 and 3.2). We further demonstrated that FEN1 removed an RNA flap and subsequently tracked down to the DNA strand and made endonucleolytic cleavage (Fig. 3.3). We then discovered that FEN1 was recruited to R-loops in normal human fibroblasts, and senataxin-deficient fibroblasts upon oxidative DNA damage, 8-oxoG induced by  $\text{KBrO}_3$  (Fig. 3.4). Further analysis on FEN1 cleavage on the RNA during BER in an R-loop showed that FEN1 cleaved the downstream 5'-DNA flap and then endonucleolytically cleaved the RNA strand (Fig. 3.5 and 3.6). Moreover, we found that APE1 exhibited an efficient 3'-5' exoribonuclease activity to cleave the RNA strand in an R-loop from its 3'-end through the coordination with FEN1 cleavage of the downstream 5'-DNA flap (Fig. 3.7 and 3.8). We showed that the RNA cleavage activity of FEN1 and APE1 during BER in an R-loop facilitated the production of the BER repaired product (Fig. 3.9). Our results supported a model during which oxidative stress induces oxidative DNA damage such as 8-oxoGs on the non-template strand of an R-loop. 8-oxoG DNA glycosylase 1 (OGG1) removes 8-oxoGs and generates abasic sites on the non-template strand. APE1 incises the AP sites and the loop converting the R-loop into a double-flap intermediate with a RNA:DNA hybrid. FEN1 removes the downstream 5'-DNA flap leaving an intermediate with a 3'-DNA flap and an RNA:DNA hybrid. Subsequently, FEN1 uses its endonucleolytic cleavage activity to remove the RNA from its 5'-end, whereas APE1 uses its 3'-5' exoribonuclease activity to remove the RNA from its 3'-end. The removal of the RNA leaves a DNA gap, which is filled by pol  $\beta$  DNA synthesis. Finally, LIG I seals the nick generated by pol  $\beta$  DNA synthesis (Fig. 10).



**Figure 3.10 A hypothetical of R-loops mitigation by BER.** R-loops accumulation can be mitigated by BER through oxidative DNA damage which targets the single stranded region of the R-loop. The damage base is removed by a damage specific DNA glycosylase leaving an abasic site that is incised by APE1 at the 5'-end to generate the strand break. Subsequently, the APE1 cleavage generates R-loop intermediate containing a double flap. The 5' downstream flap will be removed by FEN1 to allow the access of APE1 cleavage 3' to 5' of the RNA and FEN1 will cleave 5' to 3' of the RNA. The removal of the RNA will facilitate the reannealing of the upstream strand and pol β and LIG I will complete the repair of the damage. Thereby, leading to the resolution of the R-loop.



Since the Bambara group has reported that calf thymus FEN1 can cleave an RNA flap [202] indicating that mammalian FEN1 also possesses a ribonuclease activity, the work by Stevens also has shown that human FEN1 can endonucleolytically cleave mRNA and rRNA at the junction of RNA flaps, RNA hairpins, and bubbles [203]. A recent study implicates that FEN1 is involved in processing RNA:DNA hybrid to prevent telomere fragility during DNA leading strand synthesis [193] suggesting a role of FEN1 in processing RNA in telomeres. However, the mechanisms and function of FEN1 ribonuclease activity and its coordination with DNA flap cleavage activity remains unknown. Here we provide the first evidence that both human and yeast FEN1 use their ribonuclease activity to cleave RNA endonucleolytically (Fig. 3.1 and 3.2). Interestingly, we found that FEN1 removed an RNA flap efficiently and then tracked down to the junction between the RNA and DNA or the DNA region to make an endonucleolytic cleavage (Fig. 3.1 and 3.3) indicating that FEN1 preferentially makes a cleavage on DNA. Furthermore, for the first time we demonstrated that FEN1 RNA cleavage activity was employed to process the RNA strand during BER of an R-loop in human cells (Fig. 3.4 and 3.6). We showed that FEN1 5'-endoribonuclease coordinated with APE1 3'-5' exoribonuclease to remove the RNA strand in an R-loop during BER of a base lesion on the non-template strand facilitating BER and leading the resolution of the R-loop (Fig. 3.7 and 3.9). Our results indicate that FEN1 employs a unique mechanism to cleave RNAs leading to resolution of R-loops through BER. The roles of FEN1 in RNA metabolism and miRNA biogenesis and genome maintenance need to be elucidated in the future.

The cleavage of RNA by APE1 was first identified by the Hickson group in the 1990s [210]. Later studies have shown that APE1 endoribonuclease can incise the 5'-end of an abasic site in mRNAs and RNA at a site-specific manner leading to their removal [213, 214]. In this study, for the first time, we found that APE1 3'-5' exoribonuclease but not endoribonuclease activity progressively cleaved the RNA strand during BER in an R-loop (Fig. 3.7 and 3.8). We showed that the exoribonuclease activity of APE1 was dependent on the removal of the downstream 5'-DNA flap by FEN1 (Fig. 3.7 and 3.8), demonstrating functional coordination between the two BER enzymes in removing the RNA strand during BER in an R-loop. It is possible that the 5'-DNA flap in the double-flap intermediate formed during BER in an R-loop inhibited the binding of APE1 to the 3'-end of the RNA strand, thereby preventing the enzyme from making a ribonucleotide cleavage.

Our study further indicates that BER can serve as a new pathway to resolve R-loops through FEN1 endonucleolytic cleavage of DNA and RNA flaps and its coordination with APE1 3'-5' exoribonuclease cleavage of the RNA strand (Fig. 3.10). It appears that the non-template strand of an R-loop forms hotspots of oxidative DNA damage such as 8-oxoGs that are subsequently repaired by BER. This, in turn, results in the ssDNA breaks on the non-template strand leading to the opening of the R-loop converting it into a double-flap intermediate with an RNA:DNA hybrid that can be further processed by FEN1 DNA flap cleavage and endoribonuclease activity along with APE1 3'-5' exoribonuclease activity. It has been shown that other BER enzymes and cofactors, including OGG1 and NEIL2

DNA glycosylases, X-ray repair cross-complementing protein 1 (XRCC1), and poly(ADP-ribose) polymerase 1 (PARP1) are also involved in RNA metabolism and quality control. Thus, it is conceivable that these BER enzymes and cofactors can cooperate with FEN1 and APE1, leading to the efficient removal of the RNA during BER in R-loops. The molecular mechanisms underlying the resolution of R-loops through the BER pathway remains to be elucidated.

