

9-12-2014

# Oxidative DNA Damage Modulates Trinucleotide Repeat Instability Via DNA Base Excision Repair

Meng Xu

*Florida International University, xmeng005@fiu.edu*

**DOI:** 10.25148/etd.FI14110704

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

 Part of the [Biochemistry Commons](#)

---

## Recommended Citation

Xu, Meng, "Oxidative DNA Damage Modulates Trinucleotide Repeat Instability Via DNA Base Excision Repair" (2014). *FIU Electronic Theses and Dissertations*. 1576.  
<https://digitalcommons.fiu.edu/etd/1576>

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

OXIDATIVE DNA DAMAGE MODULATES TRINUCLEOTIDE REPEAT  
INSTABILITY VIA DNA BASE EXCISION REPAIR

A dissertation submitted in partial fulfillment of

the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

CHEMISTRY

by

Meng Xu

2014

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

This dissertation, written by Meng Xu, and entitled Oxidative DNA Damage Modulates Trinucleotide Repeat Instability via DNA Base Excision Repair, 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.

---

Irina Agoulnik

---

David Chatfield

---

Kathleen Rein

---

Xiaotang Wang

---

Yuan Liu, Major Professor

Date of Defense: September 12, 2014

The dissertation of Meng Xu is approved.

---

Interim Dean Michael R. Heithaus  
College of Arts and Sciences

---

Dean Lakshmi N. Reddi  
University Graduate School

Florida International University, 2014

© Copyright 2014 by Meng Xu

All rights reserved.

## DEDICATION

I dedicate this dissertation to my parents, Zhaoyong Xu and Yiping Xiao for their love, great support, and encouragement.

## ACKNOWLEDGMENTS

Words are not enough for me to express my gratefulness to all the people who gave me all kinds of suggestions and support throughout my Ph.D. study. First of all, I would like to thank my advisor Dr. Yuan Liu for her amazing mentoring and great support. Under her remarkable guidance, I gained a lot of experience on how to conduct a good scientific research through envisioning a scientific problem, developing scientific thinking and scientific presentation skill and a lot more. I am really appreciated for her great support, patience and encouragement.

I would like to thank my committee members Drs. Irina Agoulnik, David Chatfield, Kathleen Rein and Xiaotang Wang for their insightful advice and great support during my entire Ph.D. training and future career development. I couldn't have a better committee to guide me through my graduate training.

I am also very appreciated for the great support from Dr. Samuel H. Wilson from National Institute of Environmental Health Sciences/National Institutes of Health and Dr. Stephen C. West at Clare Hall Laboratories, London Research Institute, Cancer Research UK for generously providing me with purified base excision repair enzymes and vector for expressing Mus81/Eme1 and for their great support on my research.

I would like to thank Dr. Yanhao Lai for her support, help and friendship.

Finally I would like to thank Jill Beaver, Zhongliang Jiang, Yaou Ren, Eduardo Laverde, Justin Torner and other members in the Liu laboratory for their great support, helpful suggestions and persistent encouragement.

ABSTRACT OF THE DISSERTATION  
OXIDATIVE DNA DAMAGE MODULATES TRINUCLEOTIDE REPEAT  
INSTABILITY VIA DNA BASE EXCISION REPAIR

by

Meng Xu

Florida International University, 2014

Miami, Florida

Professor Yuan Liu, Major Professor

Trinucleotide repeat (TNR) expansion is the cause of more than 40 types of human neurodegenerative diseases such as Huntington's disease. Recent studies have linked TNR expansion with oxidative DNA damage and base excision repair (BER). In this research, we provided the first evidence that oxidative DNA damage can induce CAG repeat deletion/contraction via BER. We found that BER of an oxidized DNA base lesion, 8-oxoguanine in a CAG repeat tract, resulted in the formation of a CTG hairpin at the template strand. DNA polymerase  $\beta$  (pol  $\beta$ ) then skipped over the hairpin creating a 5'-flap that was cleaved by flap endonuclease 1 (FEN1) leading to CAG repeat deletion. To further investigate whether BER may help to shorten an expanded TNR tract, we examined BER in a CAG repeat hairpin loop. We found that 8-oxoguanine DNA glycosylase removed the oxidized base located in the loop region of the hairpin leaving an abasic site. Apurinic/apyrimidinic (AP) endonuclease 1 then incised the 5'-end of the abasic site leaving a nick in the loop. This further converted the hairpin into an intermediate with a 3'-flap and a 5'-flap. As a 5'-3' endonuclease, FEN1 cleaved the 5'-flap, whereas a 3'-5' endonuclease, Mus81/Eme1, removed the 3'-flap. The coordination

between FEN1 and Mus81/Eme1 ultimately resulted in removal of a CAG repeat hairpin attenuating or preventing TNR expansion. To further explore if pol  $\beta$  bypass of an oxidized base lesion, 5',8-cyclodeoxyadenosine, may affect TNR instability, we examined pol  $\beta$  DNA synthesis in bypassing this base lesion and found that the lesion preferentially induced TNR deletion during BER and Okazaki fragment maturation. The repeat deletion was mediated by the formation of a loop in the template strand induced specifically by the damage. Pol  $\beta$  then skipped over the loop structure creating a 5'-flap that was efficiently removed by FEN1 leading to repeat deletion. Our study demonstrates that pol  $\beta$ -mediated BER plays an important role in mediating TNR deletion and removing a TNR hairpin to prevent TNR expansion. Our research provides a molecular basis for further developing BER as a target for prevention and treatment of neurodegenerative diseases caused by TNR expansion.



## TABLE OF CONTENTS

CHAPTER	PAGE
INTRODUCTION .....	1
A. Oxidative DNA damage and repair.....	1
B. Trinucleotide repeat (TNR) instability and human diseases .....	12
C. TNR instability and DNA replication and RNA transcription.....	15
D. TNR instability and DNA repair .....	19
E. Base excision repair (BER) and TNR instability .....	28
OVERVIEW .....	36
CHAPTER 1: TRINUCLEOTIDE REPEAT DELETION VIA A UNIQUE HAIRPIN BYPASS BY DNA POLYMERASE $\beta$ AND ALTERNATE FLAP CLEAVAGE BY FLAP ENDONUCLEASE 1	
ABSTRACT.....	38
INTRODUCTION .....	38
MATERIALS AND METHODS .....	42
RESULTS .....	52
DISCUSSION.....	78
CHAPTER 2: BASE EXCISION REPAIR OF OXIDATIVE DNA DAMAGE COUPLED WITH REMOVAL OF A CAG REPEAT HAIRPIN ATTENUATES TRINUCLEOTIDE REPEAT EXPANSION	
ABSTRACT.....	85
INTRODUCTION .....	86
MATERIALS AND METHODS .....	88
RESULTS .....	94
DISCUSSION.....	122
CHAPTER 3: A 5', 8-CYCLODEOXYPURINE LESION INDUCES TRINUCLEOTIDE REPEAT DELETION VIA A UNIQUE LESION BYPASS BY DNA POLYMERASE $\beta$	
ABSTRACT.....	132
INTRODUCTION .....	133
MATERIALS AND METHODS.....	137
RESULTS .....	143
DISCUSSION.....	169
SUMMARY .....	177
REFERENCE .....	183
VITA .....	208

## LIST OF FIGURES

FIGURE	PAGE
INTRODUCTION	
Figure I.1 Deamination of Cytosine (C) and 5-Methylcytosine (5mC).....	2
Figure I.2 Chromosome breaks by DNA damage .....	4
Figure I.3 DNA damage, repair mechanisms and consequences .....	5
Figure I.4 DNA repair rate is an important determinant of cell pathology .....	6
Figure I.5 8-oxoG base pair with dC and dA.....	8
Figure I.6 Base excision repair (BER) of oxidized DNA base lesions .....	11
Figure I.7 Trinucleotide repeat expansion disorders caused by triplet repeats in coding and noncoding gene regions .....	12
Figure I.8 Non-B form DNA structures formed by TNR .....	14
Figure I.9 Replication mechanisms for repeat instability .....	16
Figure I.10 Correlation between replication, orientation of the (CTG/CAG) <sub>n</sub> tract, transcription, structural properties of repeated DNA and the genetic instability of the TNR .....	18
Figure I.11 Eukaryotic mismatch repair complexes .....	21
Figure I.12 Mismatch repair induce hairpin formation .....	23
Figure I.13 Speculative model for transcription-induced repeat instability .....	25
Figure I.14 Nucleotide excision repair and trinucleotide repeat loop formation .....	27
Figure I.15 FEN1 alternate flap cleavage mediated CAG repeat expansion .....	31
Figure I.16 Hypothetical models illustrating CAG repeat stability modulated by coordination among BER enzymes during BER of 8-oxoG .....	32
Figure I.17 Loops formed during base excision repair by strand displacement .....	34

## CHAPTER 1

Figure 1.1 Scheme of DNA Sequencing for Repeat Deletion Products .....	49
Figure 1.2 CAG repeat deletion during BER with MEF cell extracts. ....	54
Figure 1.3 CAG repeat deletion during BER with purified BER enzymes .....	55
Figure 1.4 Formation of trinucleotide repeat hairpins on the (CAG) <sub>20</sub> repeat substrate during BER .....	57
Figure 1.5 Formation of trinucleotide repeat hairpins on the (CTG) <sub>20</sub> repeat substrate during BER .....	58
Figure 1.6 Probing of hairpins formed by random sequences .....	59
Figure 1.7 Hairpin bypass synthesis of pol β with template hairpin containing substrates without a downstream flap .....	62
Figure 1.8 Hairpin bypass synthesis of pol β with template hairpin containing substrates with a downstream flap .....	63
Figure 1.9 Probing of hairpins formed by template strand containing (CAG) <sub>7</sub> or (CTG) <sub>7</sub> repeats. ....	64
Figure 1.10 Probing of hairpins formed by damaged strand with a 5'-(CAG) <sub>25</sub> flap .....	65
Figure 1.11 Probing of hairpins formed (CTG) <sub>25</sub> repeat in the template strand .....	66
Figure 1.12 The number of nucleotides deleted and the percentage of CAG repeat deletions resulting from BER.....	67
Figure 1.13 Spectrums of CAG repeat deletions resulting from BER .....	68
Figure 1.14 Hairpin bypass synthesis of pol β and FEN1 flap cleavage coordinate with each other leading to CTG repeat deletion. ....	69
Figure 1.15 FEN1 cleavage of CAG repeats during BER .....	73
Figure 1.16 FEN1 cleavage of repeats containing downstream flap during BER .....	74
Figure 1.17 FEN1 flap cleavage of a downstream 5'-THF-(CTG) <sub>25</sub> hairpin/flap .....	75
Figure 1.18 FEN1 promotes limited CAG repeat deletion .....	76

Figure 1.19 Western-blot analysis of FEN1 depletion in MEF cell extract.....	77
Figure 1.20 Models for CAG repeat deletion during BER .....	80
 CHAPTER 2	
Figure 2.1 Hairpin probing of (CAG) <sub>7</sub> - and (CAG) <sub>14</sub> -8-oxoG hairpin substrates by Mung Bean Nuclease .....	97
Figure 2.2 Hairpin probing of (CAG) <sub>7</sub> - and (CAG) <sub>14</sub> -8-oxoG hairpin substrates by S1 Nuclease.....	98
Figure 2.3 Hairpin probing of (CAG) <sub>7</sub> - and (CAG) <sub>14</sub> -THF hairpin substrates by Mung Bean Nuclease.....	99
Figure 2.4 Hairpin probing of (CAG) <sub>7</sub> - and (CAG) <sub>14</sub> -THF hairpin substrates by S1 Nuclease.....	100
Figure 2.5 Attenuation of CAG repeat expansion through repair of 8-oxoG in the hairpin loop and Mus81/Eme1 cleavage .....	101
Figure 2.6 Attenuation of CAG repeat expansion through repair of an abasic site in the hairpin loop and Mus81/Eme1 cleavage .....	104
Figure 2.7 Short-patch BER of (CAG) <sub>7</sub> - and (CAG) <sub>14</sub> -8-oxoG hairpin containing substrates.....	105
Figure 2.8 Mus81/Eme1 cleavage of a hairpin .....	109
Figure 2.9 Flap probing of (CAG) <sub>3</sub> /(CAG) <sub>4</sub> double-flap substrate.....	110
Figure 2.10 Flap probing of (CAG) <sub>7</sub> double-flap substrate .....	111
Figure 2.11 Processing of a CAG repeat double-flap intermediate during BER prevented or attenuated CAG repeat expansion.....	112
Figure 2.12 Mus81/Eme1 cleavage of the upstream flap of a CAG repeat double-flap intermediate .....	113
Figure 2.13 FEN1 processing of a hairpin during repair of an 8-oxoG in the hairpin loop .....	117
Figure 2.14 FEN1 processing of a hairpin during repair of an abasic site in the hairpin loop.....	118

Figure 2.15 FEN1 cleavage of CAG repeat double-flap intermediates.....	119
Figure 2.16 FEN1 cleavage of a hairpin.....	120
Figure 2.17 Prevention or attenuation of CAG repeat expansion by BER coupled with removal of a hairpin .....	124
 CHAPTER 3	
Figure 3.1 CTG repeat deletion via pol $\beta$ bypass of a 5',8-cdA in a (CAG) <sub>10</sub> tract that is opposite to a 1-nt gap during BER.....	147
Figure 3.2 CTG repeat deletion via pol $\beta$ bypass of a 5',8-cdA in a (CAG) <sub>10</sub> tract that is opposite to a 1-nt gap during DNA replication .....	148
Figure 3.3 Pol $\beta$ bypass of 5',8-cdA in a (CAG) <sub>10</sub> tract at downstream of a 1-nt gap during BER .....	149
Figure 3.4 Pol $\beta$ bypass of 5',8-cdA in a (CAG) <sub>10</sub> tract at downstream of a 1-nt gap during DNA replication .....	150
Figure 3.5 Pol $\beta$ bypass of 5',8-cdA in a (CAG) <sub>10</sub> tract at upstream of a 1-nt gap during BER .....	151
Figure 3.6 Pol $\beta$ bypass of 5',8-cdA in a (CAG) <sub>10</sub> tract at upstream of a 1-nt gap during DNA replication .....	152
Figure 3.7 Pol $\beta$ bypass of a 5',8-cdA located in a (CAG) <sub>10</sub> tract opposite to a 6-nt gap during DNA replication.....	153
Figure 3.8 Pol $\beta$ bypass of a 5',8-cdA located in a (CAG) <sub>10</sub> tract in the open template during DNA replication .....	154
Figure 3.9 Formation of a loop on the template strand of an intact double-strand DNA.....	161
Figure 3.10 Formation of a loop on the template with a 5',8-cdA lesion opposite to 1-nt gap in the absence of pol $\beta$ .....	162
Figure 3.11 Formation of a loop on the template with a 5',8-cdA lesion opposite to 1-nt gap in the presence of pol $\beta$ .....	163
Figure 3.12 Formation of a loop in the template strand of the substrates with a 1-nt gap located at the upstream of a 5',8-cdA in the absence of pol $\beta$ .....	164

Figure 3.13 Formation of a loop in the template strand of the substrates with a 1-nt gap located at the upstream of a 5',8-cdA in the presence of pol $\beta$ .....	165
Figure 3.14 FEN1 cleavage during pol $\beta$ bypass of a 1-nt gap opposite to a template 5',8-cdA .....	168
Figure 3.15 FEN1 cleavage during pol $\beta$ bypass of a 1-nt gap on the upstream of a template 5',8-cdA .....	169
Figure 3.16 CTG repeat deletion during pol $\beta$ bypass of a template 5',8-cdA.....	171

## ABBREVIATIONS AND ACRONYMS

°C	degree Celsius
APE1	AP Endonuclease 1
AP site	apurinic/apyrimidinic site
BER	base excision repair
bp	base pair
BSA	bovine serum albumin
C	cytosine
cdPu	5',8-cyclo-2'-deoxypurines
CPDs	cyclobutane–pyrimidine dimers
CSA	Cockayne syndrome proteins A
CSB	Cockayne syndrome proteins B
DMSO	dimethyl sulfoxide
DNA-PKcs	DNA-PK catalytic subunit
dNTP	deoxyribonucleotide triphosphate
DSB	double-stranded DNA breaks
DSBR	double-strand break repair
dsDNA	double stranded DNA
DTT	dithiothreitol
8-oxoG	8-oxoguanine
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	ethylenediaminetetraacetic acid

ERCC1	excision cross complementing repair 1
5',8-cdA	5',8-cyclo-2'-deoxyadenosine
5',8-cdG	5',8-cyclo-2'-deoxyguanosine
5'-dRP	5'-deoxyribose phosphate
FEN1	flap endonuclease 1
FPLC	fast protein liquid chromatography system
5mC	5-methylcytosine
G	guanine
GGR	global genome repair
HD	Huntington's disease
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HMGB1	high-mobility group protein B1
HO <sup>•</sup>	hydroxyl radical
HR	homologous recombination
ICL	interstrand cross link
IPTG	isopropyl β-D-1-thiogalactopyranoside
IR	ionizing radiation
LB	lysogeny broth
LIG I	DNA ligase I
LP-BER	long-patch BER
MD	myotonic dystrophy
MEF	mouse embryonic fibroblast



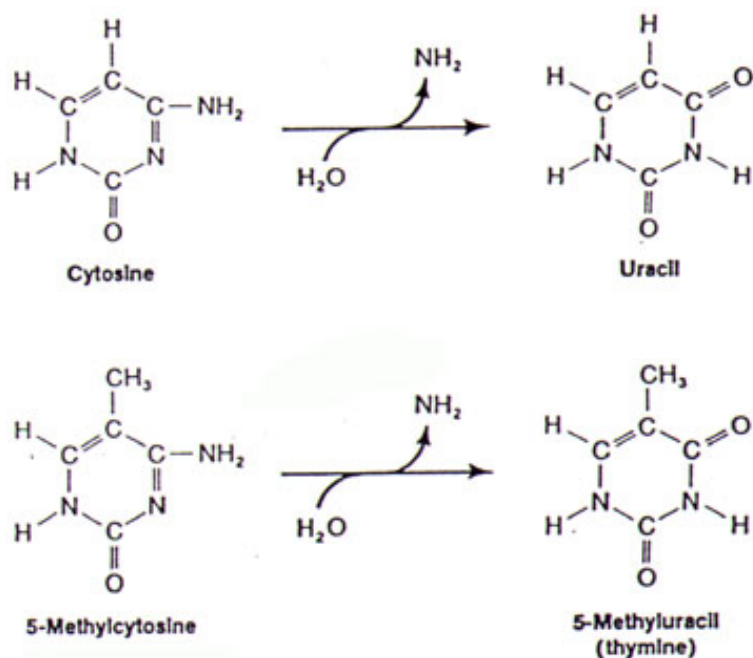
MgCl <sub>2</sub>	magnesium chloride
MMR	mismatch repair
MSH	MutS homologue
NER	nucleotide excision repair
NHEJ	nonhomologous end joining
NP-40	Nonidet P-40
nt	nucleotide
NTH1	endonuclease III-like protein 1
OGG1	8-oxoguanine DNA Glycosylase
PAGE	polyacrylamide Gel Electrophoresis
PCNA	proliferating cell nuclear antigen
PCR	polymerase chain reaction
PMSF	phenylmethylsulfonyl fluoride
poly-Q	polyglutamines
RNS	reactive nitrogen species
ROS	reactive oxygen species
RPA	replication protein A
RNAPII	RNA polymerase II complex
6-4PPs	6-4 photoproducts
SDS	sodium dodecyl sulfate
SN-BER	single-nucleotide BER
SP-BER	short patch base excision repair

ssDNA	single-stranded DNA
T	thymine
Taq polymerase	<i>Thermusaquaticus</i> polymerase
TC-NER	transcription-coupled nucleotide excision repair
TCR	transcription coupled repair
THF	tetrahydrofuran
TLS	translesion synthesis
TNR	trinucleotide repeats
XPA	xeroderma pigmentosum complementation group A
XPD	xeroderma pigmentosum complementation group D
XPF	xeroderma pigmentosum complementation group F
XPG	xeroderma pigmentosum complementation group G
XRCC1	X-ray repair cross-complementing protein 1
U	uracil
UDG	uracil DNA glycosylase
UV	ultraviolet

## INTRODUCTION

### A. Oxidative DNA damage and repair

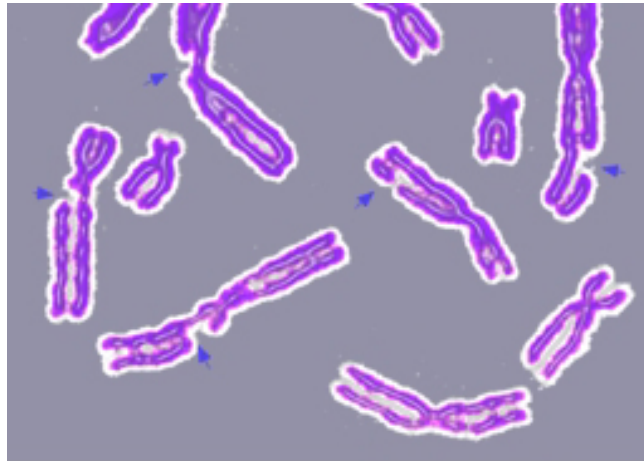
Deoxyribonucleic acid (DNA) is the essential molecule that is responsible for genetic information storage in living cells. Thus the stability and integrity of DNA is critical for passing accurate and faithful information from parental cells to daughter cells. However, DNA is chemically unstable and is constantly under spontaneous decomposition (15). The N-glycosidic bond that links a DNA base and a 2'-deoxyribose is labile and vulnerable to spontaneous hydrolysis, and the feature in turn results in spontaneous base loss under physiological conditions (15). Loss of a DNA base can further destabilize the backbone of DNA and cause DNA strand breaks through  $\beta$ -elimination leading to DNA degradation (16-18). Another major source of DNA degradation under physiological condition is DNA hydrolytic deamination. Among all four nucleotides, cytosine (C) and 5-methylcytosine (5mC) are highly susceptible to deamination, which results in C to Uracil (U) and 5mC to thymine (T) conversions (Figure I.1) (19-21). The DNA degradation caused by spontaneous base loss is defined as DNA self-decomposition.



**Figure I.1 Deamination of Cytosine (C) and 5-Methylcytosine (5mC) (2)**

Except for DNA self-decomposition resulting from the instability of the glycosidic bond and its spontaneous hydrolysis, modifications to DNA bases and DNA backbone recognized as DNA damage are also frequently induced by attack from the endogenous and environment stresses (15, 22, 23). For example, ultraviolet (UV) radiation from sunlight can directly result in the formation of bulky DNA adducts such as cyclobutane–pyrimidine dimers (CPDs) and 6–4 photoproducts (6–4PPs) thymine dimers as well as T-T thymine dimers (25). Ionizing radiation (IR) is another major source that causes intensive DNA modifications and strand breaks. A low dose of IR can induce production of reactive oxygen species (ROS) and cause oxidative DNA base lesions (26). A low level of IR can also directly attack DNA to induce a covalent linkage between nitrogenous bases as well as create a thymine dimer (27). However, a large dose of IR can induce single-stranded DNA (ssDNA) breaks and double-stranded DNA breaks (DSB)

by generating free radicals through its reaction with water (28-30). Toxins and pollutants from the environment such as tobacco smoke and bromate can result in bulky DNA adducts, oxidative DNA damage and alkylating DNA damage (31). Moreover, oxidative stresses caused by the products and byproducts from endogenous metabolism such as ROS and reactive nitrogen species (RNS) can induce oxidative DNA damage (15,22,23). Chemotherapeutic drugs that are used to treat cancer can also attack DNA and result in DNA damage including bulky DNA adducts and interstrand cross-link (ICL) (26,32). Therefore a variety of DNA damage from different sources is generated in living organisms. The number of DNA lesions in cells varies along with cell types and genome size in different organisms with an increasing number of DNA lesions correlated with increased size of the genome (33). It is estimated that approximately  $10^6$  DNA lesions occur in a single human cell per day, among which about  $10^5$  DNA lesions are caused by endogenous stresses (34). Accumulation of severe DNA damage such as DSBs in one chromosome can result in chromosome breakage (Figure I.2) that can ultimately lead to cancer.

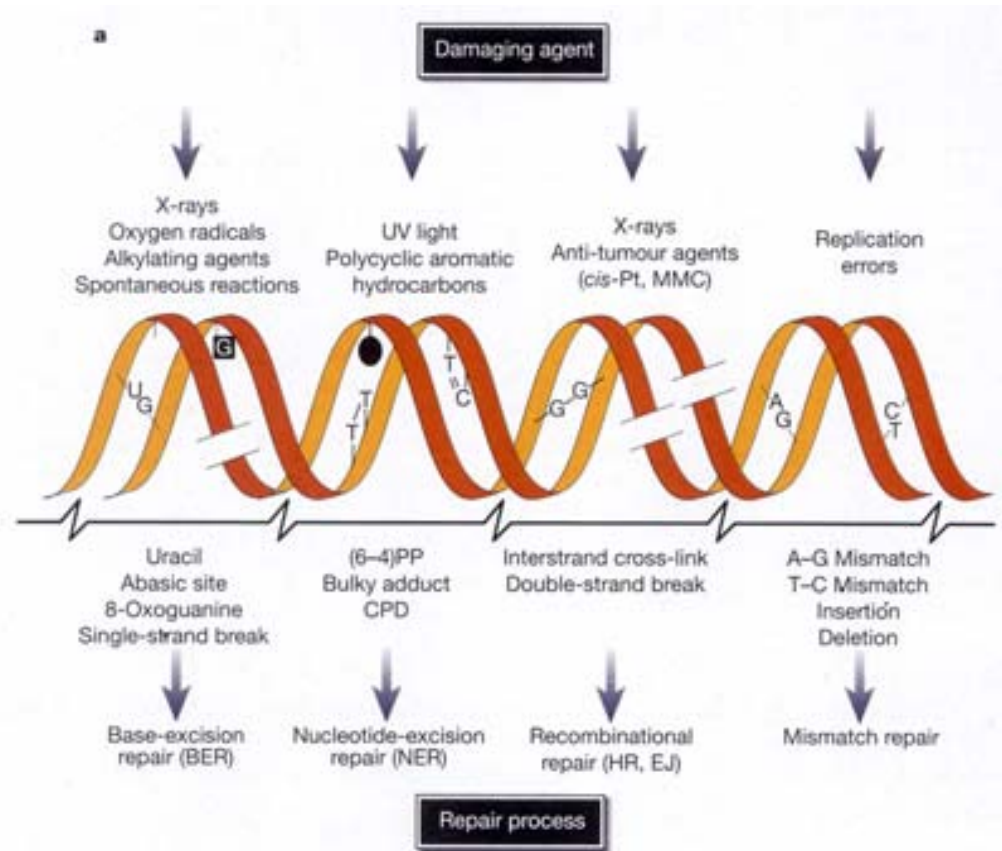


**Figure I.2 Chromosome breaks by DNA damage (1)**

Deoxyribonucleic acid lesions can also alter DNA structure by inducing DNA double helix distortion and DNA breaks resulting in inhibition of activity of replication DNA polymerases. Altered DNA structures can further prevent passage of genomic information from parent cells to daughter cells. On the other hand, RNA polymerase synthesis can also be blocked by DNA damage. The process suppresses some essential gene transcription and expression (35). Prevention of DNA replication and gene transcription by DNA damage can disrupt normal cellular function leading to cell death. Some DNA lesions are highly mutagenic and cause mutations in the genome, thereby altering genetic information that is passed from parent cells to daughter cells. Mutations that occur in the genes encoding essential enzymes or proteins can influence cell viability. In addition, some portion of DNA damage may induce mutations that alter the expression of oncogene or tumor suppressor gene, therefore lead to cancer development. Thus, unrepaired DNA damage leads to accumulation of gene mutation and DNA strand breaks resulting in the development of different types of human diseases such as cancer and aging (36). For example, excessive exposure to UV from sunlight can result in thymine

dimer, which is associated with skin cancer. Exposure to tobacco smoke induces DNA damage that in turn causes mutations in lung cells leading to development of lung cancer (34). Furthermore, accumulation of DNA damage in non-dividing cells such as neurons can result in apoptosis, thereby leading to neurodegeneration and aging (37,38).

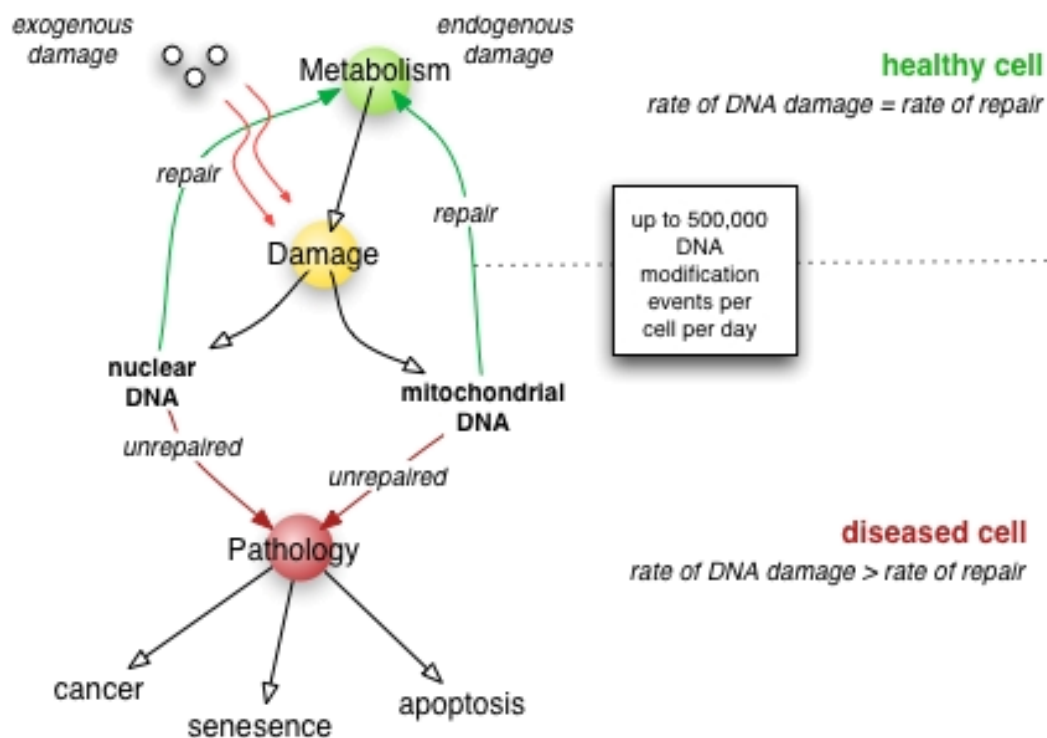
To maintain the stability and integrity of genetic information, cells have developed robust DNA repair mechanisms to remove DNA damage (Figure I.3). These include DNA base excision repair (BER), nucleotide excision repair (NER), mismatch



**Figure I.3 DNA damage, repair mechanisms and consequences (9)**

repair (MMR) and double-strand break repair (DSBR), which includes homologous recombination (HR) and nonhomologous end joining (NHEJ).

Different types of DNA damage are subject to different DNA repair pathways depending on the unique structural features of each type of DNA lesions. Base excision repair is responsible for removal of small base lesions that are chemically modified and do not induce DNA double helix distortion. The base lesions that are removed by the BER pathway include oxidized bases, alkylated bases, deaminated bases, a uracil that usually results from deamination from a cytosine or mis-incorporation by DNA polymerases as well as ssDNA breaks (39-42). NER is responsible for removal of bulky DNA damage that causes DNA double helix distortion such as CPD and thymine dimer. Mismatch repair is a repair pathway that specifically removes mis-matched nucleotide caused by replication DNA polymerase error as well as small loops with a size from 1-6 nt. The DDR pathway that includes HR and NHEJ is responsible for repairing the most severe DNA damage, DDB.



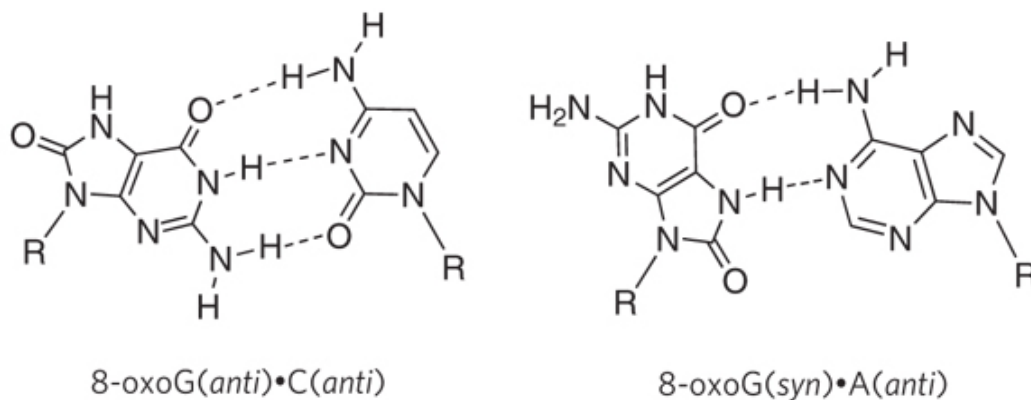
**Figure I.4 DNA repair rate is an important determinant of cell pathology (3)**



If the amount of DNA damage generated exceeds the capacity of normal cellular DNA repair, DNA damage that accumulates in cells will further lead to early senescence, apoptosis, or cancer (Figure I.4). On the other hand, deficiency of DNA repair enzymes and cofactors can also lead to early onset of aging and an increased risk of cancers (36, 38, 43). For example, transgenic mice with a deficiency of DNA-PK catalytic subunit (DNA-PKcs) are susceptible to lymphoma and infections and exhibit a shorter lifespan than wild-type mice (44). Similarly, mice with a deficiency in NER protein, xeroderma pigmentosum complementation group D (XPD) and xeroderma pigmentosum complementation group A (XPA) also exhibited the phenotypes of premature onset of aging and shortened lifespan (45, 46). Consequently, the deficiency of repair results in the accumulation of DNA damage that includes DNA crosslinks (interstrand or intrastrand crosslinks or ICLs) that ultimately lead to double-strand breaks (47). For unrepaired DNA damage, cells will adopt an error-prone DNA damage tolerance mechanism called translesion synthesis (TLS) as a backup to bypass the damaged site and allow DNA replication and gene transcription to be continued and maintained (48-51).

Oxidized base lesions are the most common form of DNA damage (52). Such type of lesions are usually created by hydrogen peroxide and ROS such as superoxide and hydroxyl radicals ( $\text{HO}^\bullet$ ) (52). These can be generated from both endogenous resources including cellular energy metabolism and inflammation and environment pollutants such as chromate, bromate, arsenate and IR (53). Accumulation of oxidative DNA damage can result in a series of adverse effects including mutations, abnormal transcription and epigenetic instability (53). Among the oxidative DNA damage detected in human cells, 7,8-dihydro-8-oxoguanine, also known as 8-oxoguanine (8-oxoG) is the most common

form that is generated by hydroxyl radical attack at the C8 position of guanine (G) because of its low redox potential (52). It has been found that 8-oxoG can not only form a normal Watson-Crick base pair with a C, but also form a Hoogsteen base pair with a A (Figure I.5). It was estimated that the ratio of 8-oxoG:C base pairing to 8-oxoG:A base pairing is almost 1:1 in the genome (54). An 8-oxoG:A mismatch leads to a misinsertion by DNA polymerases in bypassing of a 8-oxoG in the genome during replication and repair, thus creating mutations with a relatively high frequency (54). Thus, accumulation of unrepaired 8-oxoG in human genome is highly mutagenic. Moreover, accumulation of oxidative DNA damage can occur at different locations in the genome, therefore affect gene transcription and regulation in a damage location dependent manner (55-57), and this may ultimately cause human cancer (54).



**Figure I.5 8-oxoG base pair with dC and dA (7)**

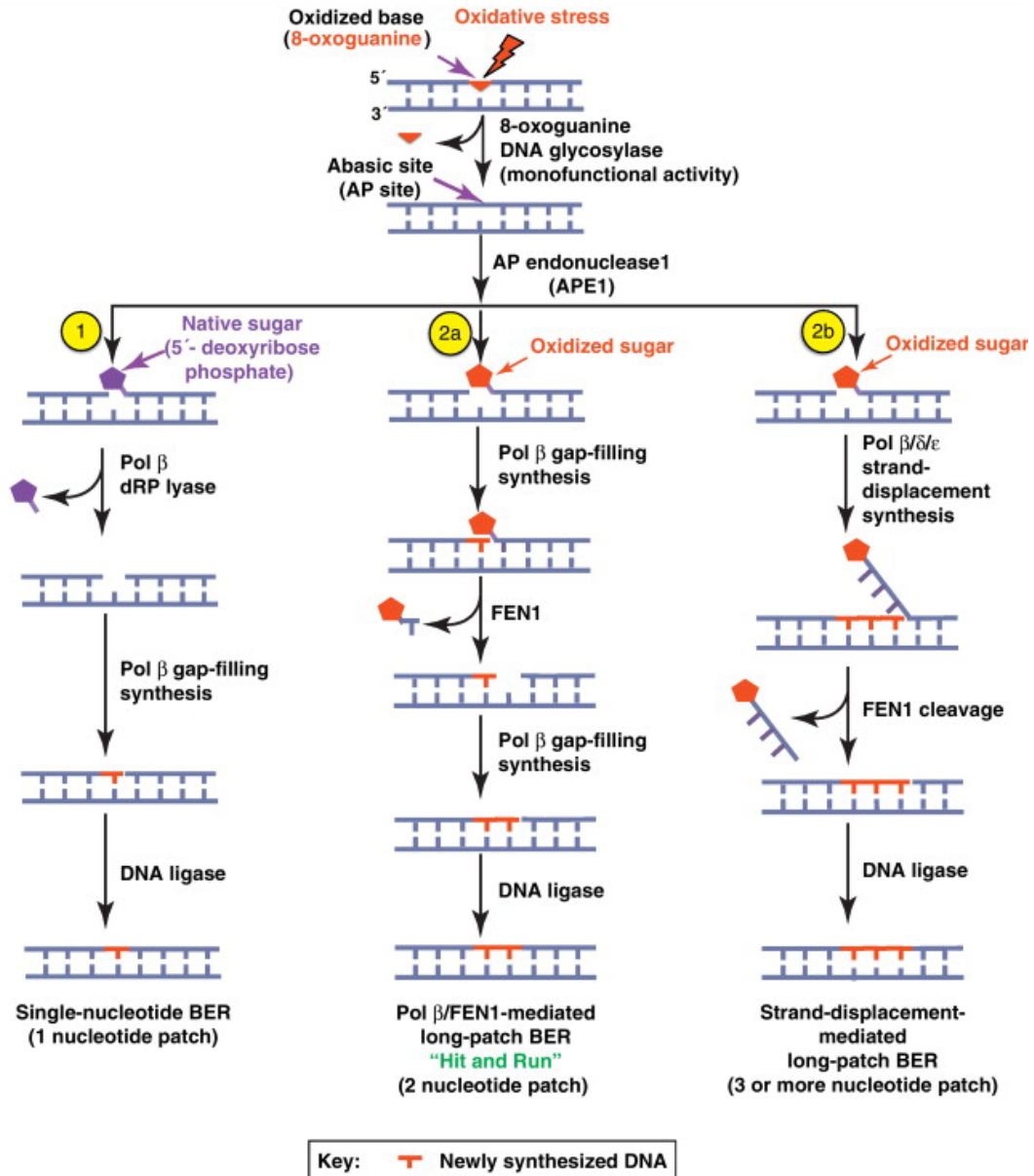
In order to protect the cells from the adverse effects of oxidative DNA damage and maintain genome stability and integrity, it is essential for cells to develop a robust DNA repair mechanism that removes oxidized DNA base lesions. As the major pathway to remove oxidized DNA base lesions, BER plays a crucial role in removal of oxidative

DNA damage with a high efficiency. It has been demonstrated that BER is recognized as the most efficient DNA repair pathway among all of DNA repair pathways identified thus far. It has been found that removal of a single oxidative DNA base lesion by one round of BER only takes a few minutes. Under a severe oxidative DNA damage condition, several rounds of BER are required, and it takes several hours to remove all base lesions because of the limited number of BER enzymes in human cells (58,59). In mammalian cells, there are two sub-pathways of BER, short-patch BER (SP-BER), also known as single-nucleotide BER (SN-BER), and long-patch BER (LP-BER). The two sub-pathways differ in the number of newly synthesized nucleotides that are used to remove a base lesion as well as the repair enzymes and cofactors that participate in the repair (60-63). SP-BER only involves the replacement of a single nucleotide. However, LP-BER involves the replacement of two or more nucleotides (60-65). Whether a base lesion is subject to SP-BER or LP-BER depends on the type of DNA lesions. A DNA base lesion is usually subject to SP-BER. A modified abasic lesion or a modified deoxyribose residue such as an oxidized abasic site or sugar can only be subject to LP-BER (11).

Both SP-BER and LP-BER are initiated by recognition and removal of DNA base lesions by DNA glycosylases. During repair of an 8-oxoG, 8-oxoguanine DNA glycosylase (OGG1) first binds to an 8-oxoG lesion, cleaves the N-glycosidic bond and removes the damaged base, leaving an apurinic/apyrimidinic site (AP site). Subsequently, AP endonuclease 1 (APE1) incises the AP site at its 5'-side, resulting in a single-stranded DNA break with a downstream 5'-deoxyribose phosphate (5'-dRP) or oxidized 5'-sugar phosphate if the deoxyribose is also oxidized by the oxidative stress. A native 5'-dRP residue is subject to the SP-BER sub-pathway, whereas an oxidized 5'-sugar phosphate

needs to be repaired by the LP-BER sub-pathway. Removal of a native 5'-dRP is accomplished by the dRP lyase activity of DNA polymerase  $\beta$  (pol  $\beta$ ), leaving a 1-nt gap that is then filled by pol  $\beta$  polymerase activity. Pol  $\beta$  gap-filling synthesis results in a nicked DNA that is sealed by a DNA ligase. In this scenario, only one nucleotide is replaced. On the other hand, an oxidized 5'-sugar phosphate is resistant to the dRP lyase activity of pol  $\beta$  (Figure I.6) (11). Removal of the oxidized sugar phosphate is dependent on cleavage of flap endonuclease 1 (FEN1). Long-patch BER can be accomplished by the "hit-and-run" mechanism where pol  $\beta$  inserts 1 nt to fill in the 1-nt gap, and subsequently FEN1 removes the oxidized sugar attached to a downstream 1 nt leaving another 1-nt gap. DNA polymerase  $\beta$  then performs its gap-filling synthesis activity again to fill in the gap leaving a nick for DNA ligase to seal. During the pol  $\beta$ /FEN1 mediated LP-BER, only two nucleotides are replaced (Figure I. 6) (11). Long-patch BER can also be achieved by a DNA strand-displacement synthesis mechanism. In the strand-displacement mediated LP-BER, pol  $\beta$  or replication DNA polymerases, pol  $\delta/\epsilon$ , perform DNA strand-displacement synthesis to insert three or more nucleotides and displace a downstream strand to create a 5'-flap. Flap endonuclease 1 removes the flap resulting in a nick for DNA ligase I to seal. The strand-displacement mediated LP-BER subpathway involves 3 or more nucleotides replacement. (Figure I.6) (11). It appears that the "Hit and Run" mechanism which only involves two nucleotides of replacement is more efficient than the "strand-displacement synthesis" mechanism which involves three or more nucleotides of replacement. In the scenario, the "Hit and Run" mechanism that is mediated by a sequential coordination between pol  $\beta$  and FEN1, is the most efficient LP-BER. The coordination between sequential BER enzymatic reactions is named the "Passing the

Baton" mechanism (66). It is proposed that the coordination may help prevent the exposure of the BER intermediates to the cellular environment to induce nonspecific

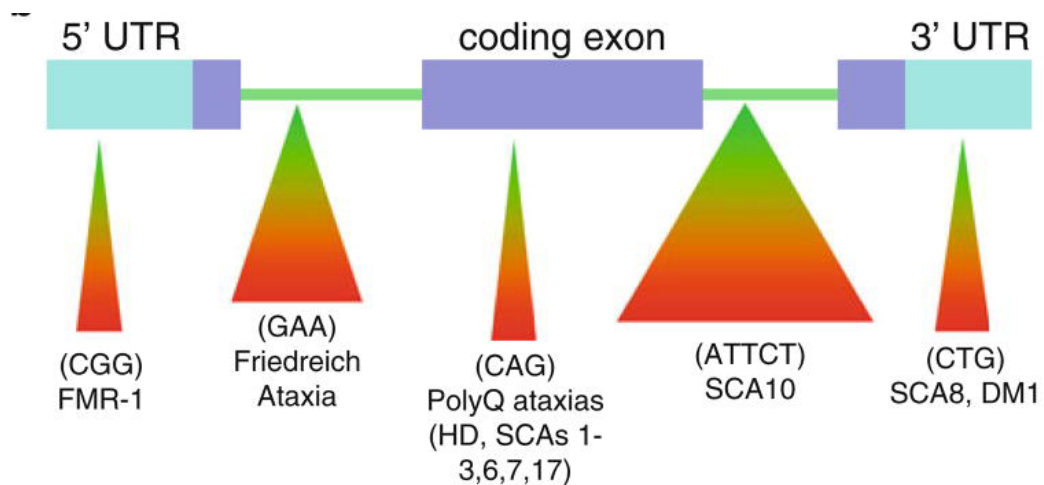


**Figure I.6 Base excision repair (BER) of oxidized DNA base lesions (11).**

nuclease cleavage, DNA recombination and cell death signaling, thereby maintaining genome stability and integrity (66-69).