## REFERENCES

1. Helena, J.M., et al., Deoxyribonucleic Acid Damage and Repair: Capitalizing on Our Understanding of the Mechanisms of Maintaining Genomic Integrity for Therapeutic Purposes. *Int J Mol Sci*, 2018. 19(4).
2. Hamperl, S. and K.A. Cimprich, The contribution of co-transcriptional RNA:DNA hybrid structures to DNA damage and genome instability. *DNA Repair (Amst)*, 2014. 19: p. 84-94.
3. Lindahl, T., The Croonian Lecture, 1996: endogenous damage to DNA. *Philos Trans R Soc Lond B Biol Sci*, 1996. 351(1347): p. 1529-38.
4. Liu, Y. and S.H. Wilson, DNA base excision repair: a mechanism of trinucleotide repeat expansion. *Trends Biochem Sci*, 2012. 37(4): p. 162-72.
5. Aguilera, A. and T. Garcia-Muse, R loops: from transcription byproducts to threats to genome stability. *Mol Cell*, 2012. 46(2): p. 115-24.
6. Jacobs, A.L. and P. Schar, DNA glycosylases: in DNA repair and beyond. *Chromosoma*, 2012. 121(1): p. 1-20.
7. Lindahl, T. and B. Nyberg, Rate of depurination of native deoxyribonucleic acid. *Biochemistry*, 1972. 11(19): p. 3610-8.
8. Lindahl, T., Instability and decay of the primary structure of DNA. *Nature*, 1993. 362(6422): p. 709-15.
9. Shapiro, R.W. and M.B. Keller, Initial 6-month follow-up of patients with major depressive disorder. A preliminary report from the NIMH collaborative study of the psychobiology of depression. *J Affect Disord*, 1981. 3(3): p. 205-20.
10. Lindahl, T. and B. Nyberg, Heat-induced deamination of cytosine residues in deoxyribonucleic acid. *Biochemistry*, 1974. 13(16): p. 3405-10.
11. Gaschler, M.M. and B.R. Stockwell, Lipid peroxidation in cell death. *Biochem Biophys Res Commun*, 2017. 482(3): p. 419-425.
12. Cadet, J., et al., Oxidative damage to DNA: formation, measurement, and biological significance. *Rev Physiol Biochem Pharmacol*, 1997. 131: p. 1-87.
13. Chaudhary, A.K., et al., Detection of endogenous malondialdehyde-deoxyguanosine adducts in human liver. *Science*, 1994. 265(5178): p. 1580-2.

14. Nath, R.G., J.E. Ocampo, and F.L. Chung, Detection of 1, N<sup>2</sup>-propanodeoxyguanosine adducts as potential endogenous DNA lesions in rodent and human tissues. *Cancer Res*, 1996. 56(3): p. 452-6.
15. Kuraoka, I., et al., Removal of oxygen free-radical-induced 5',8-purine cyclodeoxynucleosides from DNA by the nucleotide excision-repair pathway in human cells. *Proc Natl Acad Sci U S A*, 2000. 97(8): p. 3832-7.
16. Szczepanski, J.T., et al., Double-strand break formation during nucleotide excision repair of a DNA interstrand cross-link. *Biochemistry*, 2009. 48(32): p. 7565-7.
17. Mato, J.M., M.L. Martinez-Chantar, and S.C. Lu, S-adenosylmethionine metabolism and liver disease. *Ann Hepatol*, 2013. 12(2): p. 183-9.
18. Ding, W., et al., s-Adenosylmethionine Levels Govern Innate Immunity through Distinct Methylation-Dependent Pathways. *Cell Metab*, 2015. 22(4): p. 633-45.
19. Rydberg, B. and T. Lindahl, Nonenzymatic methylation of DNA by the intracellular methyl group donor S-adenosyl-L-methionine is a potentially mutagenic reaction. *EMBO J*, 1982. 1(2): p. 211-6.
20. Grollman, A.P. and M. Moriya, Mutagenesis by 8-oxoguanine: an enemy within. *Trends Genet*, 1993. 9(7): p. 246-9.
21. Boysen, G., et al., The formation and biological significance of N<sup>7</sup>-guanine adducts. *Mutat Res*, 2009. 678(2): p. 76-94.
22. Coulondre, C. and J.H. Miller, Genetic studies of the lac repressor. IV. Mutagenic specificity in the lacI gene of *Escherichia coli*. *J Mol Biol*, 1977. 117(3): p. 577-606.
23. Lindahl, T., DNA glycosylases, endonucleases for apurinic/apyrimidinic sites, and base excision-repair. *Prog Nucleic Acid Res Mol Biol*, 1979. 22: p. 135-92.
24. Kondo, N., et al., DNA damage induced by alkylating agents and repair pathways. *J Nucleic Acids*, 2010. 2010: p. 543531.
25. Yang, Y., D.A. Gordenin, and M.A. Resnick, A single-strand specific lesion drives MMS-induced hyper-mutability at a double-strand break in yeast. *DNA Repair (Amst)*, 2010. 9(8): p. 914-21.
26. Lubin, D. and E.H. Jensen, Effects of Clouds and Stratospheric Ozone Depletion on Ultraviolet-Radiation Trends. *Nature*, 1995. 377(6551): p. 710-713.

27. Alscher, R.G., J.L. Donahue, and C.L. Cramer, Reactive oxygen species and antioxidants: Relationships in green cells. *Physiologia Plantarum*, 1997. 100(2): p. 224-233.
28. Leng, J. and Y. Wang, Liquid Chromatography-Tandem Mass Spectrometry for the Quantification of Tobacco-Specific Nitrosamine-Induced DNA Adducts in Mammalian Cells. *Anal Chem*, 2017. 89(17): p. 9124-9130.
29. Mehta, A. and J.E. Haber, Sources of DNA double-strand breaks and models of recombinational DNA repair. *Cold Spring Harb Perspect Biol*, 2014. 6(9): p. a016428.
30. Xie, Z.M., et al., Ultraviolet-B exposure induces photo-oxidative damage and subsequent repair strategies in a desert cyanobacterium *Microcoleus vaginatus* Gom. *European Journal of Soil Biology*, 2009. 45(4): p. 377-382.
31. Hwang, B.J., G. Shi, and A.L. Lu, Mammalian MutY homolog (MYH or MUTYH) protects cells from oxidative DNA damage. *DNA Repair (Amst)*, 2014. 13: p. 10-21.
32. Gupta, D. and C.D. Heinen, The mismatch repair-dependent DNA damage response: Mechanisms and implications. *DNA Repair (Amst)*, 2019. 78: p. 60-69.
33. Constantin, N., et al., Human mismatch repair: reconstitution of a nick-directed bidirectional reaction. *J Biol Chem*, 2005. 280(48): p. 39752-61.
34. Hoogstraten, D., et al., Versatile DNA damage detection by the global genome nucleotide excision repair protein XPC. *J Cell Sci*, 2008. 121(Pt 17): p. 2850-9.
35. Rodriguez, A. and A. D'Andrea, Fanconi anemia pathway. *Curr Biol*, 2017. 27(18): p. R986-R988.
36. Kass, E.M., et al., Double-strand break repair by homologous recombination in primary mouse somatic cells requires BRCA1 but not the ATM kinase. *Proc Natl Acad Sci U S A*, 2013. 110(14): p. 5564-9.
37. Mao, Z., et al., Comparison of nonhomologous end joining and homologous recombination in human cells. *DNA Repair (Amst)*, 2008. 7(10): p. 1765-71.
38. Davis, A.J. and D.J. Chen, DNA double strand break repair via non-homologous end-joining. *Transl Cancer Res*, 2013. 2(3): p. 130-143.
39. Quinet, A., et al., Filling gaps in translesion DNA synthesis in human cells. *Mutat Res Genet Toxicol Environ Mutagen*, 2018. 836(Pt B): p. 127-142.