## B. Trinucleotide repeat (TNR) instability and human diseases

Trinucleotide repeats (TNR) belong to microsatellites that are composed by stretches of tandem repeated sequence with one to six nucleotides. They constitute 30% of the human genome (70). Similar to the other microsatellite repeated sequences, TNRs are very unstable and highly polymorphic in human genome (71, 72). The TNR length changes with an increasing or decreasing number of repeat units is recognized as a new type of mutation, i.e., a dynamic mutation (73). It has been found that TNR expansion is the cause of more than 40 types of human neurodegenerative diseases (4,6,74). These diseases include Huntington's disease (75) induced by  $(CAG/CTG)_n$  expansion, myotonic dystrophy (MD) caused by  $(CTG/CAG)_n$  expansion, fragile X syndrome caused by



**Figure I.7 Trinucleotide repeat expansion disorders caused by triplet repeats in coding and noncoding gene regions (6)**

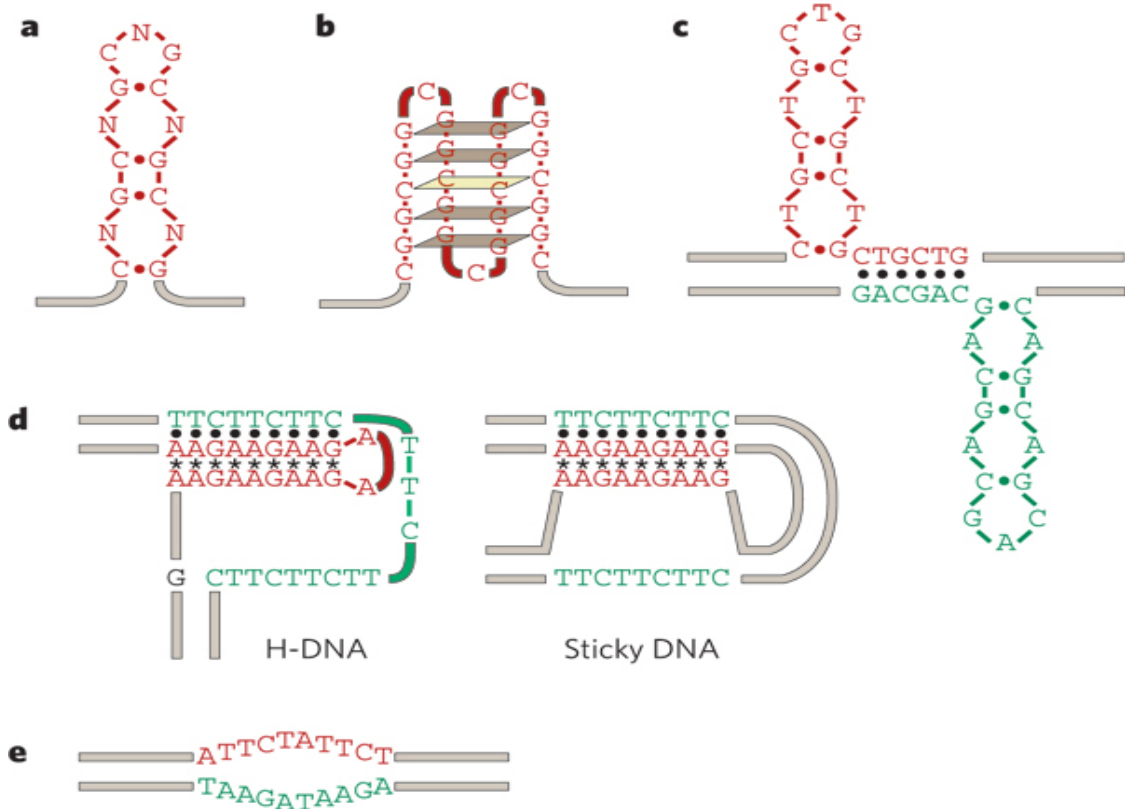
$(CGG/GCC)_n$  expansion and Friedreich's ataxia resulting from expansion of  $(GAA/CTT)_n$  (Figure I.7) (4,6,76). Expansion of TNR located in the encoding region such as CAG repeats can lead to the expression of abnormal protein, a mutant huntingtin protein containing an expanded chain of polyglutamines (poly-Q), thereby resulting in

aggregation and precipitation of huntingtin proteins along with many other proteins in neurons, further causing apoptosis and death of neuronal cells and neurodegeneration (77,78). The TNR expansion can also occur in non-encoding regions. For example, GAA repeats are located in an intron of the frataxin gene. Expansion of GAA repeats disrupts the extension of transcription of the gene by RNA polymerase, suppressing the expression of frataxin protein, which plays an important role in maintain normal level of energy metabolism. Reduction of frataxin protein in neuron cells further results in neuronal cell death via deficiency of energy metabolism (77,79,80). It has been found that TNR expansion occurs if repeat length reaches a certain length that ranges from 30-35 repeat units, which is called a threshold length for TNR expansion. If the size of TNR is over the threshold length, both the rate and the probability of TNR expansion increase (77,78). It has also been found that TNR expansion that occurs in encoding region is usually modest with an increase of  $\leq 10$  repeat units per generation of cells during parent-child transmission. However, TNR expansion in a non-encoding region is robust and can involve an increase of repeats units ranging from 100 to 10,000 units (12,77,78,81). The longer repeats can result in more severe disease symptoms and an earlier disease onset, the phenomena is named as anticipation (4,6,12).

Besides its expansion, TNR also exhibits deletion/contraction during DNA replication and repair. It has been found that TNR deletion/contraction such as CAG repeat contraction in the encoding region of the androgen receptor is associated with development of prostate cancer and ovarian cancer (82,83). Moreover, it has been found that the frequency with which CAG and CGG repeats are present in oncogenes is 2.5-fold higher than in other genes in the human genome. In addition, the size of oncogenes in

tumor cells is much longer than in the normal cells (84). This suggests that expansion of TNR in oncogenes disrupts the gene function and expression and promotes oncogenesis (84).

While the mechanisms underlying TNR instability remain to be elucidated, it has been proposed that the expansion and deletion of TNR is mediated by the formation of a series of non-B form DNA secondary structures by TNR self-base pairing (38-40). These include hairpins, triplexes and tetraplexes (Figure I.8) (85,86). For example, a CAG/CTG



**Figure I.8 Non-B form DNA structures formed by TNR (4)**

and CGG/GCC repeat tract can form thermodynamically stable hairpin and tetraplexes structures by G:C or G:G self-base pairing. With a longer repeat tract, more stable secondary structures can be formed (87,88). CGG/CCG repeats can also form tetraplexes

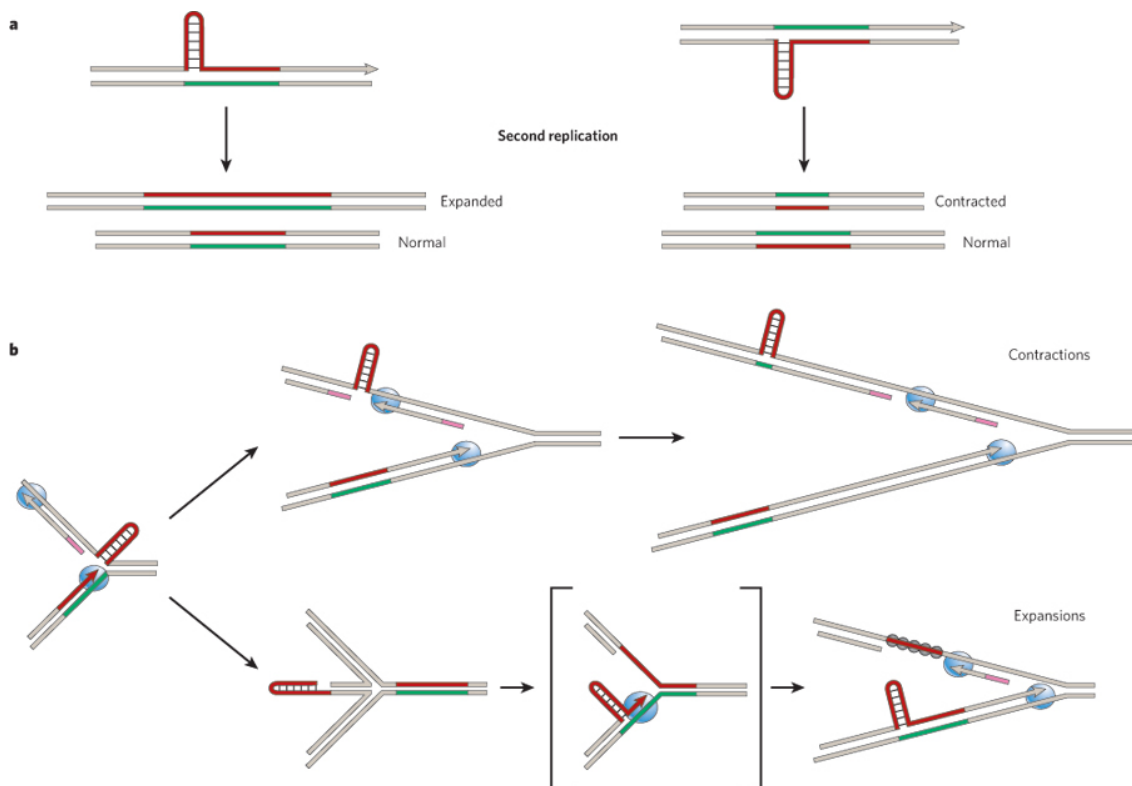


such as G-quadruplexes resulting from Hoogsteen base pairing formed among four Gs (87,88). A GAA/TTC repeat tract can result in the formation of a secondary structure called a triplex with three DNA strands winding around each other to form a triple helix. In this special structure, the third DNA strand can base pair with one strand of a B-form DNA double helix through Hoogsteen or reversed Hoogsteen hydrogen bonds (89). These secondary structures usually form during DNA metabolism, including DNA replication and repair (4,12).

### **C. TNR instability and DNA replication and RNA transcription**

In dividing cells, TNR expansion and deletion accumulate along with each cycle of DNA replication (4,6,90), indicating that DNA replication can promote TNR instability. Several studies have shown that a long tract of (CTG/CAG)<sub>n</sub> in a plasmid transformed into *E. coli* exhibited repeat length change through cell generations of *E. coli* (91). It was found that CTG repeat instability was biased towards repeat deletion in *E. coli* (91), and the position of replication origin relative to the repeat region played an essential role in determining the consequence of repeat expansion or deletion (91,92). Interestingly, in *E. coli*, deletions of (CTG/CAG)<sub>n</sub>, (CGG/CCG)<sub>n</sub>, and (GAA/TTC)<sub>n</sub> repeats preferentially occurred during DNA leading strand synthesis. In contrast, DNA lagging strand synthesis results in repeat expansion (93). The length of the repeat tract has also been found to play a crucial role in determining the outcome of TNR expansion or deletion during DNA replication (92). It has been proposed that secondary structures such as hairpins formed on the newly synthesized strand lead to TNR expansion, whereas the formation of secondary structures on the template strand causes TNR deletion. This

notion has been also confirmed by several studies with bacteria and yeast systems (94-96). The molecular mechanisms underlying TNR instability during DNA replication are proposed in the model shown in Figure I.9. In general, during DNA replication, first a DNA double strand helix is unwound to into two single-stranded templates that are annealed with RNA primers initiating new DNA strand synthesis. However, the single-stranded long tract TNR region readily forms secondary structures resulting in the misalignment between newly synthesized strand and the parental template strand. If the non-B form secondary structures formed on the newly synthesized strand, extra repeat units will be produced and integrated to fulfill the replication process, resulting in TNR expansion. On the other hand, if the secondary structures are present on the template

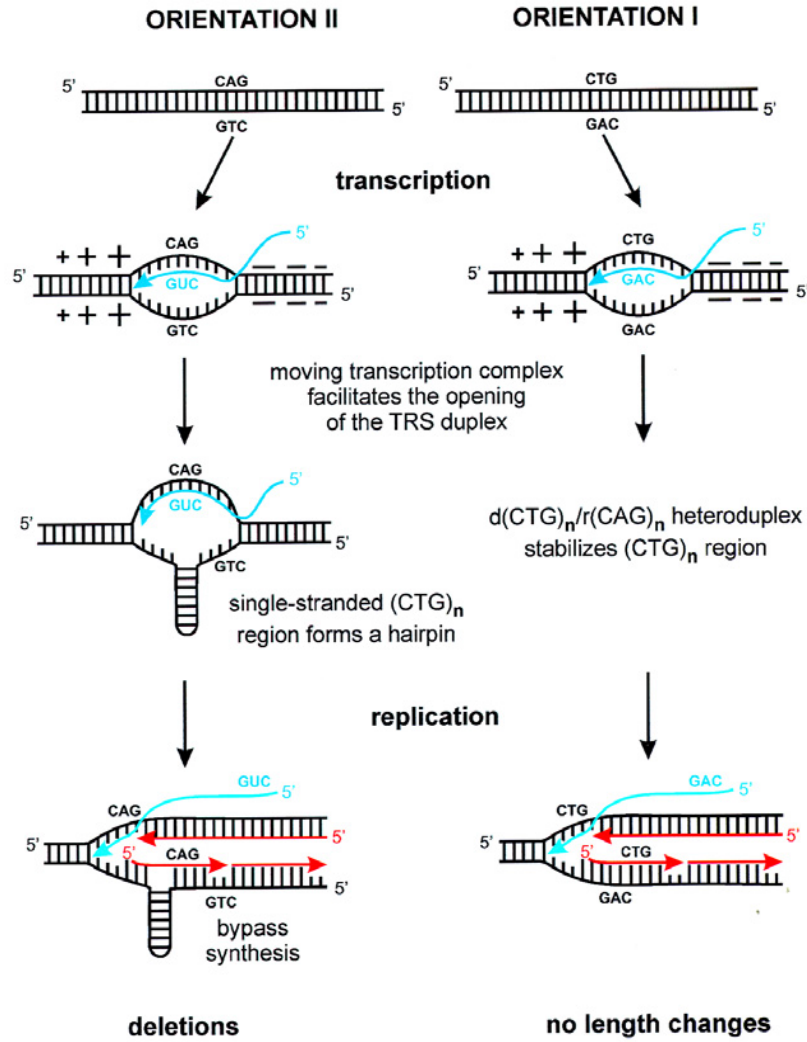


**Figure I.9 Replication mechanisms for repeat instability (4)**

strands, some repeats can be skipped over by DNA polymerase during replication resulting in TNR deletion (8). On some occasions, secondary structures formed on the template strands can block replication polymerase synthesis and result in stalling of replication fork and disassociation of replication enzymes (86,97,98). Reassembly of replication machinery on a stalled replication fork on the lagging strand synthesis results in the formation of DNA secondary structures such as hairpins, loops on the template strand or the newly synthesized strand, which lead to TNR deletion or expansion.

Interestingly, it has been found that all TNRs that result in neurodegenerative diseases are located in an actively transcribed region (Figure I.7) (6), indicating that transcription of TNR may also play an important role in modulating TNR instability. During transcription, translocation of RNA polymerase complex along an unwound DNA double helix induces DNA positive supercoiling ahead of the enzyme and leaves negative supercoiling behind it (99). It has been shown that the energy generated by negative supercoiling may facilitate the formation of hairpin or other non-B form DNA structures in a TNR tract, thereby promoting TNR instability in *E. coli* (100,101). It has been found that the outcome of TNR instability is determined by the orientation of transcription (Figure I.10). For example, during the transcription of a CAG repeat-containing strand, CAG repeat deletion occurs. Because the complementary strand that contains a CTG repeat tract becomes single-stranded due to unwinding of a double strand DNA, allowing CTG repeats to form a hairpin structure. This further promotes DNA polymerases to skip over the hairpin leading to CAG repeat deletion (left side). On the other hand, with the transcription of CTG repeats, no repeat expansion or deletion occurs. It is because the newly synthesized RNA contains CAG repeats, and these repeats can base pair with the

template CTG repeats thereby preventing the formation of a hairpin on the template CTG repeat strand.



**Figure I.10 Correlation between replication, orientation of the (CTG/CAG)<sub>n</sub> tract, transcription, structural properties of repeated DNA and the genetic instability of the TNR (8)**

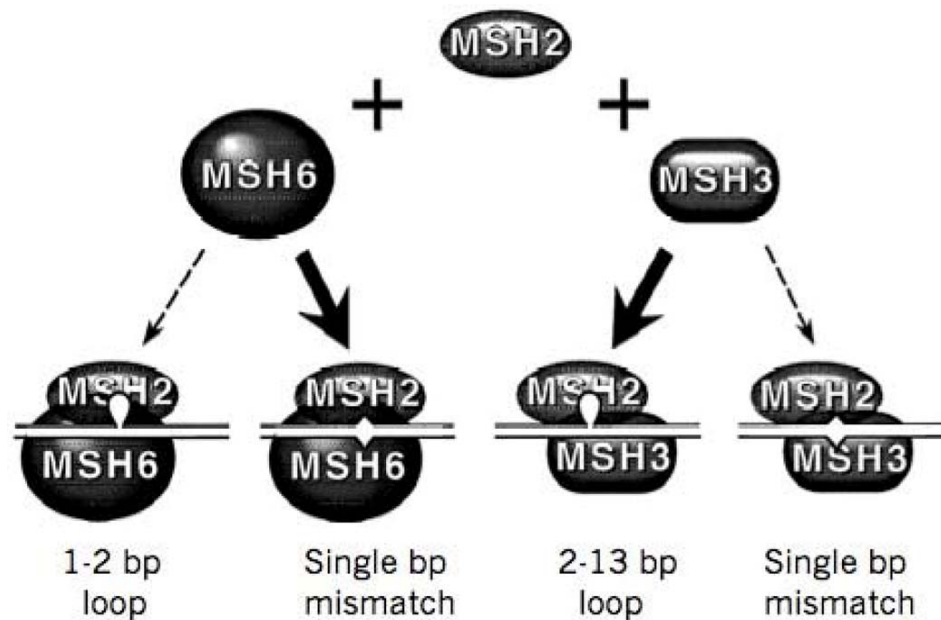
#### **D. TNR instability and DNA repair**

Although TNR instability was initially observed during DNA replication in the process of parental gene transmission and dividing somatic cells, more recently it has been found that TNR instability frequently occurs in non-dividing somatic cells (102,103). Presence of TNR instability in somatic cells indicates that TNR instability can also be induced during other DNA metabolic pathways other than DNA replication. In fact, the results from all studies have shown that TNR instability can be generated during any DNA metabolic pathways that involve formation of single-stranded DNA including replication (104,105), transcription (106), repair (11,12) and recombination (107). Because a single-stranded TNR containing region allows TNR self base-pairing that subsequently leads to the formation of non-B form DNA structures such as hairpins (85,140). The non-B form structures can further cause a series of adverse effects during DNA metabolism. For example, DNA replication polymerases cannot bypass tetraplexes formed during DNA replication and repair (108,109), thereby subsequently causing DNA polymerase pausing and DNA slippage (108,109) resulting in replication fork stalling and collapse (86,97,98). If a single-strand DNA break occurs during DNA lagging strand maturation and DNA repair, a downstream 5'-TNR-containing flap can fold back and form a hairpin structure that inhibits flap cleavage by FEN1 (110-112). In addition, a mismatch repair protein complex, MSH2/MSH3 can be trapped by hairpin structures (113,114) and stabilize the hairpin structure promoting TNR instability. Formation of a hairpin structure can also disrupt the coordination among DNA repair proteins that sustains an efficient DNA repair. For example, disruption of the coordination between pol  $\beta$  and FEN1 by the formation of a hairpin structure can promote pol  $\beta$  multi-nucleotide

gap-filling synthesis that produces extra CAG repeats for CAG repeat expansion (115). Thus, formation of a TNR secondary structure creates a challenging problem for cellular DNA repair mechanisms, and this in turn promotes the instability of TNRs. In fact, it was found that DNA repair is forced by TNR secondary structures to facilitate their integration in the genome and paradoxically, DNA repair is also found to modulate TNR instability by removing a TNR secondary structure (12).