40. Krokan, H.E. and M. Bjoras, Base excision repair. *Cold Spring Harb Perspect Biol*, 2013. 5(4): p. a012583.
41. Wilson, D.M., 3rd and D. Barsky, The major human abasic endonuclease: formation, consequences and repair of abasic lesions in DNA. *Mutat Res*, 2001. 485(4): p. 283-307.
42. Sobol, R.W., et al., Requirement of mammalian DNA polymerase-beta in base-excision repair. *Nature*, 1996. 379(6561): p. 183-6.
43. Matsumoto, Y. and K. Kim, Excision of deoxyribose phosphate residues by DNA polymerase beta during DNA repair. *Science*, 1995. 269(5224): p. 699-702.
44. Kumar, A., et al., Identification and properties of the catalytic domain of mammalian DNA polymerase beta. *Biochemistry*, 1990. 29(31): p. 7156-9.
45. Kumar, A., et al., Studies of the domain structure of mammalian DNA polymerase beta. Identification of a discrete template binding domain. *J Biol Chem*, 1990. 265(4): p. 2124-31.
46. Dianov, G., A. Price, and T. Lindahl, Generation of single-nucleotide repair patches following excision of uracil residues from DNA. *Mol Cell Biol*, 1992. 12(4): p. 1605-12.
47. Matsumoto, Y., K. Kim, and D.F. Bogenhagen, Proliferating cell nuclear antigen-dependent abasic site repair in *Xenopus laevis* oocytes: an alternative pathway of base excision DNA repair. *Mol Cell Biol*, 1994. 14(9): p. 6187-97.
48. Liu, Y., et al., DNA polymerase beta and flap endonuclease 1 enzymatic specificities sustain DNA synthesis for long patch base excision repair. *J Biol Chem*, 2005. 280(5): p. 3665-74.
49. Thapar, U. and B. Demple, Deployment of DNA polymerases beta and lambda in single-nucleotide and multinucleotide pathways of mammalian base excision DNA repair. *DNA Repair (Amst)*, 2019. 76: p. 11-19.
50. Fortini, P., et al., Different DNA polymerases are involved in the short- and long-patch base excision repair in mammalian cells. *Biochemistry*, 1998. 37(11): p. 3575-80.
51. Biade, S., et al., Impairment of proliferating cell nuclear antigen-dependent apurinic/apyrimidinic site repair on linear DNA. *J Biol Chem*, 1998. 273(2): p. 898-902.

52. Prasad, R., et al., DNA polymerase beta -mediated long patch base excision repair. Poly(ADP-ribose)polymerase-1 stimulates strand displacement DNA synthesis. *J Biol Chem*, 2001. 276(35): p. 32411-4.
53. Harrington, J.J. and M.R. Lieber, Functional domains within FEN-1 and RAD2 define a family of structure-specific endonucleases: implications for nucleotide excision repair. *Genes Dev*, 1994. 8(11): p. 1344-55.
54. Murante, R.S., et al., The calf 5'- to 3'-exonuclease is also an endonuclease with both activities dependent on primers annealed upstream of the point of cleavage. *J Biol Chem*, 1994. 269(2): p. 1191-6.
55. Shibata, Y. and T. Nakamura, Defective flap endonuclease 1 activity in mammalian cells is associated with impaired DNA repair and prolonged S phase delay. *J Biol Chem*, 2002. 277(1): p. 746-54.
56. Lindahl, T., P. Karran, and R.D. Wood, DNA excision repair pathways. *Curr Opin Genet Dev*, 1997. 7(2): p. 158-69.
57. Pettersen, H.S., et al., UNG-initiated base excision repair is the major repair route for 5-fluorouracil in DNA, but 5-fluorouracil cytotoxicity depends mainly on RNA incorporation. *Nucleic Acids Res*, 2011. 39(19): p. 8430-44.
58. Haushalter, K.A., et al., Identification of a new uracil-DNA glycosylase family by expression cloning using synthetic inhibitors. *Curr Biol*, 1999. 9(4): p. 174-85.
59. Petronzelli, F., et al., Investigation of the substrate spectrum of the human mismatch-specific DNA N-glycosylase MED1 (MBD4): fundamental role of the catalytic domain. *J Cell Physiol*, 2000. 185(3): p. 473-80.
60. Hardeland, U., et al., Modification of the human thymine-DNA glycosylase by ubiquitin-like proteins facilitates enzymatic turnover. *EMBO J*, 2002. 21(6): p. 1456-64.
61. Dizdaroglu, M., et al., Excision of products of oxidative DNA base damage by human NTH1 protein. *Biochemistry*, 1999. 38(1): p. 243-6.
62. Eide, L., et al., Human endonuclease III acts preferentially on DNA damage opposite guanine residues in DNA. *Biochemistry*, 2001. 40(22): p. 6653-9.
63. McGoldrick, J.P., et al., Characterization of a mammalian homolog of the *Escherichia coli* MutY mismatch repair protein. *Mol Cell Biol*, 1995. 15(2): p. 989-96.
64. Dherin, C., et al., Excision of oxidatively damaged DNA bases by the human alpha-hOgg1 protein and the polymorphic alpha-hOgg1(Ser326Cys) protein



which is frequently found in human populations. *Nucleic Acids Res*, 1999. 27(20): p. 4001-7.

65. Karahalil, B., et al., Substrate specificity of the Ogg1 protein of *Saccharomyces cerevisiae*: excision of guanine lesions produced in DNA by ionizing radiation- or hydrogen peroxide/metal ion-generated free radicals. *Nucleic Acids Res*, 1998. 26(5): p. 1228-33.
66. O'Connor, T.R., Purification and characterization of human 3-methyladenine-DNA glycosylase. *Nucleic Acids Res*, 1993. 21(24): p. 5561-9.
67. Lee, C.Y., et al., Recognition and processing of a new repertoire of DNA substrates by human 3-methyladenine DNA glycosylase (AAG). *Biochemistry*, 2009. 48(9): p. 1850-61.
68. Dou, H., S. Mitra, and T.K. Hazra, Repair of oxidized bases in DNA bubble structures by human DNA glycosylases NEIL1 and NEIL2. *J Biol Chem*, 2003. 278(50): p. 49679-84.
69. Liu, M.M., et al., The mouse ortholog of NEIL3 is a functional DNA glycosylase in vitro and in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, 2010. 107(11): p. 4925-4930.
70. Mol, C.D., et al., Crystal structure and mutational analysis of human uracil-DNA glycosylase: structural basis for specificity and catalysis. *Cell*, 1995. 80(6): p. 869-78.
71. Sun, B., et al., Studies on the catalytic mechanism of five DNA glycosylases. Probing for enzyme-DNA imino intermediates. *J Biol Chem*, 1995. 270(33): p. 19501-8.
72. Lipton, A.S., et al., Characterization of Mg<sup>2+</sup> binding to the DNA repair protein apurinic/aprimidic endonuclease 1 via solid-state 25Mg NMR spectroscopy. *J Am Chem Soc*, 2008. 130(29): p. 9332-41.
73. Li, M. and D.M. Wilson, 3rd, Human apurinic/aprimidic endonuclease 1. *Antioxid Redox Signal*, 2014. 20(4): p. 678-707.
74. Chen, D.S., T. Herman, and B. Dimple, Two distinct human DNA diesterases that hydrolyze 3'-blocking deoxyribose fragments from oxidized DNA. *Nucleic Acids Res*, 1991. 19(21): p. 5907-14.
75. Wilson, D.M., 3rd, Properties of and substrate determinants for the exonuclease activity of human apurinic endonuclease Ape1. *J Mol Biol*, 2003. 330(5): p. 1027-37.

76. Hill, J.W., et al., Stimulation of human 8-oxoguanine-DNA glycosylase by AP-endonuclease: potential coordination of the initial steps in base excision repair. *Nucleic Acids Res*, 2001. 29(2): p. 430-8.
77. Bennett, R.A., et al., Interaction of human apurinic endonuclease and DNA polymerase beta in the base excision repair pathway. *Proc Natl Acad Sci U S A*, 1997. 94(14): p. 7166-9.
78. Wong, D. and B. Demple, Modulation of the 5'-deoxyribose-5-phosphate lyase and DNA synthesis activities of mammalian DNA polymerase beta by apurinic/apyrimidinic endonuclease 1. *J Biol Chem*, 2004. 279(24): p. 25268-75.
79. Ranalli, T.A., S. Tom, and R.A. Bambara, AP endonuclease 1 coordinates flap endonuclease 1 and DNA ligase I activity in long patch base excision repair. *J Biol Chem*, 2002. 277(44): p. 41715-24.
80. Frank, G., et al., Stimulation of eukaryotic flap endonuclease-1 activities by proliferating cell nuclear antigen (PCNA) is independent of its in vitro interaction via a consensus PCNA binding region. *J Biol Chem*, 2001. 276(39): p. 36295-302.
81. Sukhanova, M., S. Khodyreva, and O. Lavrik, Poly(ADP-ribose) polymerase 1 regulates activity of DNA polymerase beta in long patch base excision repair. *Mutat Res*, 2010. 685(1-2): p. 80-9.
82. Lai, Y., et al., Crosstalk between MSH2-MSH3 and polbeta promotes trinucleotide repeat expansion during base excision repair. *Nat Commun*, 2016. 7: p. 12465.
83. Odell, I.D., et al., Nucleosome disruption by DNA ligase III-XRCC1 promotes efficient base excision repair. *Mol Cell Biol*, 2011. 31(22): p. 4623-32.
84. Dianova, II, et al., XRCC1-DNA polymerase beta interaction is required for efficient base excision repair. *Nucleic Acids Res*, 2004. 32(8): p. 2550-5.
85. Wiederhold, L., et al., AP endonuclease-independent DNA base excision repair in human cells. *Mol Cell*, 2004. 15(2): p. 209-20.
86. Paulson, H.L. and K.H. Fischbeck, Trinucleotide repeats in neurogenetic disorders. *Annu Rev Neurosci*, 1996. 19: p. 79-107.
87. Verma, A.K., et al., Exploring the Potential of Small Molecule-Based Therapeutic Approaches for Targeting Trinucleotide Repeat Disorders. *Mol Neurobiol*, 2020. 57(1): p. 566-584.

88. McMurray, C.T., Mechanisms of trinucleotide repeat instability during human development. *Nat Rev Genet*, 2010. 11(11): p. 786-99.
89. Kovtun, I.V., et al., OGG1 initiates age-dependent CAG trinucleotide expansion in somatic cells. *Nature*, 2007. 447(7143): p. 447-52.
90. Liu, Y., et al., Coordination between polymerase beta and FEN1 can modulate CAG repeat expansion. *J Biol Chem*, 2009. 284(41): p. 28352-66.
91. Xu, M., J. Gabison, and Y. Liu, Trinucleotide repeat deletion via a unique hairpin bypass by DNA polymerase beta and alternate flap cleavage by flap endonuclease 1. *Nucleic Acids Res*, 2013. 41(3): p. 1684-97.
92. Lai, Y.H., et al., Instability of CTG Repeats Is Governed by the Position of a DNA Base Lesion through Base Excision Repair. *Plos One*, 2013. 8(2).
93. Xu, M., et al., Base excision repair of oxidative DNA damage coupled with removal of a CAG repeat hairpin attenuates trinucleotide repeat expansion. *Nucleic Acids Res*, 2014. 42(6): p. 3675-91.
94. Beaver, J.M., et al., AP endonuclease 1 prevents trinucleotide repeat expansion via a novel mechanism during base excision repair. *Nucleic Acids Res*, 2015. 43(12): p. 5948-60.
95. Beaver, J.M., et al., Proliferating cell nuclear antigen prevents trinucleotide repeat expansions by promoting repeat deletion and hairpin removal. *DNA Repair (Amst)*, 2016. 48: p. 17-29.
96. Beaver, J.M., et al., An oxidized abasic lesion inhibits base excision repair leading to DNA strand breaks in a trinucleotide repeat tract. *PLoS One*, 2018. 13(2): p. e0192148.
97. Lai, Y., Y. Weizmann, and Y. Liu, The deoxyribose phosphate lyase of DNA polymerase beta suppresses a processive DNA synthesis to prevent trinucleotide repeat instability. *Nucleic Acids Res*, 2018. 46(17): p. 8940-8952.
98. Groh, M. and N. Gromak, Out of balance: R-loops in human disease. *PLoS Genet*, 2014. 10(9): p. e1004630.
99. Shaw, N.N. and D.P. Arya, Recognition of the unique structure of DNA:RNA hybrids. *Biochimie*, 2008. 90(7): p. 1026-39.
100. Roy, D. and M.R. Lieber, G clustering is important for the initiation of transcription-induced R-loops in vitro, whereas high G density without clustering is sufficient thereafter. *Mol Cell Biol*, 2009. 29(11): p. 3124-33.