The role of DNA repair has been implicated by the fact that TNR expansion in non-dividing neurons exhibits an age dependency (102,103). In neuronal cells, TNR exhibits a small expansion ranging from 5-20 repeat units (102,103). The small size of repeat expansion cannot be explained by DNA replication and recombination, which usually cause much larger TNR instability (116). The findings suggest that DNA repair plays a crucial role in modulating TNR expansion. Several DNA repair pathways have been shown to modulate TNR instability during both DNA replication and damage repair. The MMR pathway is the first repair pathway identified to play an important role in modulating TNR instability. The MMR pathway is mainly responsible for removing mismatched bases generated from a replication error as well as small loop structures (Figure I.11). However, recent studies have found that MMR plays a dual role in modulating TNR instability (117-119). On one hand, MMR can facilitate small CAG/CTG repeat expansion and deletion by binding and stabilizing small hairpin structures during replication and repair (117,118). On the other hand, MMR can remove small TNR loops composed of (CAG)<sub>1</sub>-(CAG)<sub>3</sub> by stabilizing TNR (119). The MMR system is made up of two types of heterodimeric MutS homologue (MSH) complexes, MSH2/MSH6 (MutS $\alpha$ ) and MSH2/MSH3 (MutS $\beta$ ). The two MSH complexes have a

different substrate preference (11,120,121). Although both MSH2/MSH6 and MSH2/MSH3 complexes can recognize a single base mismatch and small extra-helical loop, the MSH2/MSH6 complex prefers to recognize and repair a single base mismatch, whereas the MSH2/MSH3 complex specifically repairs small loops (120-125) (Figure I.11). Since CAG/CTG repeats can form hairpin or loop structures during DNA



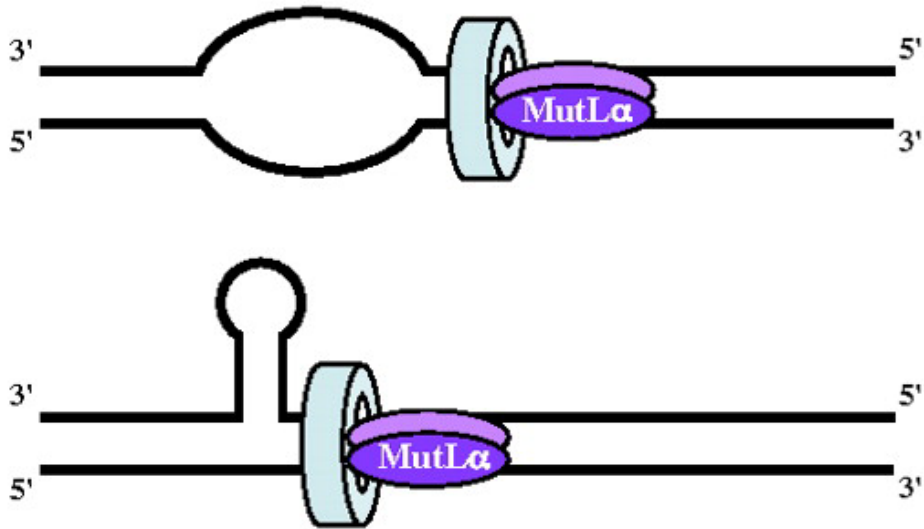
**Figure I.11 Eukaryotic mismatch repair complexes (11)**

replication and repair (126), these structures in the genome can be recognized and bound by MSH complexes, especially the MSH2/MSH3 complex. However, in HD transgenic mice, CAG repeat expansion depends on the presence of MSH2/MSH3 because CAG repeat expansion was abolished in the HD mice with MSH3 knock-out, but not in the mice with MSH6 knock-out (113,127), indicating that MSH2/MSH3 plays a crucial role in promoting CAG repeat expansion. It has been found that although the MSH2/MSH3 complex is able to induce removal of small extra-helix loops, a stable hairpin stem formed by CAG/CTG repeat can inhibit this process by trapping the protein complex

(128). It has been found that once MSH2/MSH3 complex binds to CAG hairpin, the A:A mismatch in the hairpin stem region inhibits the ATPase activity of the protein complex and alters the binding affinity of the complex to DNA, therefore prevents the protein complex from sliding off from the hairpin and repair of the hairpin by MMR and facilitate CAG repeat expansion (113). On the other hand, during DNA leading strand synthesis, the MSH2/MSH3 complex can bind and stabilize a hairpin formed on the template strand, facilitating DNA replication polymerase skipping over the hairpin leading to TNR deletion (93). Thus, the MSH2/MSH3 complex can bind and stabilize CAG/CTG repeat hairpin formed during DNA replication and repair pathways to facilitate TNR instability. In contrast, for a short CAG repeat tract, usually an unstable hairpin or loop may form. MSH2/MSH3 cannot bind to the small unstable hairpin or loop structure. Thus, in this scenario, MSH2/MSH3 can perform normal MMR activity to remove the unstable CAG repeat hairpin or loop and help to maintain TNR stability (119).

The complex of MSH2/MSH6 is not directly involved in facilitating TNR expansion. However the protein complex can induce ssDNA breaks during a process where it removes a mismatched nucleotide in the context of TNR. The ssDNA breaks induced by MSH2/MSH6 can further result in the formation of a hairpin structure that leads to repeat expansion (Figure I.12) (13).

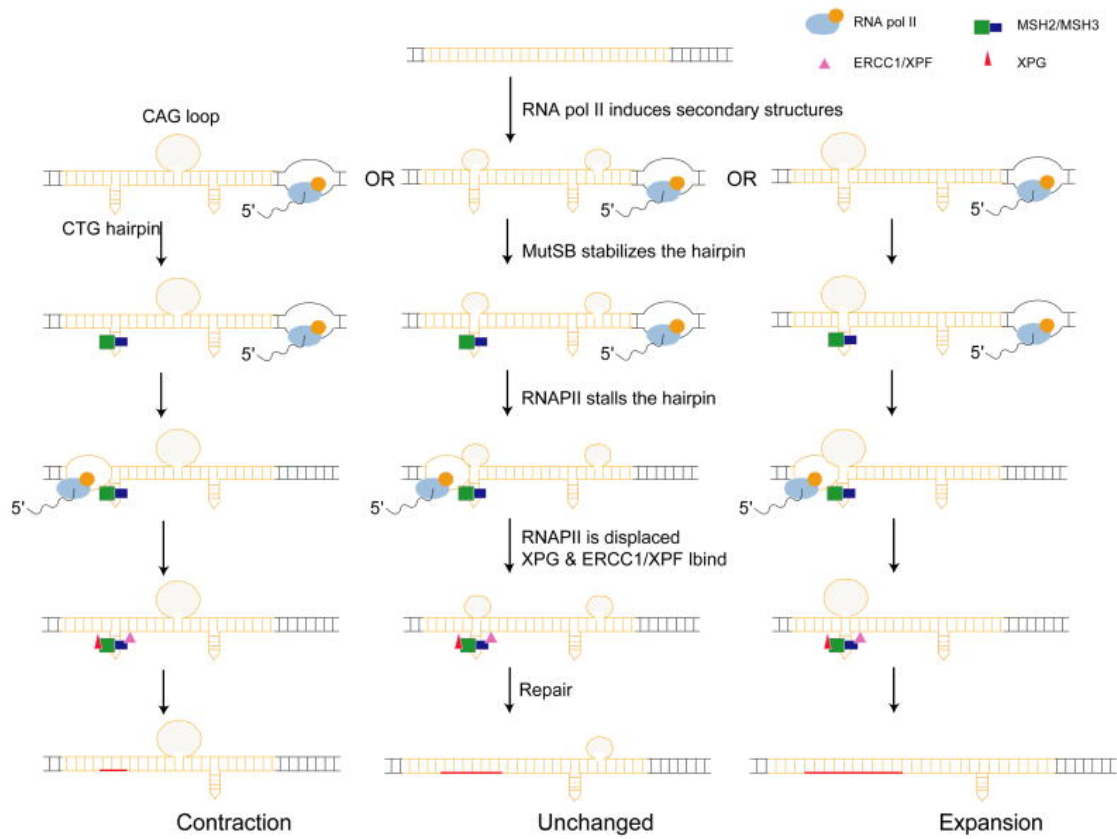




**Figure I.12 Mismatch repair induces hairpin formation (13)**

Nucleotide excision repair is another important repair pathway that is associated with TNR instability. Nucleotide excision repair is a major repair pathway that defends cells from DNA double helix distortion damages such as bulky DNA adducts and DNA intrastrand cross-links as well as some methylated DNA bases (45, 46). *In vitro* studies showed that some NER proteins are also involved in resolving bubble or loop structures in the genome (129,130). Because TNR-induced hairpins and loops can also induce double-strand DNA distortion, which may recruit NER enzymes to TNR hairpins and loops and facilitate TNR instability (101). Previous studies showed that TNR instability in *E. coli* was modulated by NER enzymes during transcription (101). In addition, NER in human cells and *Drosophila* also participates in transcription-induced repeat instability (131-133). A hypothetical model that demonstrates how NER may modulate TNR instability during transcription has been proposed as shown in Figure I.13 (5). During the first round of transcription of a CTG repeat-containing template, progression of RNA

polymerase II complex (RNAPII) leads to the formation of a single-stranded complementary strand that subsequently induces the formation of CTG and CAG repeat hairpins and bubbles in both the template strand and the complementary strand, respectively. Because MSH2/MSH3 readily binds to hairpin and loop structures, in the second round of transcription, the template hairpins block RNAPII progression and causes transcription to stall further triggering transcription-coupled nucleotide excision repair (TC-NER) to resolve the template hairpins so that transcription can be continued. In this scenario, NER proteins dislodge the RNAPII to access to the template hairpin and initiate the repair process. Depending on how a CAG repeat hairpin or loop in the template strand is rearranged in relation to a CTG hairpin or loop in the complementary strand, three types of repair consequences may occur, i.e., deletion, no change, and expansion of the template CTG repeats (5). If several CAG repeat small bubbles merge into a big bubble and migrate into downstream of the CTG hairpin bound by MSH2/MSH3, ERCC1/XPF and XPG remove the stabilized CTG hairpin subsequently forcing replicative DNA polymerase, pol  $\delta$  and pol  $\epsilon$  to use the complementary strand without bubbles as their template to fill in the gaps created by the removal of the hairpin. In this scenario, repaired products with repeat deletion are generated. If the small bubbles formed by CAG repeat stay at their original locations, after removal of CTG repeat hairpin by ERCC1/XPF and XPG, an unstable CAG repeat bubble can be unwound, and DNA polymerases can copy the entire length of the CAG repeat in a small bubble to fulfill the gap-filling synthesis leading to maintenance of the repeat length. If the CAG repeat small bubbles merge into a big bubble and migrate to the opposite side of the CTG hairpin stabilized by MSH2/MSH3, DNA polymerases will copy the entire CAG repeats

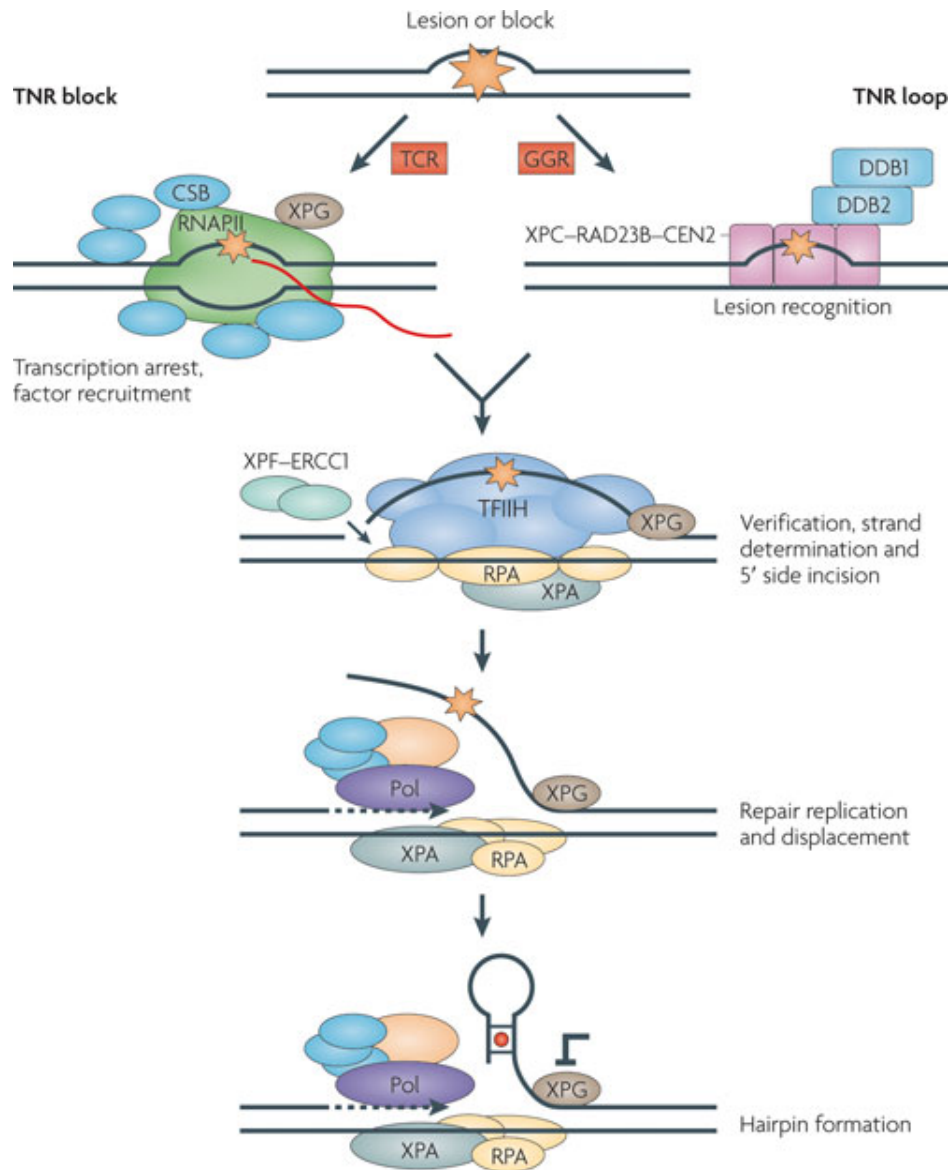


**Figure I.13 Speculative model for transcription-induced repeat instability (5)**

in the big bubble and synthesize extra CTG repeat units leading to CTG repeat expansion (5).

On some occasions, non-B form DNA structures can also be induced by ssDNA breaks via endonuclease cleavage during NER in the context of TNR repeats. In the scenario, TNR instability is directly facilitated by the NER pathway, and both sub-pathways of NER, global genome repair (GGR) that is the general repair pathway for correcting DNA lesions throughout the entire genome, and TC-NER or transcription coupled repair (TCR), can be involved in modulating of TNR instability (134,135). Except for the damage recognition step that initiates the repair, all other steps in the repair process are identical in these two sub-pathways. Global genome repair is initiated

by DNA-damage binding (DDB) and the XPC-Rad23B complex that is responsible for scanning the entire genome for strand distortion damage recognition (136), whereas TCR is initiated by a stalled RNAPII resulting from DNA lesions or secondary structures. To restart transcription, accessory proteins including Cockayne syndrome proteins CSA and CSB protein are recruited to the damaged site (137). This is followed by recognition of the damaged site by xeroderma pigmentosum complementation group F (XPF) and 5'-incision of the lesion by excision cross complementing repair 1 (ERCC1) complex and 3'-incision of the lesion by XPG endonuclease (137). The incision of the lesion by XPF-ERCC1 induces strand breaks in the repeated region allowing formation of dissociation of the downstream strand from its template strand that leads to the formation of a TNR flap. Subsequently, the flap can fold back and form a stable hairpin that inhibits XPG removal of the flap generated in the repair process. This further results in repeat expansion (Figure I.14) (12).



Nature Reviews | **Genetics**

**Figure I.14 Nucleotide excision repair and trinucleotide repeat hairpin or loop formation (12)**

Theoretically, TNR expansion can be mediated by either GGR or TCR with an equal probability. However, it has been shown that XPC deficiency in HD transgenic mice has little effect on promoting CAG expansion (138). It could be mainly because GGR prefers to remove pre-formed looped-out structures and bulky lesions (12). In

contrast, TCR plays a very active role in modulating TNR deletion by removing the TNR secondary structure through ERCC1/XPF and XPG nuclease activity (131-133,139). The fact is consistent with the results showing that TNR deletion was suppressed in CSB deficient human cells (133) further indicating that NER may actively be involved in modulating TNR instability by processing TNR hairpins and loops formed during resolution of a stalled RNAP II caused by TNR hairpins or loops.

Besides MMR and NER, other DNA repair pathways are also involved in TNR instability. For example, DDR, along with DNA recombination initiated during DNA replication and repair can modulate TNR expansion in yeast, bacteria and human cells (71,140-146). In the double-strand DNA break repair pathway induced by TNR-induced replication blockage and fork stalling, large TNR expansion products are usually generated (86,97,147). However, TNR instability resulting from these repair pathways is dependent on DNA replication and occurs in dividing cells. It appears that an age-dependent small TNR expansion in non-dividing neuronal cells is mediated by a different repair pathway. Recent studies indicate that BER is the main pathway for modulating an age-dependent TNR expansion that is induced by oxidative stress (12,103).

#### **E. Base excision repair (BER) and TNR instability**

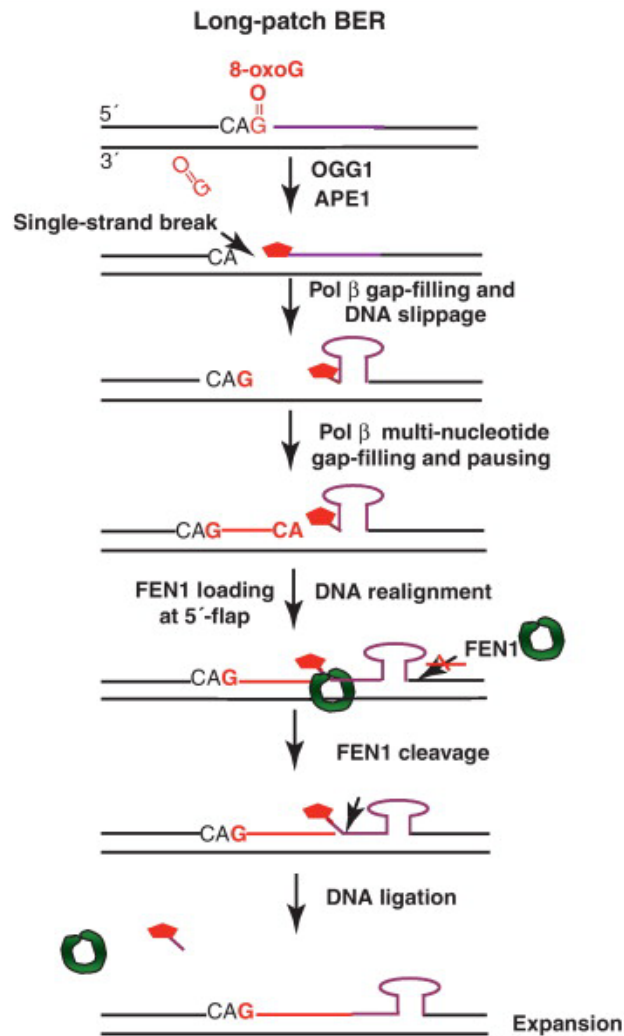
The genome of human is constantly exposed to environmental oxidative stresses and endogenous ROS generated from metabolic oxidative process that results in a series of oxidative base lesions including 8-oxoG, AP sites and ssDNA breaks (148). Because of its G rich sequence and the low redox potential of G, TNR is a hot spot for oxidative DNA base lesions generated by ROS resulting from both environmental and endogenous

sources (149). Base excision repair, as the major repair pathway for removing oxidized base lesions in human genome, may modulate TNR instability by creating a single-strand DNA break during repair of an oxidized base in a TNR tract. During base lesion repair, removal of oxidative DNA damage by a DNA glycosylase can result in an AP site that is subsequently 5'-incised by APE1, leaving a series of ssDNA break intermediates (20,150,151). Single-stranded DNA breaks formed in the context of TNRs during BER allow the formation of non-B form DNA structures in the same manner as during DNA replication and other repair pathways (12). Consistent with the notion, previous studies from bacteria, mice and human cells have demonstrated an association among TNR instability, oxidative DNA damage and BER (26,32,103,152-154). In a HD transgenic mouse model, an age-dependent progressive CAG repeat expansion in somatic cells has been found to be associated with an increased level of 8-oxoG. (103). Treatment of fragile X syndrome transgenic mice with an environmental oxidative agent, potassium bromate, resulted in CGG expansion (154). Treatment of fibroblasts from HD patients with hydrogen peroxide also causes CAG repeat expansion (103). Interestingly, the CAG repeat expansion in the HD mice is dependent on the presence of OGG1 because knock-out OGG1 in mice can prevent the progressive CAG repeat expansion and stabilize repeat length. All the results indicate that TNR expansion is directly associated with removal of oxidized bases by OGG1, which initiates BER of an oxidative DNA damage. In fact, in addition to OGG1, other BER enzymes and cofactors that are involved in LP-BER such as pol  $\beta$ , FEN1 and HMGB1 can also modulate somatic TNR instability *in vitro* (115) and in different types of nerve cells in HD mouse model that exhibit different BER capacity (155).

Furthermore, it has been found that decreased efficiency and capacity of BER during aging results in accumulation of oxidative DNA damage and ssDNA breaks in the genome that in turn may induce age-dependent somatic TNR expansion (156,157). For example, it has been found that a small repeat duplication can be induced by exposing a PARP1 knockout mouse model to alkylating DNA damage reagents (158). Since PARP1 is responsible for recognizing an intermediate with a single-strand break and promotes ssDNA break repair, deficiency of PARP1 may result in the accumulation of ssDNA breaks in the genome, thereby facilitating the formation of TNR hairpins and loops. In this manner, PARP1 deficiency could promote TNR expansion by stimulating the accumulation of ssDNA breaks in cells. As a central BER enzyme, pol  $\beta$  plays a crucial role in replacing the damaged base with the correct one with its dRP lyase activity and polymerase activity. In addition, pol  $\beta$  can also perform lesion bypass synthesis to bypass a template AP site with a low efficiency, resulting in DNA slippage and TNR expansion (159). Moreover, pol  $\beta$  multi-nucleotide gap-filling synthesis on a slipped TNR DNA strand also promotes TNR expansion by synthesizing extra repeat units (115,160). Although FEN1, another critical enzyme in BER, plays an essential role in preventing TNR expansion in bacteria, yeast and mouse by efficiently removing a TNR containing flap during DNA replication (71,161-166); however, a pre-formed stable hairpin created by DNA slippage can inhibit FEN1 cleavage of the entire TNR hairpin and result in TNR expansion (111,167). In this scenario, FEN1 is forced to make a cleavage on a short downstream 5'-TNR flap attached to a hairpin that is created by repeat rearrangement. The cleavage of FEN1 at an alternate cleavage site is named the FEN1 alternate flap cleavage (Figure I.15) (9,115). As an enzyme that seals a nick, LIG I is responsible for



sealing a ligatable nick that integrates a TNR hairpin into the genome. Thus, during BER, LIG I can compete with FEN1 to ligate a downstream hairpin and promote TNR expansion (168,169). All of these findings indicate that BER is actively involved in



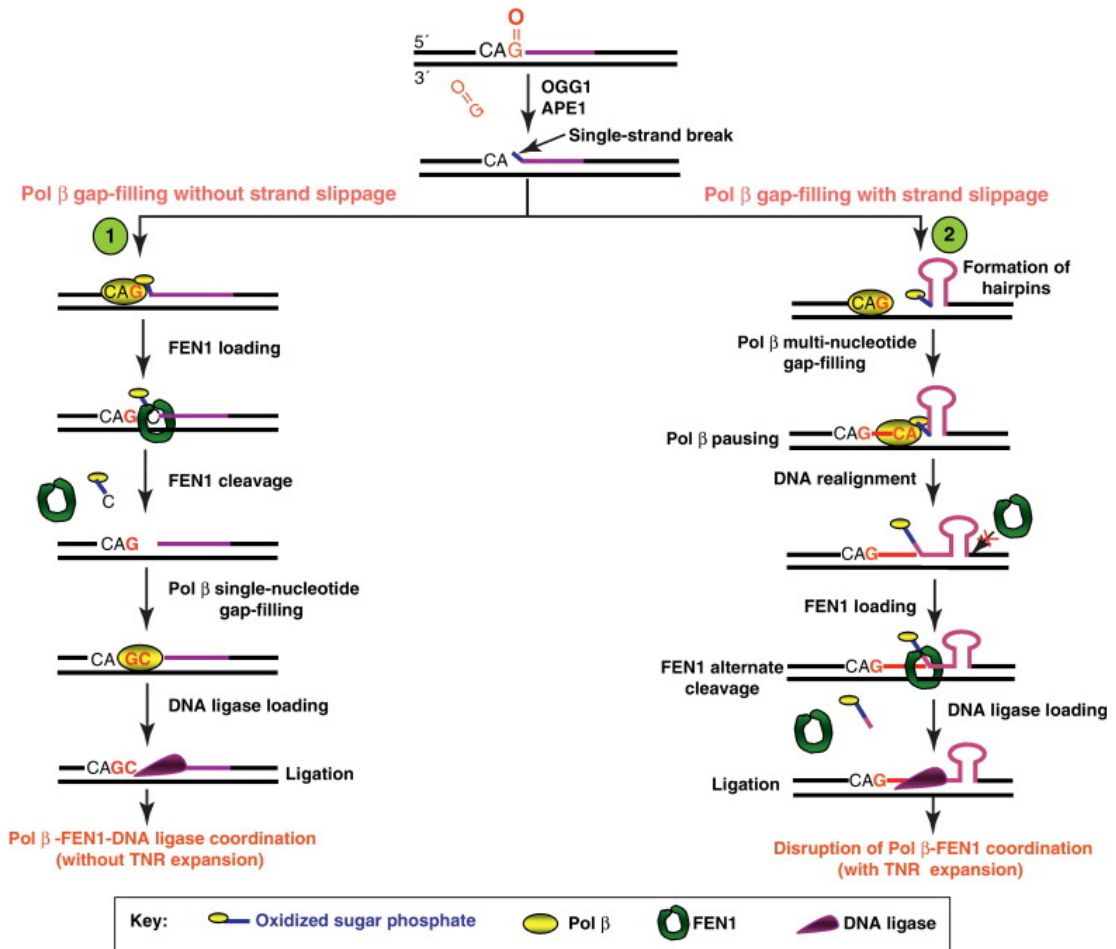
**Figure I.15 FEN1 alternate flap cleavage mediated CAG repeat expansion (11)**

modulating TNR expansion in non-dividing somatic cells.

It is proposed that TNR instability is initiated by the production of ssDNA breaks by APE1 cleavage on an AP site after removal of an oxidized base in a TNR region by OGG1, and the repeat expansion may be mediated by production of extra repeat units via

pol  $\beta$  multi-nucleotide gap-filling synthesis. To test this hypothesis, several *in vitro* studies with reconstituted BER of an oxidized DNA base lesion in the context of TNR tracts were performed with purified BER enzymes (115,160). These studies have shown that TNR expansion can only be induced by LP-BER through pol  $\beta$  multi-nucleotide gap-filling synthesis that produces extra repeat units. In this sub-pathway, FEN1 employs its alternate flap cleavage to process the 5'-end of a downstream TNR hairpin creating a ligatable nick allowing the sealing of the hairpin (Figure I.16) (9,115).

Thus, a model is proposed regarding how TNR expansion may occur as shown in



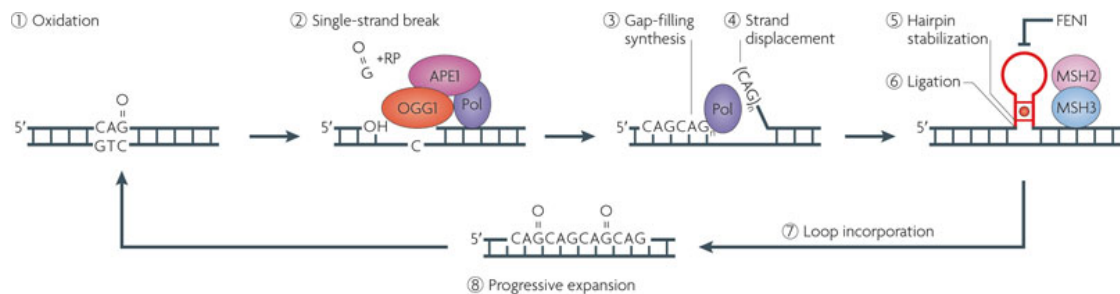
**Figure I.16 Hypothetical models illustrating CAG repeat stability modulated by coordination among BER enzymes during BER of 8-oxoG (11)**

Figure I.16. When an 8-oxoG occurs in the context of TNR such as CAG repeat, OGG1 removes the 8-oxoG leaving an AP site initiating BER (170). Subsequently, APE1 incises the 5'-end of the AP site leaving a 1-nt gapped DNA with a 5'-sugar phosphate in the context of CAG repeat tract. In the absence of DNA strand slippage, pol  $\beta$  fills in the gap, and FEN1 cleaves 1nt that is attached to the sugar phosphate residue. The strand break intermediate is efficiently processed by the coordination between pol  $\beta$  and FEN1 through the 'hit-and run' mechanism. In this scenario, no extra repeats are synthesized, and CAG repeat expansion is prevented. If DNA strand slippage occurs in the context CAG repeats, this will results in the formation a CAG repeat hairpin. A stabilized hairpin will disrupt the coordination between pol  $\beta$  and FEN1. In this scenario, pol  $\beta$  performs multi-nucleotide gap-filling synthesis to produce extract repeat units. Flap endonuclease 1 will be forced to perform its alternate flap cleavage to remove a part of the TNR hairpin facilitating CAG repeat expansion (9).

On some occasions, an oxidized DNA base lesion is subject to SN-BER. Repair of the 1-nt gap can be processed efficiently through the coordination between pol  $\beta$  and XRCC1-LIG III complex. DNA slippage can be prevented by efficient removal of sugar phosphate residue via pol  $\beta$  dRP lyase activity and ligation reaction mediated by XRCC1-LIG III complex. Thus, BER plays a dual role in modulating TNR instability, i.e., the SP-BER prevents TNR expansion, whereas LP-BER promotes TNR expansion.

The BER cofactor and proteins from other repair pathways that can modulate the stability of TNR hairpins during BER can facilitate or prevent TNR expansion. For example, the MSH2/MSH3 complex can stabilize hairpin structures (117,118). It could then promote TNR expansion through multiple cycles of BER of oxidative DNA damage

in the context of TNR (Figure I.17). On the other hand, in human cells, although each round of BER in TNR tract only results in 1-5 repeat expansion through pol  $\beta$  multi-nucleotide gap-filling synthesis via LP-BER (171,172), approximately 50,000 DNA base lesions are generated per day (173). Removal of these base lesions could result in numerous ssDNA breaks that may need to be repaired by many rounds of oxidation-repair process (i.e., 50,000 rounds of BER). Thus it could strongly promote TNR expansion (103,174).



Nature Reviews | Genetics

**Figure I.17** Loops formed during base excision repair by strand displacement (12)

Except for TNR expansion, previous studies also showed that DNA repair including BER and NER also mediate TNR deletions. In human lymphoblasts (32), bacterial cells (26), and mouse sperm (175), TNR deletion can be induced by DNA damaging agents including chemotherapeutic agents such as ethyl methanesulfonate, mitomycin C, N-ethylnitrosourea and cisplatin, or environmental damage agents such as ultraviolet radiation and ionizing radiation (32,175,176). Trinucleotide repeat deletion can also be induced by H<sub>2</sub>O<sub>2</sub> in bacteria and mouse kidney cells (26,177). Because these DNA damaging agents can induce alkylated or oxidized base lesions, DNA strand breaks

and DNA cross-links usually subjected to BER and NER, the results suggest that BER and NER are also actively involved in mediating TNR deletion. However, the molecular mechanism underlying TNR instability during BER remains to be unknown. Understanding the molecular basis of TNR expansion and deletion induced by DNA base lesions and repair will help to identify new targets for prevention and treatment of TNR expansion-related neurodegenerative diseases for which there is no treatment thus far. In my Ph.D. dissertation research, I conducted a series of studies on the molecular mechanisms of TNR deletion via oxidative DNA damage and BER. In CHAPTER 1, we studied the mechanism underlying TNR deletion via BER of an oxidative DNA damage. In CHAPTER 2, we identified the role of BER in repairing an oxidized base lesion located in the loop region of a TNR hairpin that ultimately results in attenuation or prevention of TNR expansion. Finally, in CHAPTER 3 we investigated the role of an oxidized DNA lesion, cyclo-deoxynucleotides, in inducing TNR deletion.

## OVERVIEW

Trinucleotide repeat expansion has been demonstrated as the cause of more than 40 types of human neurodegenerative diseases, whereas TNR deletion in the androgen receptor has been shown to be linked with human prostate cancer and ovarian cancer (12,72,74,82,178,179). A study of molecular mechanisms underlying somatic TNR instability during DNA repair is important for understanding the roles of DNA damage and repair on TNR instability as well as for determining whether DNA damage and repair can be developed as targets for prevention and treatment of TNR expansion-related neurodegenerative diseases. Although many studies focus on the mechanisms of TNR expansion during DNA replication and repair, the mechanisms underlying TNR deletion remain unknown and need to be elucidated. A study of somatic TNR deletion could lead to identification of novel targets for shortening the expanded TNR repeats facilitating development of a new strategy for treatment of TNR-related neurodegenerative diseases. Previous studies have shown that DNA damage can result in TNR deletion in both human cells and *E. coli*. The results lead to a hypothetical model for TNR deletion that needs to be examined in the context of DNA base lesions and BER. My dissertation project explored the mechanisms underlying TNR deletion and prevention of TNR expansion via BER and oxidative base lesion in three different ways. In CHAPTER 1, examining BER of oxidative DNA damage in the context of CAG or CTG repeats allowed us to establish a model for TNR deletion mediated by the coordination between a unique pol  $\beta$  hairpin bypass synthesis and FEN1 alternated flap cleavage. In CHAPTER 2, we examined the role of BER in preventing TNR expansion during repair of oxidized DNA base lesions in

the hairpin loop. We demonstrated that in coordinating with a 3'-5' endonuclease, Mus81/Eme1, BER of a base lesion in the loop region of a hairpin can lead to removal of the entirety or a part of the hairpin, demonstrating that BER in a TNR loop can be coupled with removal of the hairpin, thereby preventing or attenuating TNR expansion. In CHAPTER 3, we further explored a mechanism for TNR deletion via pol  $\beta$  bypass of an oxidized base lesion, 5',8-dcA, in a CAG repeat tract. We found that a 5',8-dcA on the template strand of a CAG repeat tract can induce the formation of a small CAG repeat loop that promotes pol  $\beta$  skipping over the loop structure, thereby resulting in repeat deletion.

**CHAPTER 1: TRINUCLEOTIDE REPEAT DELETION VIA A UNIQUE  
HAIRPIN BYPASS BY DNA POLYMERASE  $\beta$  AND ALTERNATE FLAP  
CLEAVAGE BY FLAP ENDONUCLEASE 1**

**ABSTRACT**

Trinucleotide repeat (TNR) expansions and deletions are associated with human neurodegenerative diseases and prostate cancer. Recent studies have pointed to a linkage between oxidative DNA damage, base excision repair (BER) and TNR expansion, which is demonstrated by the observation that DNA polymerase  $\beta$  (pol  $\beta$ ) gap-filling synthesis acts in concert with alternate flap cleavage by flap endonuclease 1 (FEN1) to mediate CAG repeat expansions. In this study, we provide the first evidence that the repair of a DNA base lesion can also contribute to CAG repeat deletions that were initiated by the formation of hairpins on both the template and the damaged strand of a continuous run of (CAG)<sub>20</sub> or (CAG)<sub>25</sub> repeats. Most important, we found that pol  $\beta$  not only bypassed one part of the large template hairpin but also managed to pass through almost the entire length of small hairpin. The unique hairpin bypass of pol  $\beta$  resulted in large and small deletions in coordination with FEN1 alternate flap cleavage. Our results provide new insight into the role of BER in modulating genome stability that is associated with human diseases.

**INTRODUCTION**

Trinucleotide repeats (TNRs) are highly polymorphic in the human genome and confer their instability via expansions or deletions (71,72). Trinucleotide repeat



expansions are known to be the cause of more than 40 neurodegenerative diseases, including Huntington's disease, myotonic dystrophy and fragile X syndrome among others (12,72,74). Similarly, TNR deletions are also closely related to human disorders. Epidemiological studies have demonstrated that CAG repeat deletions in the androgen receptor are correlated with human ovarian (82) and prostate cancer (178,179).

Trinucleotide repeat expansion has been studied extensively in bacteria (91), yeast (71), mouse (180), human (85,181) and human cell extracts (182). It was initially shown that TNR expansions occur during DNA replication of dividing somatic cells and during recombination of germ cells (12,85,97,183). However, later studies have demonstrated that expansions can also be mediated by DNA repair in dividing and non-dividing somatic cells (12,103,152). It is proposed that the formation of non-B form DNA secondary structures during DNA replication, recombination and repair contributes to expansions (85,140). In the human genome, typical non-B form structures arising from  $(CAG)_n/(CTG)_n$ ,  $(CTG)_n/(CAG)_n$ ,  $(CGG)_n/(CCG)_n$  and  $(GAA)_n/(TTC)_n$  sequence contexts include hairpins, tetraplexes and triplexes (85,86). These DNA secondary structures can cause DNA polymerase pausing and DNA slippage (108,109), replication fork stalling and collapse (86,97,98), inhibition of flap cleavage by flap endonuclease 1 (FEN1) (110-112), trapping of mismatch repair proteins (113,114) and disruption of coordination between DNA repair proteins (115). All of these events can compromise cellular repair mechanisms for removing the non-B form structures, thereby allowing them to be integrated into the human genome for TNR expansions. Hence, TNR expansion appears to be a consequence of the interactions between DNA secondary structures and DNA replication, repair and recombination. Compared with the other DNA metabolic pathways,

DNA repair plays a more versatile role in modulating TNR stability because it can repair the non-B form DNA structures generated from both DNA damages and other DNA metabolic pathways.

Several DNA repair pathways have been shown to modulate TNR expansion during both DNA replication and damage repair. The repair pathways that are initiated during DNA replication include mismatch repair (117,118,184-186) and double-stranded DNA break repair, along with DNA recombination (71,140-146). A mismatch repair protein complex, MSH2/MSH3 can bind and stabilize hairpin structures, thereby promoting TNR expansion (113,114). Double-stranded break repair processes the intermediates resulting from TNR-induced replication blockage and fork stalling (86,97,147), therefore facilitates large expansions. A specific DNA repair pathway can be initiated directly by DNA damage in the context of TNR to modulate repeat stability (103,115). For example, 8-oxoguanine (8-oxoG) in TNRs initiates DNA base excision repair (BER) through 8-oxoG DNA glycosylase (OGG1), which plays an essential role in mediating age-dependent somatic CAG or CGG repeat expansions in a Huntington's disease and fragile X syndrome mouse models (103,187,188). The DNA glycosylase removes the oxidized base lesion leaving an abasic site (115), which is subsequently incised by apurinic/apyrimidinic endonuclease 1 (APE1). Apurinic/apyrimidinic endonuclease 1 incision results in single-stranded DNA (ssDNA) breakage, which leads to DNA slippage and the formation of hairpins and a multi-nucleotide gap (115). The formation of hairpin further causes the disruption of functional coordination between DNA polymerase  $\beta$  (pol  $\beta$ ) and FEN1 leading to an inefficient pol  $\beta$  multi-nucleotide gap-filling synthesis and FEN1 alternate cleavage, i.e., the cleavage of a short 5'-flap

attached to the 5'-end of a hairpin subsequently resulting in production of a ligatable nick for ligating a hairpin with newly synthesized repeats leading to repeat expansion (115). Thus, inefficient BER leads to TNR expansion (11). A multi-nucleotide gap generated during BER could also allow the formation of a hairpin on the template strand that may mediate TNR deletion induced by DNA damages.

Trinucleotide repeat deletions can be induced in human lymphoblasts (32), bacterial cells (26) and mouse sperm (175), by DNA-damaging agents including ethyl methanesulfonate, mitomycin C, ethylnitrosourea, ultraviolet radiation, cisplatin and ionizing radiation (32,175,176). Because the DNA-damaging agents can induce alkylating DNA damage, DNA strand breaks, DNA cross-links that are usually subjected to BER and nucleotide excision repair, it suggests that BER and NER are actively involved in mediating TNR deletion. The transcription-coupled NER was found to be associated with TNR deletion in mammalian cells (132,133,139) and in *Drosophila* (131). Thereby supporting the notion, NER proteins, Xeroderma pigmentosum group A-complementing protein (139), Xeroderma pigmentosum group G-complementing protein and Cockayne syndrome B protein are also known to promote TNR deletion (131-133). These findings suggest an important role of DNA damage and repair in modulating TNR deletion.

Although significant progress has been made in exploring TNR deletion in recent years, our knowledge about the process remains limited, therefore results in poor understanding of the mechanisms by which deletions arise. Understanding the molecular basis of TNR deletions and how chemicals may induce deletions through DNA damage and repair is critically important for the development of new methods for prevention and

treatment of human diseases caused by both TNR deletions and expansions (189). Interestingly, in addition to its role in promoting CAG repeat expansion (103,115), oxidative DNA damage was also shown to induce repeat deletions in bacteria (26), suggesting that oxidized base lesions and BER play a dual role in causing both expansions and deletions. The concept has been further supported by *in vitro* evidence that pol  $\beta$  slippage is involved in CAG repeat deletion (190). However, it remains unknown if or how DNA base lesions and BER cause deletions. To explore the potential role of BER in TNR deletions, we examined CAG repeat deletions during BER and identified the roles of major BER enzymes in the steps that could lead to TNR deletions. Surprisingly, we discovered that repair of a DNA base lesion in the context of (CAG)<sub>20</sub> and (CAG)<sub>25</sub> repeats led to repeat deletions with the concomitant loss of 1-21 triplet repeats per repair event, depending upon the length of repeats. The deletions were mediated by pol  $\beta$  hairpin bypass in coordination with FEN1 alternate flap cleavage. As a result, we propose a hypothetical model to describe a unique mechanism underlying TNR deletions involving a base lesion-induced loss of TNR through pol  $\beta$  hairpin bypass and FEN1 alternate flap cleavage. Our studies provide the first evidence that TNR deletion can be initiated by DNA base lesions and mediated by BER.

## **MATERIALS AND METHODS**

### **Materials**

Dexoyribose nucleic acid oligonucleotides were purchased from Integrated DNA Technologies (IDT Inc., Coralville, IA, USA). The radionucleotides [ $\gamma$ -<sup>32</sup>P] ATP (6000

mCi/mmol) and Cordycepin 5'-triphosphate 3'-[ $\alpha$ - $^{32}$ P] (5000 mCi/mmol) were from PerkinElmer Inc. (Boston, MA, USA). Micro Bio-Spin 6 chromatography columns were purchased from Bio-Rad (Hercules, CA, USA). Deoxynucleoside 5'-triphosphates (dNTPs) and terminal deoxynucleotidyl transferase were from Fermentas (Glen Burnie, MD, USA). The T4 polynucleotide kinase was from USB Corp. (Cleveland, OH, USA). Mung Bean Nuclease was from Epicenter (Madison, WI, USA). The original TA cloning kit was from Invitrogen/Life Technologies Corp. (Grand Island, NY, USA). The Big Dye Terminator version 3.1 Cycle Sequencing Kit was from Applied Biosystems/Life Technologies Corp (Grand Island, NY, USA). Polyclonal and monoclonal antibodies of FEN1 and tubulin and non-specific IgG were from Abcam (Boston, MA, USA). All standard chemical reagents were purchased from Sigma-Aldrich (St. Louis, MO, USA) and Thermo Fisher Scientific (Pittsburgh, PA, USA). Purified *E. coli* uracil-DNA glycosylase (UDG), APE1, pol  $\beta$ , FEN1, DNA ligase I (LIG I) and a monoclonal pol  $\beta$  antibody were generous gifts from Dr. Samuel Wilson at the National Institute of Environmental Health Sciences/National Institutes of Health, Research Triangle Park, NC, USA.

### **Oligonucleotide substrates**

Substrates containing (CAG)<sub>20</sub> or (CTG)<sub>20</sub> with a deoxyuridine or a tetrahydrofuran (THF), an abasic site analog, were designed to mimic a repair intermediate with a native or modified abasic site in the context of TNR. The THF residue was used to represent an oxidized abasic site in the study. The guanine of the first CAG or CTG of a substrate was substituted with a THF residue or a deoxyuridine residue.