101. Roy, D., et al., Competition between the RNA transcript and the nontemplate DNA strand during R-loop formation in vitro: a nick can serve as a strong R-loop initiation site. *Mol Cell Biol*, 2010. 30(1): p. 146-59.
102. Westover, K.D., D.A. Bushnell, and R.D. Kornberg, Structural basis of transcription: separation of RNA from DNA by RNA polymerase II. *Science*, 2004. 303(5660): p. 1014-6.
103. Xu, B. and D.A. Clayton, RNA-DNA hybrid formation at the human mitochondrial heavy-strand origin ceases at replication start sites: an implication for RNA-DNA hybrids serving as primers. *EMBO J*, 1996. 15(12): p. 3135-43.
104. Baldacci, G., B. Cherif-Zahar, and G. Bernardi, The initiation of DNA replication in the mitochondrial genome of yeast. *EMBO J*, 1984. 3(9): p. 2115-20.
105. Pavri, R., R Loops in the Regulation of Antibody Gene Diversification. *Genes (Basel)*, 2017. 8(6).
106. Chaudhuri, J. and F.W. Alt, Class-switch recombination: interplay of transcription, DNA deamination and DNA repair. *Nat Rev Immunol*, 2004. 4(7): p. 541-52.
107. Guikema, J.E., et al., APE1- and APE2-dependent DNA breaks in immunoglobulin class switch recombination. *J Exp Med*, 2007. 204(12): p. 3017-26.
108. Masani, S., L. Han, and K. Yu, Apurinic/aprimidinic endonuclease 1 is the essential nuclease during immunoglobulin class switch recombination. *Mol Cell Biol*, 2013. 33(7): p. 1468-73.
109. Rada, C., J.M. Di Noia, and M.S. Neuberger, Mismatch recognition and uracil excision provide complementary paths to both Ig switching and the A/T-focused phase of somatic mutation. *Mol Cell*, 2004. 16(2): p. 163-71.
110. Muramatsu, M., et al., Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell*, 2000. 102(5): p. 553-63.
111. Skourti-Stathaki, K., N.J. Proudfoot, and N. Gromak, Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. *Mol Cell*, 2011. 42(6): p. 794-805.
112. Chen, Y.Z., et al., DNA/RNA helicase gene mutations in a form of juvenile amyotrophic lateral sclerosis (ALS4). *Am J Hum Genet*, 2004. 74(6): p. 1128-35.

113. Moreira, M.C., et al., Senataxin, the ortholog of a yeast RNA helicase, is mutant in ataxia-ocular apraxia 2. *Nat Genet*, 2004. 36(3): p. 225-7.
114. Suraweera, A., et al., Senataxin, defective in ataxia oculomotor apraxia type 2, is involved in the defense against oxidative DNA damage. *J Cell Biol*, 2007. 177(6): p. 969-79.
115. Yuce, O. and S.C. West, Senataxin, defective in the neurodegenerative disorder ataxia with oculomotor apraxia 2, lies at the interface of transcription and the DNA damage response. *Mol Cell Biol*, 2013. 33(2): p. 406-17.
116. Alzu, A., et al., Senataxin associates with replication forks to protect fork integrity across RNA-polymerase-II-transcribed genes. *Cell*, 2012. 151(4): p. 835-846.
117. Grabczyk, E., M. Mancuso, and M.C. Sammarco, A persistent RNA.DNA hybrid formed by transcription of the Friedreich ataxia triplet repeat in live bacteria, and by T7 RNAP in vitro. *Nucleic Acids Res*, 2007. 35(16): p. 5351-9.
118. Mclvor, E.I., U. Polak, and M. Napierala, New insights into repeat instability: role of RNA\*DNA hybrids. *RNA Biol*, 2010. 7(5): p. 551-8.
119. Groh, M., et al., R-loops associated with triplet repeat expansions promote gene silencing in Friedreich ataxia and fragile X syndrome. *PLoS Genet*, 2014. 10(5): p. e1004318.
120. Su, X.A. and C.H. Freudenreich, Cytosine deamination and base excision repair cause R-loop-induced CAG repeat fragility and instability in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A*, 2017. 114(40): p. E8392-E8401.
121. Reddy, K., et al., Processing of double-R-loops in (CAG).(CTG) and C9orf72 (GGGGCC).(GGCCCC) repeats causes instability. *Nucleic Acids Res*, 2014. 42(16): p. 10473-87.
122. Bhatia, V., et al., BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2. *Nature*, 2014. 511(7509): p. 362-5.
123. Hill, S.J., et al., Systematic screening reveals a role for BRCA1 in the response to transcription-associated DNA damage. *Genes Dev*, 2014. 28(17): p. 1957-75.
124. Gomez-Gonzalez, B. and A. Aguilera, Activation-induced cytidine deaminase action is strongly stimulated by mutations of the THO complex. *Proc Natl Acad Sci U S A*, 2007. 104(20): p. 8409-14.

125. Aguilera, A., The connection between transcription and genomic instability. *EMBO J*, 2002. 21(3): p. 195-201.
126. Wahba, L., et al., RNase H and multiple RNA biogenesis factors cooperate to prevent RNA:DNA hybrids from generating genome instability. *Mol Cell*, 2011. 44(6): p. 978-88.
127. Chib, S., A.K. Byrd, and K.D. Raney, Yeast Helicase Pif1 Unwinds RNA:DNA Hybrids with Higher Processivity than DNA:DNA Duplexes. *J Biol Chem*, 2016. 291(11): p. 5889-901.
128. Boule, J.B. and V.A. Zakian, The yeast Pif1p DNA helicase preferentially unwinds RNA DNA substrates. *Nucleic Acids Res*, 2007. 35(17): p. 5809-18.
129. Cristini, A., et al., RNA/DNA Hybrid Interactome Identifies DXH9 as a Molecular Player in Transcriptional Termination and R-Loop-Associated DNA Damage. *Cell Rep*, 2018. 23(6): p. 1891-1905.
130. Groh, M., et al., Senataxin: Genome Guardian at the Interface of Transcription and Neurodegeneration. *J Mol Biol*, 2017. 429(21): p. 3181-3195.
131. Sollier, J., et al., Transcription-coupled nucleotide excision repair factors promote R-loop-induced genome instability. *Mol Cell*, 2014. 56(6): p. 777-85.
132. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group. *Cell*, 1993. 72(6): p. 971-83.
133. Kim, S.D. and V.S. Fung, An update on Huntington's disease: from the gene to the clinic. *Curr Opin Neurol*, 2014. 27(4): p. 477-83.
134. Vonsattel, J.P. and M. DiFiglia, Huntington disease. *J Neuropathol Exp Neurol*, 1998. 57(5): p. 369-84.
135. Sakazume, S., et al., A patient with early onset Huntington disease and severe cerebellar atrophy. *Am J Med Genet A*, 2009. 149A(4): p. 598-601.
136. McMurray, C.T., Mechanisms of DNA Expansion. *Chromosoma*, 1995. 104(1): p. 2-13.
137. Tang, W., et al., Friedreich's ataxia (GAA)<sub>n</sub>\*(TTC)<sub>n</sub> repeats strongly stimulate mitotic crossovers in *Saccharomyces cerevisiae*. *PLoS Genet*, 2011. 7(1): p. e1001270.