A substrate with a native abasic site was generated by incubating an oligonucleotide containing a deoxyuridine residue, with *E. coli* UDG (2.5 nM) at 37 °C for 15 min. Substrates were constructed by annealing the damaged strands to their template strands at a molar ratio of 1:1.5. Substrates with a perfect 15-nt base-paired hairpin or a preformed (CTG)<sub>7</sub> or a (CAG)<sub>7</sub> or a (CTG)<sub>25</sub> hairpin in their template strands were constructed to mimic the repair intermediates containing a template hairpin with or without a 5'-downstream flap. The downstream primers contained a 5'-phosphate-THF alone or a 5'-phosphorylated THF with a 21-nt-dT flap or a (CAG)<sub>7</sub> or a (CTG)<sub>7</sub> or a (CAG)<sub>12</sub> or a (CAG)<sub>25</sub> flap/hairpin. Upstream and downstream primers were annealed to their templates at a molar ratio of 1:1:1.5. Substrates were radiolabeled at the 5'-end of the damaged strand or that of the upstream primer or at the 3'- or 5'-end of the template strand for measuring various types of enzymatic activities. The sequences of oligonucleotide substrates are listed in Table 1.1. DNA size markers that correspond to the cleavage products of Mung Bean Nuclease and the products of pol β nucleotide insertions were synthesized and purified by polyacrylamide gel electrophoresis (191). A DNA fragment that contained (CAG)<sub>20</sub> without any DNA base lesions was used as a marker for DNA fragment analysis.



U1	30	CTGCAGCTGATGCGCAATATGAGTAAGTCT
U2	20	CGAGTCATCTAGCATCCGTA
U3	20	GCAATGAGTAAGTCTACGTA

---

F: tetrahydrofuran, THF U: deoxyuridine



### ***In vitro* BER in mouse embryonic fibroblast cell extracts**

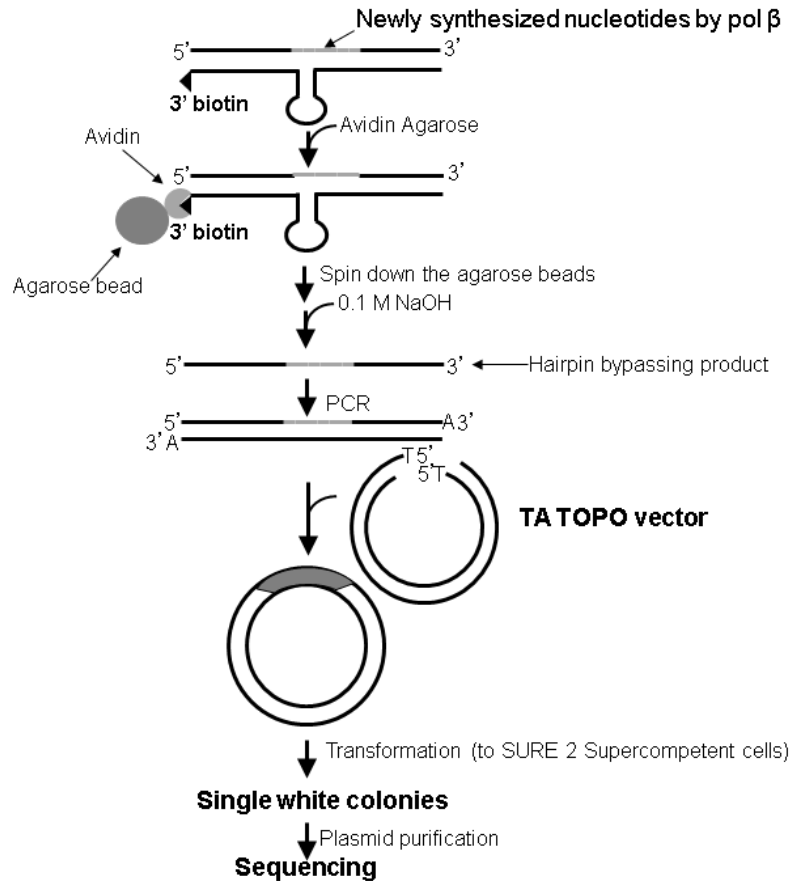
Cell extracts were made as described (63). Briefly, pol  $\beta$  null (pol  $\beta^{-/-}$ ) and wild-type pol  $\beta$  (pol  $\beta^{+/+}$ ) mouse embryonic fibroblast (MEF) cell lines were grown to near confluence. Cells were harvested and lysed and cell lysates were centrifuged at 12,000 rpm for 30 min to obtain whole-cell extracts that were subsequently dialyzed into BER reaction buffer containing 50 mM Tris-HCl, pH 7.5, 50 mM KCl, 0.1 mM EDTA, 0.1 mg/ml bovine serum albumin and 0.01% Nonidet P-40 (NP-40). The FEN1 deficient cell extracts were prepared by incubating wild-type MEF cell extracts (100  $\mu$ g) with 5  $\mu$ g FEN1 polyclonal antibodies at 4 °C overnight. Protein-antibody complexes were precipitated with protein A agarose beads (Pierce-Thermo Scientific, Rockford, IL, USA) for FEN1 depletion. Control cell extracts were prepared under the same conditions by incubating cell extracts with 5  $\mu$ g IgG. Levels of FEN1, pol  $\beta$  and tubulin (loading control) were determined by immunoblotting. Base excision repair was performed by incubating 60  $\mu$ g cell extracts with the (CAG)<sub>20</sub>-THF or a (CAG)<sub>20</sub>-uracil-containing substrate (25 nM). A reaction (25  $\mu$ l) was assembled in BER reaction buffer with 5 mM MgCl<sub>2</sub>, 2 mM ATP and 50  $\mu$ M dNTPs. Base excision repair reactions were also performed by incubating substrates with the pol  $\beta^{-/-}$  cell extracts supplied with increasing concentrations of purified pol  $\beta$ . Substrates were pre-incubated with 50 nM APE1 at 37 °C for 30 min to generate ssDNA break intermediates. The reaction mixture was assembled on ice and incubated at 37 °C for 30 min. The reactions were terminated by transfer to 95 °C for 5 min and the reaction mixture was subsequently subjected to protease K digestion at 55 °C for 30 min. DNA was precipitated and dissolved into buffer containing 95% formamide and 2 mM EDTA. Substrates and repair products were

separated by 15% urea-denaturing PAGE and detected using a Pharos FX Plus PhosphorImager from Bio-Rad. Repair products were excised from gels and eluted at room temperature overnight with TE buffer (10 mM Tris-HCl, pH 7.5 and 1 mM EDTA) and then were precipitated and dissolved in TE buffer. Twenty nanograms of the repaired DNA was used for subsequent DNA fragment analysis. All substrates were <sup>32</sup>P-radiolabeled at the 5'-end of the strand that contained a deoxyuridine or a THF residue.

### **In vitro BER reconstituted with purified enzymes**

*In vitro* BER of a native or modified sugar in the context of (CAG)<sub>20</sub> repeats was performed with purified UDG, APE1, pol β, FEN1, LIG I and (CAG)<sub>20</sub>-containing substrates with a deoxyuridine or a THF residue. The 20-μl reaction was reconstituted with the indicated concentrations of substrates and BER enzymes in BER reaction buffer that contained 5 mM MgCl<sub>2</sub>, 2 mM ATP and 50 μM dNTPs. Reaction mixtures were assembled on ice and incubated at 37 °C for 15 min. Reactions were then terminated by transfer to 95 °C for 10 min in buffer containing 95% formamide and 10 mM EDTA. Substrates and products were separated by 15% urea-denaturing PAGE and detected by PhosphorImager. Repair products were isolated for size analysis. Substrates were <sup>32</sup>P-radiolabeled at the 5'-end of the strand with a deoxyuridine or a THF residue. Base excision repair mediated by pol β hairpin bypass was reconstituted by incubating purified pol β, FEN1 and LIG I with 25 nM substrates that contained a hairpin in the template strands with or without a downstream flap/hairpin. Reactions were performed in BER reaction buffer with 5 mM MgCl<sub>2</sub> and 50 μM dNTPs. To isolate repair products, the template or upstream strands of the substrates were biotinylated at the 3'- or 5'-end.

Repair products were incubated with Pierce avidin agarose beads (Pierce-Thermo Scientific, Rockford, IL, USA) in binding buffer that contained 0.1 M phosphate, 0.15 M NaCl, pH 7.2 and 1% Nonidet P-40 at 4 °C for 2 h with rotation. The agarose beads were centrifuged at 5,000 rpm for 1 min and were washed with the binding buffer three times. The repaired strands were dissociated and separated from their template strands with 0.15



**Figure 1.1 Scheme of DNA Sequencing for Repeat Deletion Products.**

The repair products of pol β hairpin bypass synthesis were incubated with Pierce avidin agarose beads allowing the biotinylated template strand binding to avidin agarose beads. Agarose beads were spun down, and the repaired strand was released by incubating the beads with 0.15 M NaOH. Repaired strands were then subjected to PCR amplification. Amplified repair products were then cloned into the TA TOPO vector. Plasmids that contained repair products were then transformed into bacteria for amplification. Plasmids were isolated and purified by Qiagen miniprep kit and subjected to DNA sequencing.

M NaOH and centrifugation at 5,000 rpm for 2 min. Repaired strands were then precipitated and dissolved in TE buffer for subsequent amplification by polymerase chain reaction (PCR), cloning and DNA sequencing (Figure 1.1).

### **Enzymatic activity assays**

Pol  $\beta$  DNA synthesis was measured using substrate that contained (CAG)<sub>20</sub> repeats with a THF alone or substrates that contained a template hairpin with or without a downstream 5'-THF-flap/hairpin. The DNA synthesis was measured at 37 °C in a 20- $\mu$ l reaction mixture that contained BER reaction buffer with 5 mM MgCl<sub>2</sub> and 50  $\mu$ M dNTPs. Flap endonuclease 1 cleavage activity was examined in the same buffer with 5 mM MgCl<sub>2</sub> and 50  $\mu$ M dNTPs in the absence or presence of increasing concentrations of pol  $\beta$  at 37 °C for 15 min. Substrates and products were separated by 15% urea-denaturing PAGE and detected by a PhosphorImager.

### **Probing of hairpin structures by Mung Bean Nuclease digestion**

The formation of hairpin structures by CAG and CTG repeats was probed by incubating 2 U Mung Bean Nuclease with 200 nM substrates that contained (CAG)<sub>20</sub> or (CTG)<sub>20</sub> repeats with a THF embedded in the first CAG or CTG or in a random DNA sequence. Hairpins present in the substrates containing a preformed template (CTG)<sub>7</sub> and (CTG)<sub>25</sub> hairpin were also probed by the enzyme digestion. Substrates were pre-incubated with 10 nM APE1 at 37 °C for 30 min to generate an ssDNA breakage intermediate. The 10- $\mu$ l reaction mixture was assembled by incubating 2 U of Mung Bean Nuclease with 200 nM substrates in reaction buffer containing 30 mM sodium

acetate (pH 4.6), 50 mM NaCl, 1 mM zinc acetate and 0.01% Triton X-100. The reaction was incubated at 37 °C for 1, 2, 3, 5 and 8 min and subsequently subjected to protease K digestion at 55 °C for 30 min to remove Mung Bean Nuclease. Substrates and products were separated by 15% urea-denaturing PAGE and detected by a PhosphorImager.

### **Sizing of repair products with CAG repeats by DNA fragment analysis**

Repair products (20 ng) resulting from BER in the context of (CAG)<sub>20</sub> repeats were amplified by PCR with a forward primer tagged by a 6-carboxyfluorescein and an untagged reverse primer that were annealed to the 5'- and 3'-end of random sequence regions of repair products. PCR amplification was performed under the following conditions: 95 °C for 10 min, 1 cycle; 95 °C for 30 s, 50 °C for 30 s and 72 °C for 1.5 min, 35 cycles; 72 °C for 1 h, 1 cycle. The 6-carboxyfluorescein-labeled PCR products were then subjected to capillary electrophoresis. The size of repair products was determined by DNA fragment analysis (Florida International University DNA Sequencing Core Facility) with Peak Scanner version 1.0 software (Applied Biosystems, Foster City, CA, USA).

### **Cloning and sequencing of repair products**

The repair products resulting from pol β hairpin bypass were amplified by PCR under the following conditions: 95 °C for 5 min, 1 cycle; 95 °C for 1 min, 54 °C for 30 s and 72 °C for 1 min, 25 cycles; 72 °C for 10 min, 1 cycle. PCR products were then ligated into the pCR 2.1-TOPO vector (Invitrogen/Life Technologies Corp., Grand Island, NY, USA). Recombinant plasmids that contained repair products were then transformed

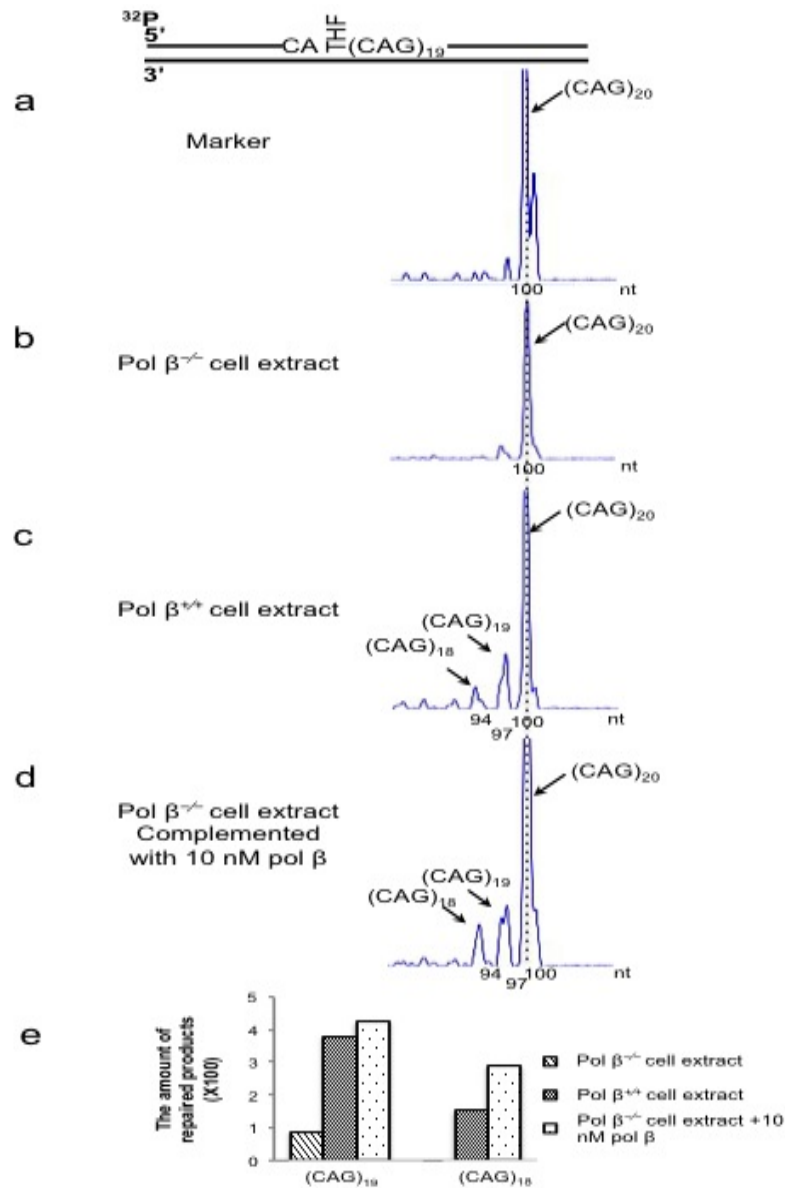
into SURE 2 Supercompetent Cells (Stratagene, Santa Clara, CA, USA). The plasmids were isolated and purified by the Miniprep Kit from Qiagen (Valencia, CA, USA) and were subjected to DNA sequencing. PCR sequencing reactions were performed with the Big Dye Terminator version 3.1 Cycle Sequencing Kit from Applied Biosystems/Life Technologies under the following conditions: 94 °C for 15 s, 50 °C for 10 s and 60 °C for 4 min, 35 cycles. The sequences of repair products were determined using ABI Genetic Analyzer 3130 version 1 (Applied Biosystems/Life Technologies Corp) by the Florida International University DNA Sequencing Core Facility and analyzed with Mac Vector version 12.0.2 software (MacVector Inc., Cary, NC, USA).

## **RESULTS**

### **Pol $\beta$ promotes CAG repeat deletion during BER**

Pol  $\beta$ , a central component of BER, removes a native deoxyribose phosphate residue and fills in a single-nucleotide gap to fulfill the efficient single nucleotide BER (11). Pol  $\beta$  can also perform multi-nucleotide gap-filling synthesis and strand displacement synthesis that leads to the relatively inefficient long-patch BER (11). Pol  $\beta$  multi-nucleotide gap-filling synthesis was found to promote CAG repeat expansion by producing extra repeat units (115), whereas its slippage synthesis is suggested to be involved in repeat deletion. Thus the enzyme plays a pivotal role in modulating TNR instability. To further determine how pol  $\beta$  could be involved in TNR deletion, we initially studied deletions in the cell extract-based BER of a THF embedded at the first CAG of a (CAG)<sub>20</sub> repeat-containing substrate. Experiments were performed using pol

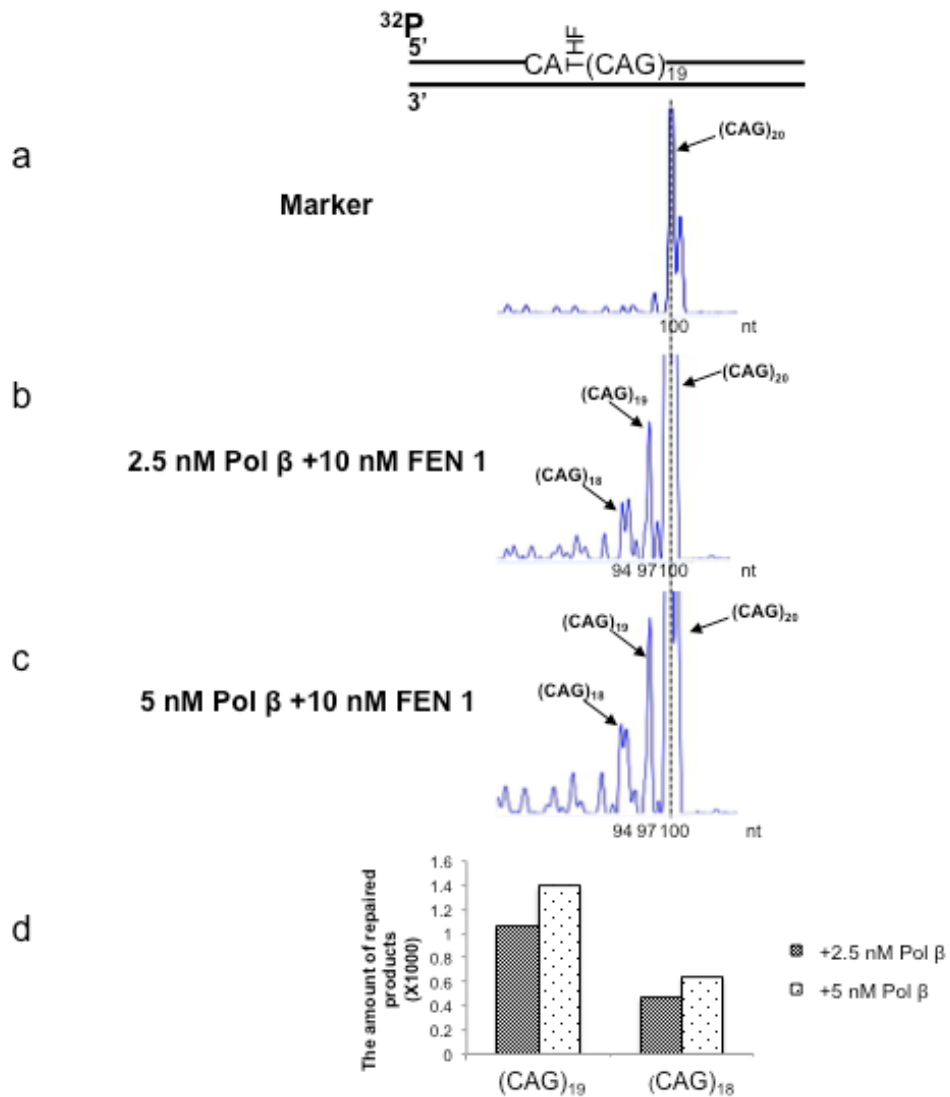
$\beta^{-/-}$  (Figure 1.2, panel b), pol  $\beta^{+/+}$  cell extracts (Figure 1.2, panel c) or pol  $\beta^{-/-}$  cell extracts complemented with 10 nM purified pol  $\beta$  (Figure 1.2, panel d). The results revealed that significant amounts of deletion products formed with pol  $\beta^{+/+}$  cell extracts and pol  $\beta^{-/-}$  cell extracts complemented with purified pol  $\beta$  (Figure 1.2, panels c and d). However, only a small amount of deletion product was produced by pol  $\beta^{-/-}$  cell extracts (Figure 1.2, panel b). The amounts of deletion products increased by 3- to 4-fold by the presence of pol  $\beta$  (Figure 1.2, panel e). The repair mainly resulted in (CAG)<sub>19-</sub> and (CAG)<sub>18-</sub>-repeat containing products with one or two repeat deletion (Figure 1.2, panels c and d). To further confirm the role of pol  $\beta$  in CAG repeat deletion, we determined the deletions in a reconstituted BER by incubating the (CAG)<sub>20</sub>-containing substrate with purified APE1, FEN1, LIG I and increasing concentrations of pol  $\beta$ . Consistent with the results from the pol  $\beta$  cell extracts, the reconstituted BER also resulted in deletion of one or two CAG repeats with the amounts of the products slightly increased with increasing concentrations of pol  $\beta$  (Figure 1.3, panels b, c and d). Thus, the results indicate that pol  $\beta$  promotes CAG repeat deletion. To exclude the possibility that deletion products could result from PCR artifacts due to Taq polymerase DNA slippage, a (CAG)<sub>20</sub> repeat-containing marker without any base lesions was PCR-amplified and subjected to sizing analysis under the same conditions as the ones used for repair products. The results showed that little deletion product was from the amplified marker (Figure 1.2 and 1.3, panel a), indicating that repeat deletions were from BER rather than from PCR artifacts.



**Figure 1.2 CAG repeat deletion during BER with MEF cell extracts.**

Repair of an abasic site in the context of CAG repeats was examined with a (CAG)<sub>20</sub>-containing substrate with a THF embedded in the first CAG. BER reactions were reconstituted with cell extracts from pol  $\beta^{+/+}$ , pol  $\beta^{-/-}$  MEFs and pol  $\beta^{-/-}$  MEF cell extracts complemented with 10 nM of purified wild-type pol  $\beta$ . Gel-isolated repair products were amplified by PCR with 6-carboxyfluorescein-labeled primers. PCR products were separated by capillary electrophoresis and their size was determined by DNA fragment analysis. The products are illustrated as different individual peaks. The height of a peak represents the abundance of repair products. The sizes of repair products are illustrated in nucleotides and in number of repeat units.





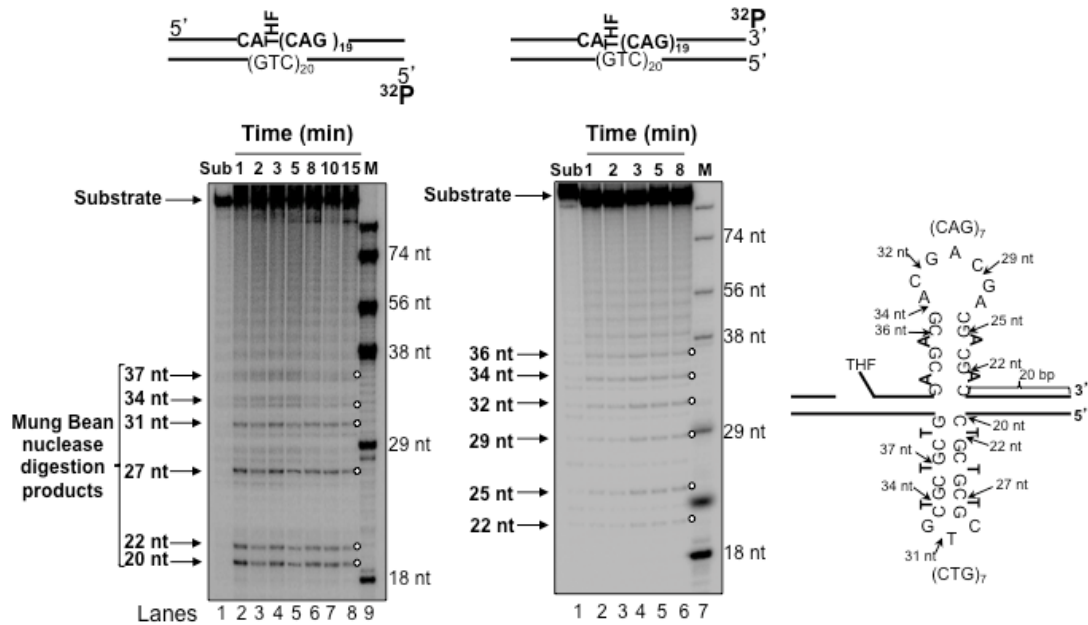
**Figure 1.3 CAG repeat deletion during BER with purified BER enzymes.**

BER reactions were reconstituted by incubating 25nM substrate radiolabeled at the 5'-end of its damaged strand with pol β (2.5 and 5 nM), 50 nM APE1, 10 nM FEN1 and 5 nM LIG I along with 50 mM dNTPs under the conditions described in the 'In vitro BER in mouse embryonic fibroblast cell extracts of 'Experimental Procedures'.

### **Formation of CAG and CTG repeat hairpin structures during BER.**

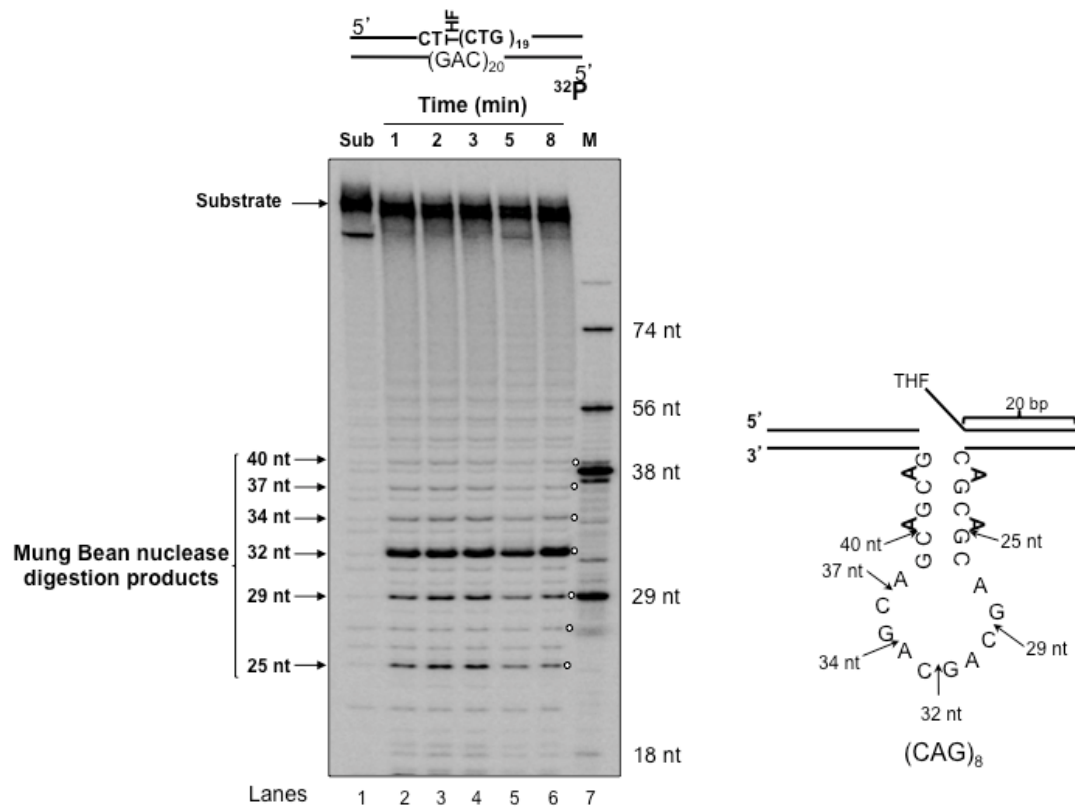
It is proposed that TNR deletion involves the formation of hairpin structures on the template DNA strand and hairpin bypass by DNA polymerases during DNA replication (4,85). Thus, it is conceivable that hairpins may also form on the template strands during BER leading to pol  $\beta$  hairpin bypass DNA synthesis. To test this possibility, we initially determined the formation of hairpin structures in the damaged and template strands of the substrates containing (CAG)<sub>20</sub> or (CTG)<sub>20</sub> with a THF residue, using Mung Bean Nuclease. The nuclease preferentially makes cleavages at a single-strand loop region and the sites with mismatched base pairs in the stem region of a hairpin. A specific cleavage pattern of Mung Bean Nuclease on CAG or CTG repeats indicates the formation of a hairpin with a specific size. The enzyme cleavage on the template strand of the (CAG)<sub>20</sub>-repeat substrate resulted in products with 20, 22, 27, 31, 34 and 37 nt, respectively (Figure 1.4, left panel). The cleavage pattern indicated that a (CTG)<sub>7</sub> hairpin formed next to the 5'-side 20 nt-random sequence of the template strand. The hairpin consisted of a one-CTG loop and a six-CTG repeat stem as indicated (Figure 1.4). The enzyme cleavage on the damaged strand mainly generated products with 25, 29, 32, 34 and 36 nt, respectively (Figure 1.4, right panel), indicating the formation of a (CAG)<sub>7</sub> repeat hairpin next to the 3'-side 20 nt-random sequence of the damaged strand. The hairpin was composed of a (CAG)<sub>3</sub>-repeat loop and a stem with (CAG)<sub>4</sub> repeats (Figure 1.4). The nuclease cleavage on the template strand resulted in products with 25, 29, 32, 34, 37 and 40 nt (Figure 1.5), respectively. The cleavage pattern demonstrated the formation of a (CAG)<sub>8</sub> repeat hairpin next to the 5'-side 20-nt random sequence of the template strand. The hairpin contained a (CAG)<sub>4</sub>-repeat loop region and a stem region

with (CAG)<sub>4</sub> repeats (Figure 1.5). The nuclease digestion of a substrate containing random sequence failed to generate any cleavage products (Figure 1.6). The results indicated that the formation of hairpins on both the damaged and template strands during BER is repeat sequence-dependent.



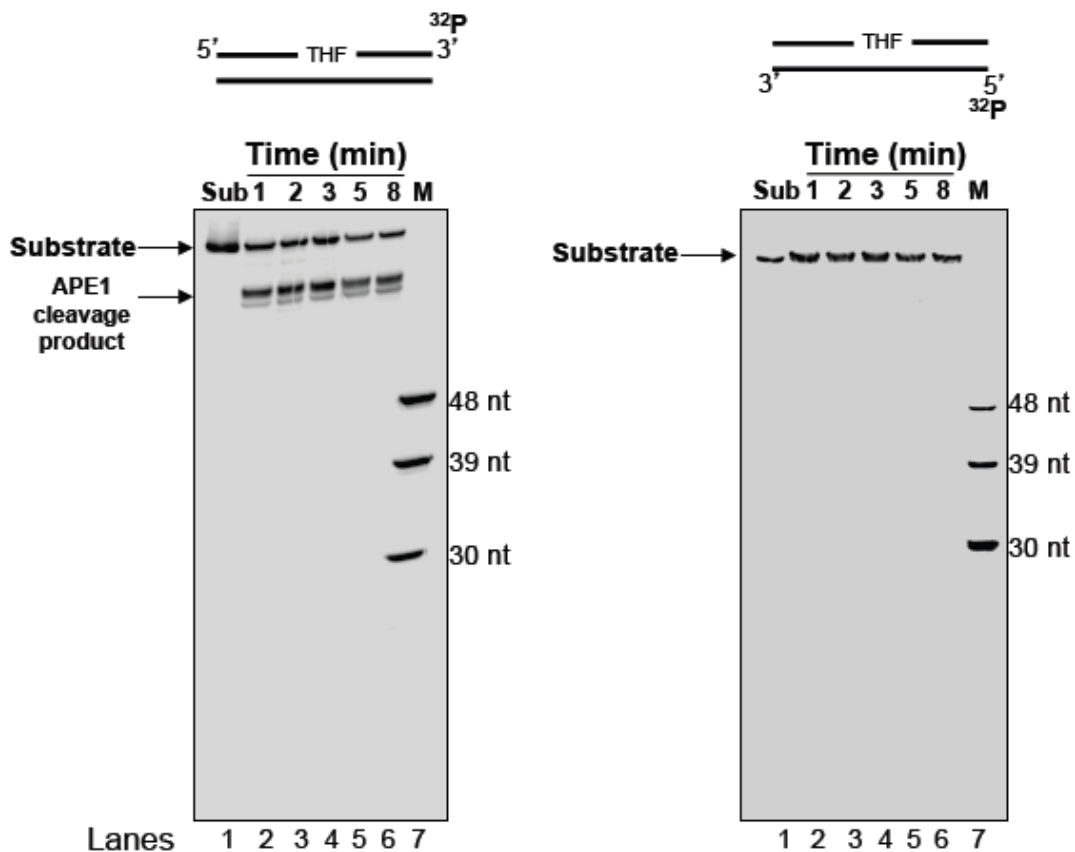
**Figure 1.4 Formation of trinucleotide repeat hairpins on the (CAG)<sub>20</sub> repeat substrate during BER.**

Formation of a template hairpin on the (CAG)<sub>20</sub> repeat substrate with a THF residue was probed by Mung Bean Nuclease digestion (left panel). The substrate was radiolabeled at the 5'-end of its template strand and was pre-cut by 10 nM APE1. Subsequently, the substrate was incubated with 2 U of Mung Bean Nuclease at 1-, 2-, 3-, 5-, 8-, 10- and 15-min time intervals (lanes 2–8). Lane 1 represents the undigested substrate. Lane 9 represents synthesized size markers (M) with 18, 29, 38, 56 and 74 nucleotides, respectively. Hairpins formed on the damaged strand were probed by the enzyme digestion of the substrate radiolabeled at the 30-end of the strand (right panel, lanes 2-6). Lane 1 represents the undigested substrate. Lane 7 represents size markers (M) with 18, 29, 38, 56 and 74 nucleotides, respectively. The substrates are illustrated schematically above the gel. A hairpin deduced by a specific nuclease cleavage pattern and the nuclease digestion sites is illustrated schematically beside the gel.



**Figure 1.5 Formation of trinucleotide repeat hairpins on the (CTG)<sub>20</sub> repeat substrate during BER.**

The template strand that contained (CTG)<sub>20</sub> repeats was labeled at its 5'-end. The substrate was incubated with 2 U of Mung Bean Nuclease at 1-, 2-, 3-, 5- and 8-min time intervals, respectively (lanes 2-6). Lane 1 represents the undigested substrate. Lane 7 represents markers (M). For all the experiments, 200 nM of substrate was used. Arrows indicate the major digestion products. The substrates are illustrated schematically above the gel. A hairpin deduced by a specific nuclease cleavage pattern and the nuclease digestion sites is illustrated schematically beside the gel.



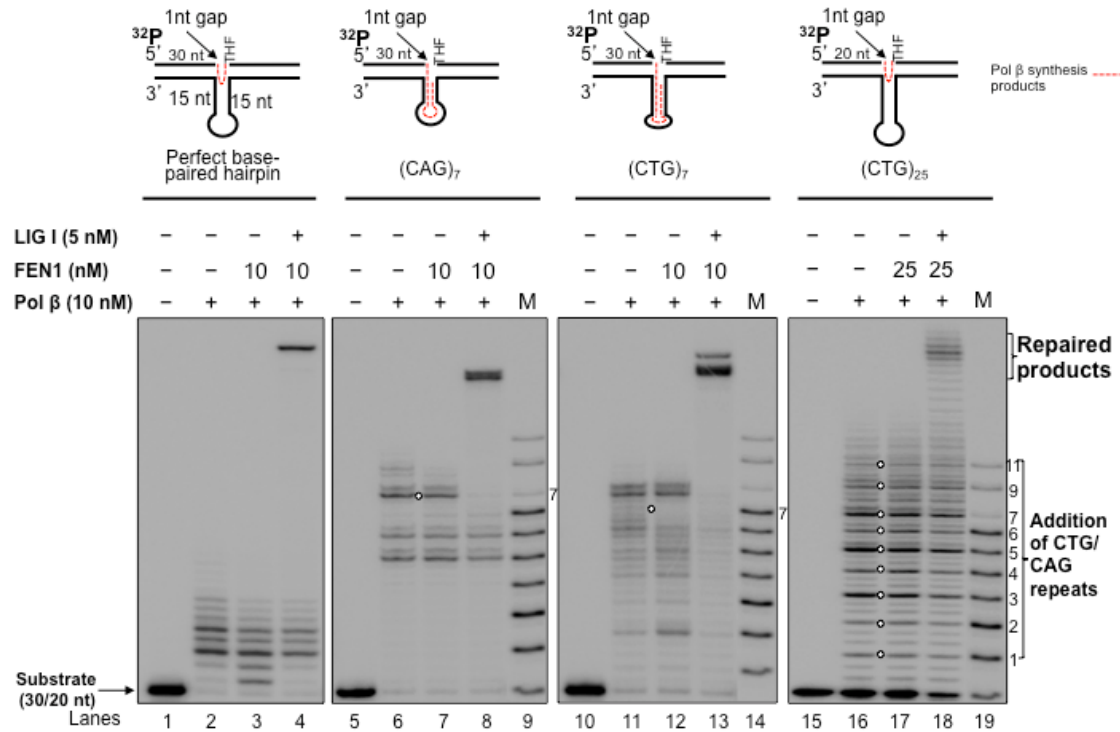
**Figure 1.6 Probing of hairpins formed by random sequences.**

The damaged strand of a substrate with random sequence was radiolabeled at the 3'-end and was pre-cut by 10 nM APE1 (left panel). Subsequently, the substrate was incubated with two units of Mung Bean Nuclease at 1-, 2-, 3-, 5-, and 8-minute time intervals (lanes 2-6). Lane 1 represents the undigested substrate. Lane 7 represents synthesized size markers (M) with 30, 39, and 48 nucleotides, respectively. The template strand was radiolabeled at its 5'-end and was digested by the enzyme (right panel, lanes 2-6). Lane 1 represents the undigested substrate. Lane 7 represents size markers (M) with 30, 39, and 48 nucleotides, respectively. Arrows and white dots indicate APE1 pre-cut product. The substrates are illustrated schematically above the gel.

## **Pol $\beta$ bypasses a template hairpin effectively, thereby facilitating CAG repeat deletions during BER**

Since TNR deletion is proposed to be mediated by DNA polymerase bypass of a template hairpin generated during DNA replication (4), a similar mechanism could also be adopted by pol  $\beta$  to mediate TNR deletion during BER. To test this hypothesis, pol  $\beta$  DNA synthesis was examined with substrates containing a template hairpin consisting of a 6-nt loop and a 15-nt perfect base-paired stem (Figure 1.7 and 1.8, lanes 2–4) or a template hairpin with (CAG)<sub>7</sub> (Figure 1.7 and 1.8, lanes 6–8) or (CTG)<sub>7</sub> (Figure 1.7 and 1.8, lanes 11–13) or (CTG)<sub>25</sub> (Figure 1.7 and 1.8, lanes 16–18). These substrates mimic the BER intermediates containing a small or large template hairpin with or without a 5'-downstream flap that result from APE1 5'-incision and DNA slippage. The formation of the hairpins on both the template and damaged strands of these substrates was verified by Mung Bean Nuclease digestion (Figure 1.9–1.11). Substrates were radiolabeled at the 5'-end of the upstream primers for characterizing any effects of downstream flaps/hairpins on pol  $\beta$  primer extension for its hairpin bypass. With all the substrates, pol  $\beta$  exhibited efficient multi-nucleotide synthesis (illustrated schematically by red lines) indicating that it readily bypassed the template hairpins with varying sizes (Figure 1.7 and 1.8, lanes 2, 6, 11 and 16). For both (CAG)<sub>7</sub> and (CTG)<sub>7</sub> template hairpins, pol  $\beta$  inserted five to seven repeat units (Figure 1.7, lanes 6, 7, 11 and 12), indicating that the enzyme passed through almost the entire length of the hairpins. In contrast, pol  $\beta$  only inserted 3–10 repeats in bypassing (CTG)<sub>25</sub> hairpin (Figure 1.7 and 1.8, lanes 16 and 17), suggesting that the enzyme can skip over the hairpin by passing through its portion. Interestingly, the presence of FEN1 did not affect pol  $\beta$  DNA synthesis on the substrates without a

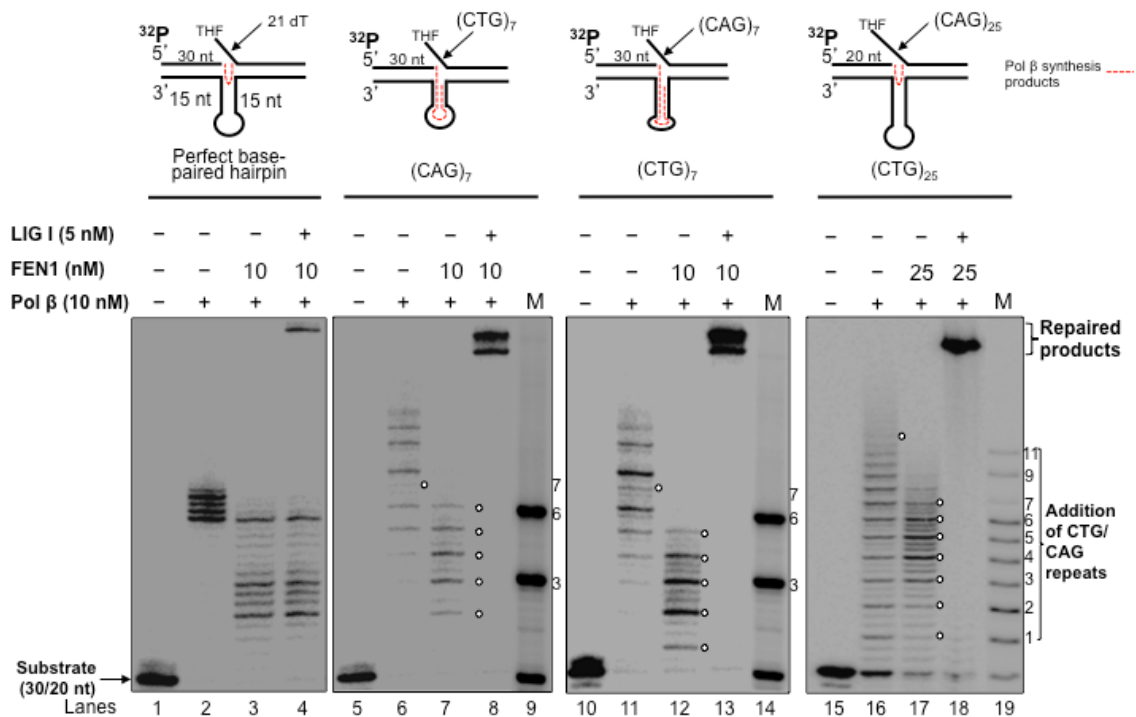
downstream flap, but inhibited the DNA synthesis on the substrates with a 5'-flap (compare lanes 2 and 3, 6 and 7, 11 and 12, and 16 and 17 of Figure 1.8). These results suggested that FEN1 suppressed pol  $\beta$  hairpin bypass synthesis, possibly by competing with the polymerase to bind a flap. To determine whether pol  $\beta$  hairpin bypass could result in repeat deletions, BER reactions were reconstituted using the same set of template hairpin-containing substrates and purified pol  $\beta$ , FEN1 and LIG I. The results showed that pol  $\beta$  bypass of the perfect base-paired hairpin and the (CTG)<sub>25</sub> hairpin led to repair products with deletions of 4–39 and 3–63 nt (1–21 repeats), respectively, which were identified by DNA sequencing (Figure 1.12 and 1.13). Consistent with the observation, pol  $\beta$  bypass of a (CAG)<sub>25</sub> hairpin also inserted 1–6 repeats and resulted in 3–60 nt (1–20 repeats) (Figure 1.14). Surprisingly, pol  $\beta$  bypass of (CTG)<sub>7</sub> and (CAG)<sub>7</sub> hairpins failed to cause any repeat deletion (Figure 1.12), indicating that it passed through only a portion of the large hairpins and the entire small hairpins. Thus, pol  $\beta$  can bypass a large hairpin directly leading to repeat unit loss, but can faithfully copy almost all the repeats that constitute a small hairpin. However, other repair steps that can shorten the repeats may also contribute to repeat deletion. During BER, this could result from FEN1 flap cleavage activity.