138. Lin, Y.F. and J.H. Wilson, Transcription-induced DNA toxicity at trinucleotide repeats Double bubble is trouble. *Cell Cycle*, 2011. 10(4): p. 611-618.
139. Sinden, R.R., M.J. Pytlos-Sinden, and V.N. Potaman, Slipped strand DNA structures. *Front Biosci*, 2007. 12: p. 4788-99.
140. Polidori, M.C., et al., Oxidative damage to mitochondrial DNA in Huntington's disease parietal cortex. *Neurosci Lett*, 1999. 272(1): p. 53-6.
141. Bogdanov, M.B., et al., Increased oxidative damage to DNA in a transgenic mouse model of Huntington's disease. *J Neurochem*, 2001. 79(6): p. 1246-9.
142. Chan, N.L., et al., Coordinated processing of 3' slipped (CAG)<sub>n</sub>/(CTG)<sub>n</sub> hairpins by DNA polymerases beta and delta preferentially induces repeat expansions. *J Biol Chem*, 2013. 288(21): p. 15015-22.
143. Xu, M., et al., A 5', 8-cyclo-2'-deoxypurine lesion induces trinucleotide repeat deletion via a unique lesion bypass by DNA polymerase beta. *Nucleic Acids Res*, 2014. 42(22): p. 13749-63.
144. Goula, A.V., et al., The nucleotide sequence, DNA damage location, and protein stoichiometry influence the base excision repair outcome at CAG/CTG repeats. *Biochemistry*, 2012. 51(18): p. 3919-32.
145. Goula, A.V., et al., Stoichiometry of base excision repair proteins correlates with increased somatic CAG instability in striatum over cerebellum in Huntington's disease transgenic mice. *PLoS Genet*, 2009. 5(12): p. e1000749.
146. Imam, S.Z., et al., Mitochondrial and nuclear DNA-repair capacity of various brain regions in mouse is altered in an age-dependent manner. *Neurobiol Aging*, 2006. 27(8): p. 1129-36.
147. Wheeler, V.C., et al., Mismatch repair gene Msh2 modifies the timing of early disease in Hdh(Q111) striatum. *Hum Mol Genet*, 2003. 12(3): p. 273-81.
148. Dragileva, E., et al., Intergenerational and striatal CAG repeat instability in Huntington's disease knock-in mice involve different DNA repair genes. *Neurobiol Dis*, 2009. 33(1): p. 37-47.
149. Genetic Modifiers of Huntington's Disease, C., Identification of Genetic Factors that Modify Clinical Onset of Huntington's Disease. *Cell*, 2015. 162(3): p. 516-26.

150. Bettencourt, C., et al., DNA repair pathways underlie a common genetic mechanism modulating onset in polyglutamine diseases. *Ann Neurol*, 2016. 79(6): p. 983-90.
151. Goold, R., et al., FAN1 modifies Huntington's disease progression by stabilizing the expanded HTT CAG repeat. *Hum Mol Genet*, 2019. 28(4): p. 650-661.
152. Liu, T., et al., FAN1 acts with FANCI-FANCD2 to promote DNA interstrand cross-link repair. *Science*, 2010. 329(5992): p. 693-6.
153. Cannavo, E., et al., Characterization of the interactome of the human MutL homologues MLH1, PMS1, and PMS2. *J Biol Chem*, 2007. 282(5): p. 2976-86.
154. Kratz, K., et al., Deficiency of FANCD2-associated nuclease KIAA1018/FAN1 sensitizes cells to interstrand crosslinking agents. *Cell*, 2010. 142(1): p. 77-88.
155. Smogorzewska, A., et al., A genetic screen identifies FAN1, a Fanconi anemia-associated nuclease necessary for DNA interstrand crosslink repair. *Mol Cell*, 2010. 39(1): p. 36-47.
156. MacKay, C., et al., Identification of KIAA1018/FAN1, a DNA repair nuclease recruited to DNA damage by monoubiquitinated FANCD2. *Cell*, 2010. 142(1): p. 65-76.
157. Pizzolato, J., et al., FANCD2-associated nuclease 1, but not exonuclease 1 or flap endonuclease 1, is able to unhook DNA interstrand cross-links in vitro. *J Biol Chem*, 2015. 290(37): p. 22602-11.
158. Wang, R., et al., DNA repair. Mechanism of DNA interstrand cross-link processing by repair nuclease FAN1. *Science*, 2014. 346(6213): p. 1127-30.
159. Ren, Y., et al., Modulation of trinucleotide repeat instability by DNA polymerase beta polymorphic variant R137Q. *PLoS One*, 2017. 12(5): p. e0177299.
160. Li, W., et al., Rapid induction and long-term self-renewal of primitive neural precursors from human embryonic stem cells by small molecule inhibitors. *Proc Natl Acad Sci U S A*, 2011. 108(20): p. 8299-304.
161. Porro, A., et al., FAN1 interaction with ubiquitylated PCNA alleviates replication stress and preserves genomic integrity independently of BRCA2. *Nat Commun*, 2017. 8(1): p. 1073.



162. Henricksen, L.A., et al., Inhibition of flap endonuclease 1 by flap secondary structure and relevance to repeat sequence expansion. *J Biol Chem*, 2000. 275(22): p. 16420-7.
163. Liu, Y. and R.A. Bambara, Analysis of human flap endonuclease 1 mutants reveals a mechanism to prevent triplet repeat expansion. *J Biol Chem*, 2003. 278(16): p. 13728-39.
164. Liu, Y., et al., *Saccharomyces cerevisiae* flap endonuclease 1 uses flap equilibration to maintain triplet repeat stability. *Mol Cell Biol*, 2004. 24(9): p. 4049-64.
165. Liu, Y., H.I. Kao, and R.A. Bambara, Flap endonuclease 1: a central component of DNA metabolism. *Annu Rev Biochem*, 2004. 73: p. 589-615.
166. Balakrishnan, L. and R.A. Bambara, Flap endonuclease 1. *Annu Rev Biochem*, 2013. 82: p. 119-38.
167. Gloor, J.W., L. Balakrishnan, and R.A. Bambara, Flap endonuclease 1 mechanism analysis indicates flap base binding prior to threading. *J Biol Chem*, 2010. 285(45): p. 34922-31.
168. Zhou, W., et al., FAN1 mutations cause karyomegalic interstitial nephritis, linking chronic kidney failure to defective DNA damage repair. *Nat Genet*, 2012. 44(8): p. 910-5.
169. Segui, N., et al., Germline Mutations in FAN1 Cause Hereditary Colorectal Cancer by Impairing DNA Repair. *Gastroenterology*, 2015. 149(3): p. 563-6.
170. Lans, H. and J.H. Hoeijmakers, Genome stability, progressive kidney failure and aging. *Nat Genet*, 2012. 44(8): p. 836-8.
171. Ionita-Laza, I., et al., Scan statistic-based analysis of exome sequencing data identifies FAN1 at 15q13.3 as a susceptibility gene for schizophrenia and autism. *Proc Natl Acad Sci U S A*, 2014. 111(1): p. 343-8.
172. Tom, S., L.A. Henricksen, and R.A. Bambara, Mechanism whereby proliferating cell nuclear antigen stimulates flap endonuclease 1. *J Biol Chem*, 2000. 275(14): p. 10498-505.
173. Zhang, Z., et al., Structure of monoubiquitinated PCNA: implications for DNA polymerase switching and Okazaki fragment maturation. *Cell Cycle*, 2012. 11(11): p. 2128-36.
174. Mirkin, S.M., DNA structures, repeat expansions and human hereditary disorders. *Curr Opin Struct Biol*, 2006. 16(3): p. 351-8.

175. Freudenreich, C.H., S.M. Kantrow, and V.A. Zakian, Expansion and length-dependent fragility of CTG repeats in yeast. *Science*, 1998. 279(5352): p. 853-6.
176. Garcia-Muse, T. and A. Aguilera, R Loops: From Physiological to Pathological Roles. *Cell*, 2019. 179(3): p. 604-618.
177. Wahba, L., et al., S1-DRIP-seq identifies high expression and polyA tracts as major contributors to R-loop formation. *Genes Dev*, 2016. 30(11): p. 1327-38.
178. Lin, Y., et al., R loops stimulate genetic instability of CTG.CAG repeats. *Proc Natl Acad Sci U S A*, 2010. 107(2): p. 692-7.
179. Reddy, K., et al., Determinants of R-loop formation at convergent bidirectionally transcribed trinucleotide repeats. *Nucleic Acids Res*, 2011. 39(5): p. 1749-62.
180. Loomis, E.W., et al., Transcription-associated R-loop formation across the human FMR1 CGG-repeat region. *PLoS Genet*, 2014. 10(4): p. e1004294.
181. Becherel, O.J., et al., A new model to study neurodegeneration in ataxia oculomotor apraxia type 2. *Hum Mol Genet*, 2015. 24(20): p. 5759-74.
182. Abbotts, J., et al., Studies on the mechanism of Escherichia coli DNA polymerase I large fragment. Effect of template sequence and substrate variation on termination of synthesis. *J Biol Chem*, 1988. 263(29): p. 15094-103.
183. Keijzers, G., D. Liu, and L.J. Rasmussen, Exonuclease 1 and its versatile roles in DNA repair. *Crit Rev Biochem Mol Biol*, 2016. 51(6): p. 440-451.
184. Zhao, X., et al., MutLgamma promotes repeat expansion in a Fragile X mouse model while EXO1 is protective. *PLoS Genet*, 2018. 14(10): p. e1007719.
185. Zhao, X.N. and K. Usdin, FAN1 protects against repeat expansions in a Fragile X mouse model. *DNA Repair (Amst)*, 2018. 69: p. 1-5.
186. Keum, J.W., et al., The HTT CAG-Expansion Mutation Determines Age at Death but Not Disease Duration in Huntington Disease. *Am J Hum Genet*, 2016. 98(2): p. 287-98.
187. Grunseich, C., et al., Senataxin Mutation Reveals How R-Loops Promote Transcription by Blocking DNA Methylation at Gene Promoters. *Mol Cell*, 2018. 69(3): p. 426-437 e7.

188. Lieber, M.R., The FEN-1 family of structure-specific nucleases in eukaryotic DNA replication, recombination and repair. *Bioessays*, 1997. 19(3): p. 233-40.
189. Harrington, J.J. and M.R. Lieber, The characterization of a mammalian DNA structure-specific endonuclease. *EMBO J*, 1994. 13(5): p. 1235-46.
190. Zheng, L., et al., Novel function of the flap endonuclease 1 complex in processing stalled DNA replication forks. *EMBO Rep*, 2005. 6(1): p. 83-9.
191. Bornarth, C.J., et al., Effect of flap modifications on human FEN1 cleavage. *Biochemistry*, 1999. 38(40): p. 13347-54.
192. Saharia, A., et al., FEN1 ensures telomere stability by facilitating replication fork re-initiation. *J Biol Chem*, 2010. 285(35): p. 27057-66.
193. Teasley, D.C., et al., Flap Endonuclease 1 Limits Telomere Fragility on the Leading Strand. *J Biol Chem*, 2015. 290(24): p. 15133-45.
194. Liu, L., et al., Functional FEN1 genetic variants contribute to risk of hepatocellular carcinoma, esophageal cancer, gastric cancer and colorectal cancer. *Carcinogenesis*, 2012. 33(1): p. 119-23.
195. Yang, M., et al., Functional FEN1 polymorphisms are associated with DNA damage levels and lung cancer risk. *Hum Mutat*, 2009. 30(9): p. 1320-8.
196. Zheng, L., et al., Fen1 mutations result in autoimmunity, chronic inflammation and cancers. *Nat Med*, 2007. 13(7): p. 812-9.
197. Larsen, E., et al., Early-onset lymphoma and extensive embryonic apoptosis in two domain-specific Fen1 mice mutants. *Cancer Res*, 2008. 68(12): p. 4571-9.
198. Zheng, L., et al., Functional regulation of FEN1 nuclease and its link to cancer. *Nucleic Acids Res*, 2011. 39(3): p. 781-94.
199. Kitamura, K., et al., Flap endonuclease 1 is involved in cccDNA formation in the hepatitis B virus. *PLoS Pathog*, 2018. 14(6): p. e1007124.
200. Zeng, X., et al., FEN1 mediates miR-200a methylation and promotes breast cancer cell growth via MET and EGFR signaling. *FASEB J*, 2019. 33(10): p. 10717-10730.
201. Liu, S.B., et al., Transcriptome Analysis of FEN1 Knockdown HEK293T Cell Strain Reveals Alteration in Nucleic Acid Metabolism, Virus Infection, Cell Morphogenesis and Cancer Development. *Comb Chem High Throughput Screen*, 2019. 22(6): p. 379-386.

202. Murante, R.S., et al., Calf RTH-1 nuclease can remove the initiator RNAs of Okazaki fragments by endonuclease activity. *J Biol Chem*, 1996. 271(42): p. 25888-97.
203. Stevens, A., Endonucleolytic cleavage of RNA at 5' endogenous stem structures by human flap endonuclease 1. *Biochem Biophys Res Commun*, 1998. 251(2): p. 501-8.
204. Yu, K., et al., R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nat Immunol*, 2003. 4(5): p. 442-51.
205. Jinek, M., et al., A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, 2012. 337(6096): p. 816-21.
206. Lin, Y., et al., Convergent transcription through a long CAG tract destabilizes repeats and induces apoptosis. *Mol Cell Biol*, 2010. 30(18): p. 4435-51.
207. Patterson, T.A., et al., Molecular cloning and high-level expression of human polymerase beta cDNA and comparison of the purified recombinant human and rat enzymes. *Protein Expr Purif*, 2000. 18(1): p. 100-10.
208. Ballmaier, D. and B. Epe, DNA damage by bromate: mechanism and consequences. *Toxicology*, 2006. 221(2-3): p. 166-71.
209. Lai, Y., et al., AP endonuclease 1 prevents the extension of a T/G mismatch by DNA polymerase beta to prevent mutations in CpGs during base excision repair. *DNA Repair (Amst)*, 2016. 43: p. 89-97.
210. Barzilay, G., et al., Site-directed mutagenesis of the human DNA repair enzyme HAP1: identification of residues important for AP endonuclease and RNase H activity. *Nucleic Acids Res*, 1995. 23(9): p. 1544-50.
211. Kim, W.C., D. King, and C.H. Lee, RNA-cleaving properties of human apurinic/apyrimidinic endonuclease 1 (APE1). *Int J Biochem Mol Biol*, 2010. 1(1): p. 12-25.
212. Chohan, M., et al., Human apurinic/apyrimidinic endonuclease 1 (APE1) has 3' RNA phosphatase and 3' exoribonuclease activities. *J Mol Biol*, 2015. 427(2): p. 298-311.
213. Barnes, T., et al., Identification of Apurinic/apyrimidinic endonuclease 1 (APE1) as the endoribonuclease that cleaves c-myc mRNA. *Nucleic Acids Res*, 2009. 37(12): p. 3946-58.

214. Berquist, B.R., D.R. McNeill, and D.M. Wilson, 3rd, Characterization of abasic endonuclease activity of human Ape1 on alternative substrates, as well as effects of ATP and sequence context on AP site incision. *J Mol Biol*, 2008. 379(1): p. 17-27.

## VITA

EDUARDO E. LAVERDE

Miami, FL

- 2015-2020            Biochemistry Ph.D. Candidate  
Florida International University  
Miami, FL
- 2014                 B.S. Biology  
Minor: Nutrition/Chemistry  
Florida International University,  
Miami, FL
- 2007                 Associate Degree/Mortgage Broker  
Miami Dade College, Miami, FL

### Publications

Beaver, J. M., Lai, Y., Xu, M., Casin, A. H., Laverde, E. E., & Liu, Y. (2015). AP endonuclease 1 prevents trinucleotide repeat expansion via a novel mechanism during base excision repair. *Nucleic Acids Research*, 43(12), 5948–5960.

Jiang, Z., Xu, M., Lai, Y., Laverde, E. E., Terzidis, M. A., Masi, A., Chatgililoglu, C., Liu, Y. (2015). Bypass of a 5',8-cyclopurine-2'-deoxynucleoside by DNA polymerase  $\beta$  during DNA replication and base excision repair leads to nucleotide misinsertions and DNA strand breaks. *DNA Repair*, 33, 24–34.

Masi A, Sabbia A, Ferreri C, Manoli F, Lai Y, Laverde E.E., Liu Y, Krokidis MG, Chatgililoglu C, Faraone MR. (2019) Diastereomeric Recognition of 5',8-cyclo-2'-Deoxyadenosine Lesions by Human Poly(ADP-ribose) Polymerase 1 in a Biomimetic Model. *Cells*, 8, 116.

Laverde, E.E, Lai, Y., and Liu, Y.; Proliferating Cell Nuclear Antigen Stimulates Flap Endonuclease 1 to Modulate GAA Repeat Stability Via Base Excision Repair. Poster presented at the 46th Annual Meeting of Environmental Mutagenesis and Genomics Society (EMGS), Sep. 26th-30th, 2015, New Orleans, LA. ENVIRONMENTAL AND MOLECULAR MUTAGENESIS Volume: 56 Supplement: 1 Pages: S62-S62 Meeting Abstract: P22 Published: AUG 2015

Laverde, E.E, Lai, Y., and Liu, Y.; Base Excision Repair in a GAA Repeat R-loop Leads to GAA Repeats Instability. Poster presented at the 48th Annual Meeting of Environmental Mutagenesis and Genomics Society (EMGS), Sep. 9th-13th, 2017, Raleigh, NC. ENVIRONMENTAL AND MOLECULAR MUTAGENESIS Volume: 58 Supplement: 1 Pages: S57-S57 Meeting Abstract: P10 Published: SEP 2017

Laverde, E.E, Lai, Y., and Liu, Y.; R-loop at GAA Repeats Mediates Repeat Instability Through Modulation of BER Enzymatic Activity. Poster presented at the 49th Annual Meeting of Environmental Mutagenesis and Genomics Society (EMGS), Sep. 22th-26th, 2018, San Antonio, TX. ENVIRONMENTAL AND MOLECULAR MUTAGENESIS Volume: 59 Supplement: 1 Pages: 88-88 Meeting Abstract: P17 Published: SEP 2018

Ren Y, Lai Y, Laverde E.E., Lei R, Rein HL, Liu Y. (2017). Modulation of trinucleotide repeat instability by DNA polymerase  $\beta$  polymorphic variant R137Q. PLoS ONE 12(5): e0177299.

Suzol SH, Howlader AH, Wen Z, Ren Y, Laverde E.E., Garcia C, Liu Y, and Wnuk SF. (2018). Pyrimidine Nucleosides with a Reactive ( $\beta$ -Chlorovinyl)sulfone or ( $\beta$ -Keto)sulfone Group at the C5 Position, Their Reactions with Nucleophiles and Electrophiles, and Their Polymerase-Catalyzed Incorporation into DNA. ACS Omega 3(4), 4276-4288.

Wen Z, Tuttle P, Howlader H, Vasilyeva A, Gonzalez L, Tangar A, Lei R, Laverde E.E, Liu Y, Miksovska J, and Wnuk SF. (2019). Fluorescent 5-Pyrimidine and 8-Purine Nucleosides Modified with an N-Unsubstituted 1,2,3-Triazol-4-yl Moiety. The Journal of Organic Chemistry, 84 (6), 3624-3631