**Figure 1.7 Hairpin bypass synthesis of pol  $\beta$  with template hairpin containing substrates without a downstream flap.**

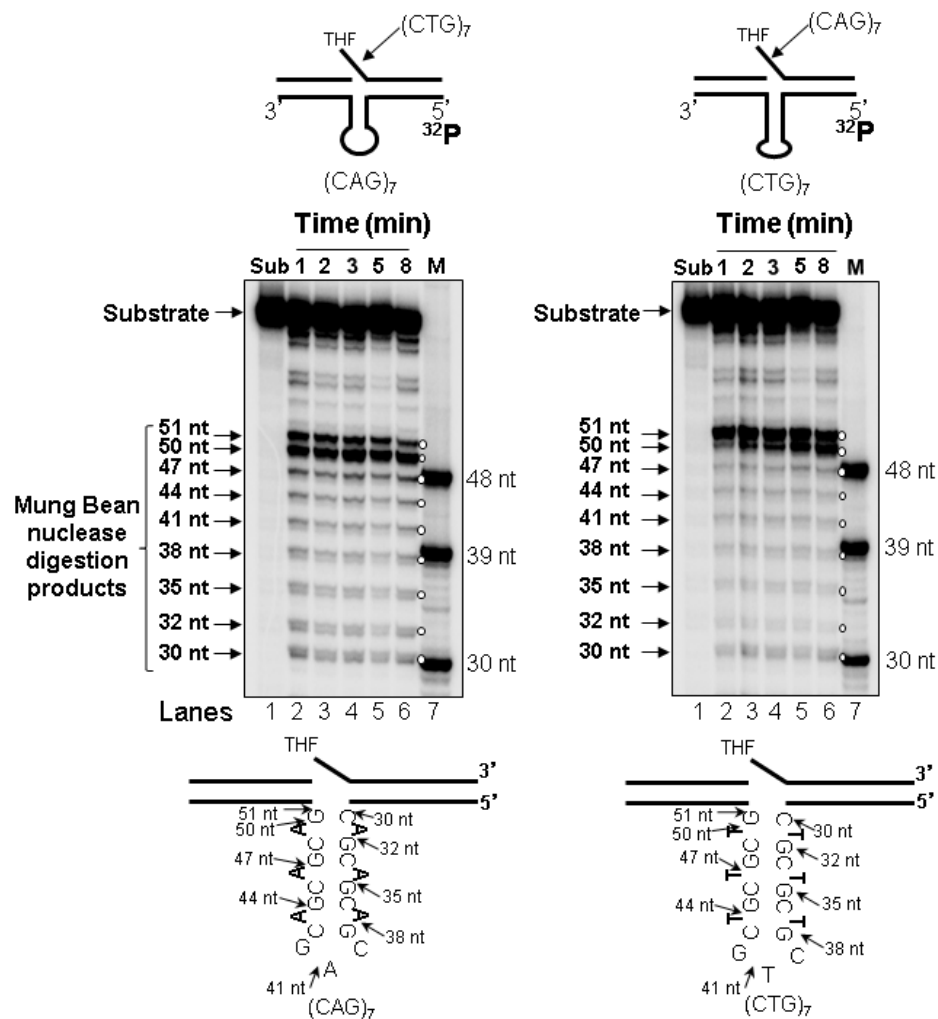
Pol  $\beta$  hairpin bypass synthesis and BER were examined in the context of a template (CAG)<sub>7</sub> or a (CTG)<sub>7</sub> hairpin or a (CTG)<sub>25</sub> hairpin or a template hairpin with a 6-nt loop and 15-nt perfect base pairs at its stem region. The substrates contained THF residues at the 5'-end of the downstream primers and were radiolabeled at the 5'-end of the upstream primers. BER reactions were reconstituted with 10 nM pol  $\beta$ , 10 or 25 nM FEN1 and 5 nM LIG I in the presence of 50  $\mu$ M dNTPs under the conditions described under 'Experimental Procedures' (lanes 4, 8, 13 and 18). Lanes 1, 5, 10 and 15 represent substrates only. Lanes 2, 6, 11 and 16 correspond to the reaction mixture with 10 nM pol  $\beta$ . Lanes 3, 7, 12 and 17 represent the reaction mixture in the presence of 10 nM pol  $\beta$  and 10 or 25 nM FEN1. Lanes 9, 14 and 19 represent size markers of pol  $\beta$  DNA synthesis. In all the reactions, 25 nM DNA substrates were used. The substrates are illustrated schematically above the gel. Pol  $\beta$  DNA synthesis products are illustrated by dotted lines.





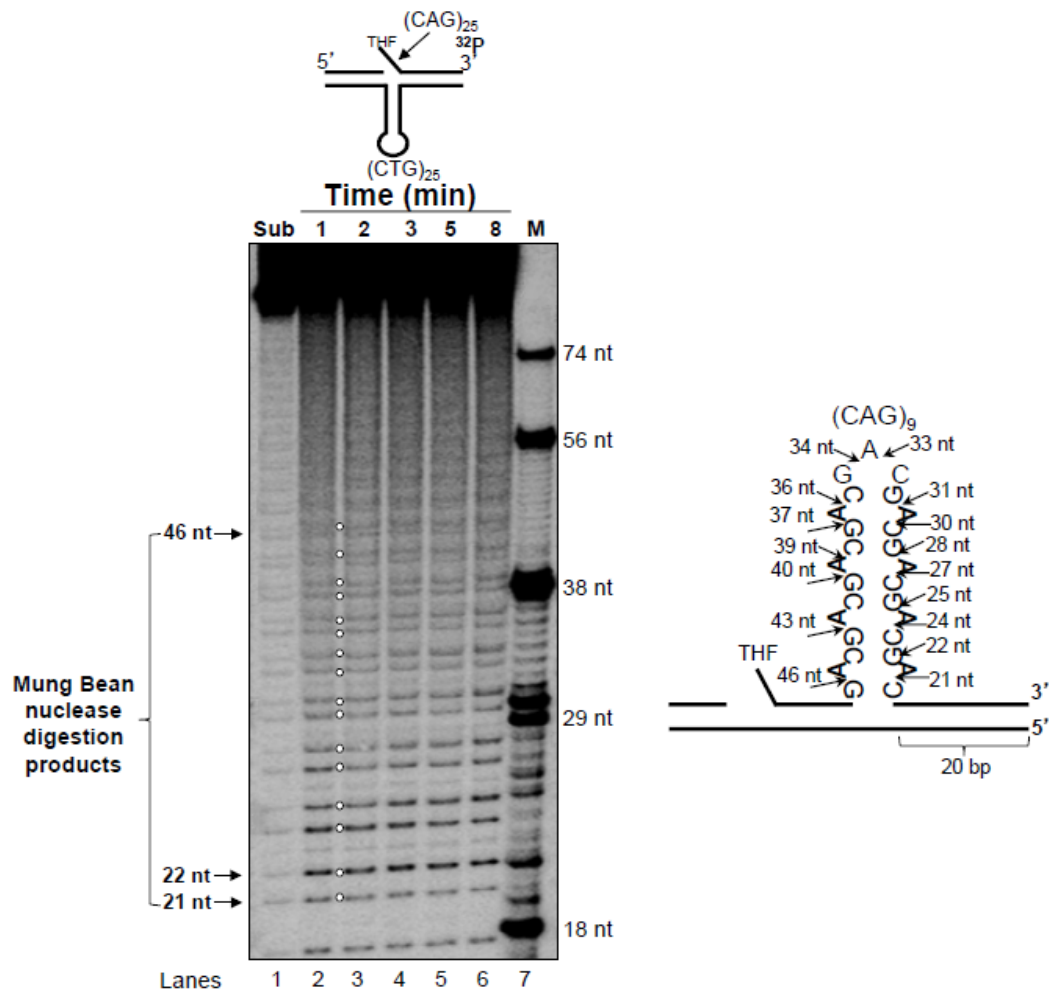
**Figure 1.8 Hairpin bypass synthesis of pol β with template hairpin containing substrates with a downstream flap.**

Pol β DNA synthesis and repair products were determined with substrates that contained a downstream 5'-THF containing (CTG)<sub>7</sub> or (CAG)<sub>7</sub> or (CAG)<sub>25</sub> flap/hairpin or 15 dT flap with a template hairpin. Lanes 1, 5, 10 and 15 represent substrates only. Lanes 2, 6, 11 and 16 correspond to the reaction mixtures with 10 nM pol β. Lanes 3, 7, 12 and 17 represent the reaction mixture in the presence of 10 nM pol β and 10 or 25 nM FEN1. Lanes 4, 8, 13 and 18 represent the reaction mixture reconstituted with 10 nM pol β, 10 or 25 nM FEN1 and 5 nM LIG I. Lanes 9, 14 and 19 represent size markers of pol β DNA synthesis. In all the reactions, 25 nM 5'-radiolabeled substrates were used. The substrates are illustrated schematically above the gel. Pol β DNA synthesis products are illustrated by dotted lines.



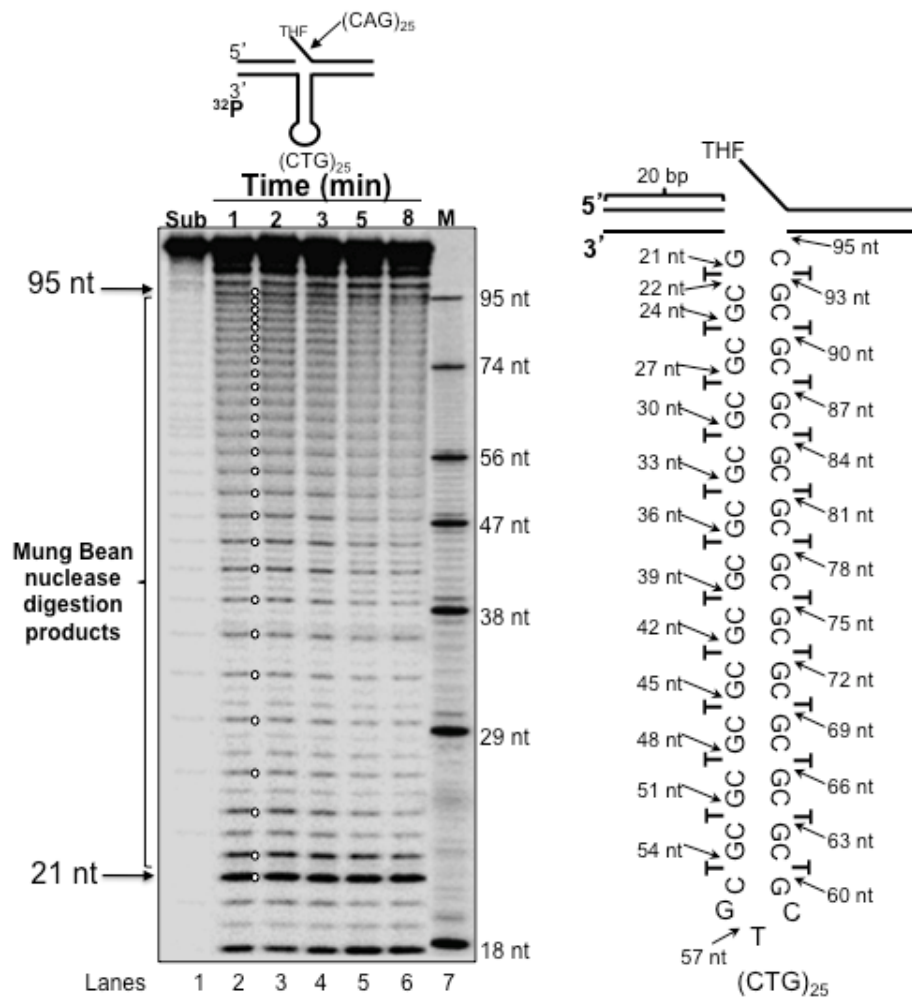
**Figure 1.9 Probing of hairpins formed by template strand containing (CAG)<sub>7</sub> or (CTG)<sub>7</sub> repeats.**

The template strand containing (CAG)<sub>7</sub> (left panel) or (CTG)<sub>7</sub> repeats (right panel) was radiolabeled at its 5'-end. The substrate was incubated with two units of Mung Bean Nuclease at 1-, 2-, 3-, 5-, and 8-minute time intervals (lanes 2–6). Lane 1 represents the undigested substrate. Lane 7 represents markers (M). Arrows and white dots indicate the major digestion products. The substrates are illustrated schematically above the gel. A hairpin deduced by a specific nuclease cleavage pattern and the nuclease digestion sites is illustrated schematically below the gel.



**Figure 1.10 Probing of hairpins formed by damaged strand with a 5'-(CAG)<sub>25</sub> flap.**

A substrate that contained a downstream 5'-flap/hairpin at the damaged strand and a template (CTG)<sub>25</sub> hairpin was radiolabeled at the 3'-end of the 5'-flap/hairpin. The substrate was digested with two units of Mung Bean Nuclease at 1-, 2-, 3-, 5-, and 8-minute time intervals (lanes 2–6). Lane 1 represents the undigested substrate. Lane 7 represents markers (M). For all the experiments, 200 nM of substrate was used. Arrows and white dots indicate the major digestion products. The substrates are illustrated schematically above the gel. A hairpin deduced by a specific nuclease cleavage pattern and the nuclease digestion sites is illustrated schematically beside the gel.



**Figure 1.11 Probing of hairpins formed (CTG)<sub>25</sub> repeat in the template strand.**

A substrate that contained a downstream 5'-flap/hairpin at the damaged strand and a template (CTG)<sub>25</sub> hairpin was radiolabeled at the 3'-end of its template strand. The substrate was digested with two units of Mung Bean Nuclease at 1-, 2-, 3-, 5-, and 8-minute time intervals (lanes 2–6). Lane 1 represents the undigested substrate. Lane 7 represents markers (M). For all the experiments, 200 nM of substrate was used. Arrows and white dots indicate the major digestion products. The substrates are illustrated schematically above the gel. A hairpin deduced by a specific nuclease cleavage pattern and the nuclease digestion sites is illustrated schematically beside the gel.

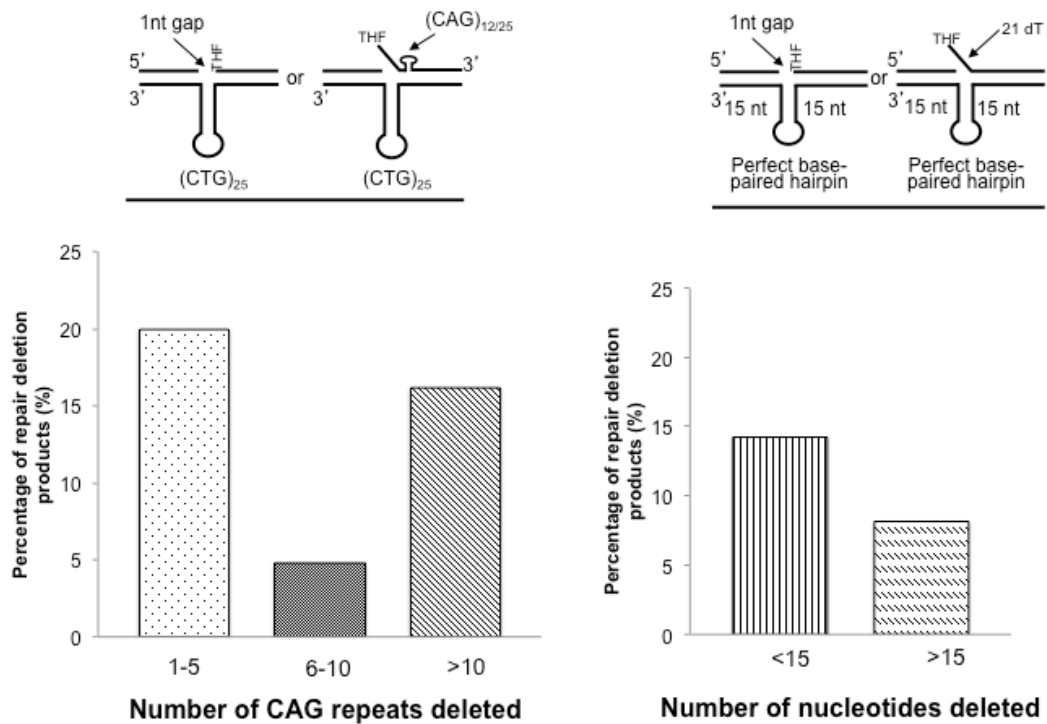
### DNA Sequencing Results of Repair Products

Substrate	Number of nucleotides deleted	Percentage of deletions
	4-39 nt	22% (11*/49*)
	0	0% (0/21)
	0	0% (0/18)
	3- 63nt	41% (43/105)

\*Number of the clones containing repair deletion products that were sequenced  
 ♦Total number of the clones containing repair products that were sequenced

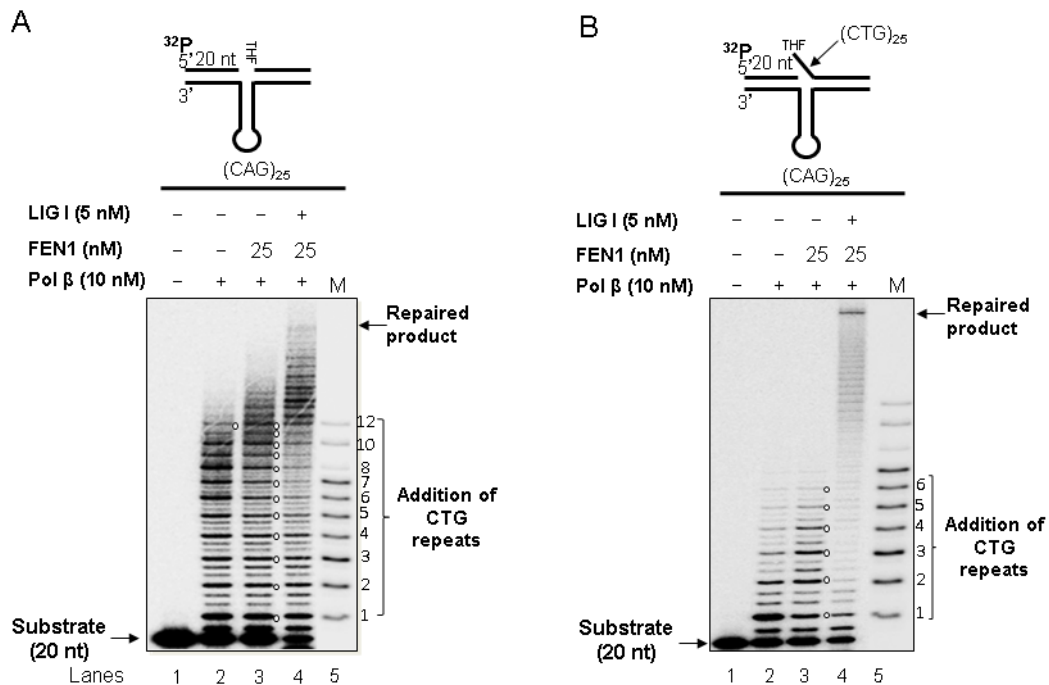
**Figure 1.12 The number of nucleotides deleted and the percentage of CAG repeat deletions resulting from BER.**

The products resulting from BER in the context of a small or large template hairpin were isolated with avidin beads, amplified by PCR and cloned into TA vector for DNA sequencing. The number of nucleotides deleted and the percentage of deletions were listed.



**Figure 1.13 Spectrums of CAG repeat deletions resulting from BER.**

The spectrums of repeat deletions from the substrate with a (CTG)<sub>25</sub> repeat-containing template hairpin and a perfect base-paired template stem-loop structure were illustrated.



**Figure 1.14 Hairpin bypass synthesis of pol β and FEN1 flap cleavage coordinate with each other leading to CTG repeat deletion.**

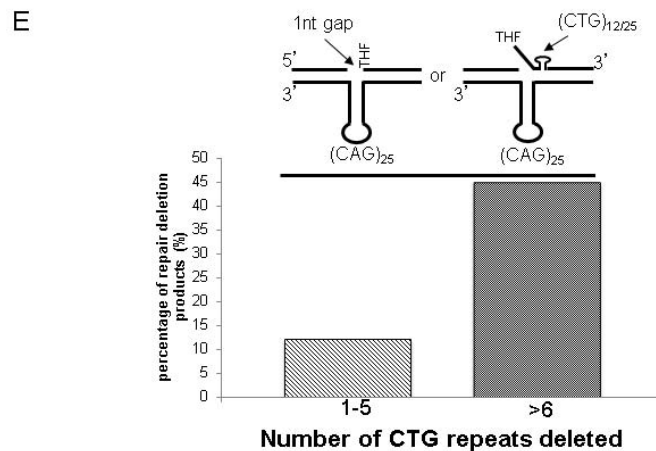
(A) Pol β hairpin bypass synthesis and BER were examined in the context of a template (CAG)<sub>25</sub> hairpin. The substrate contained THF residue at the 5'-end of the downstream primers, and were radiolabeled at the 5'-end of the upstream primers. BER reactions were reconstituted with 10 nM pol β, 25 nM FEN1, and 5 nM LIG I in the presence of 50 μM dNTPs under the conditions described under "Experimental Procedures" (lane 4). Lane 1 represents substrate only. Lane 2 corresponds to the reaction mixture with 10 nM pol β. Lane 3 represents the reaction mixture in the presence of 10 nM pol β and 25 nM FEN1. Lane 5 represents size markers of pol β DNA synthesis. (B) Pol β DNA synthesis and repair products were determined with substrates that contained a downstream 5'-THF-(CTG)<sub>25</sub> flap/hairpin. Lane 1 represents substrate only. Lane 2 corresponds to the reaction mixtures with 10 nM pol β. Lane 3 represents the reaction mixture in the presence of 10 nM pol β and 25 nM FEN1. Lane 4 represents the reaction mixture reconstituted with 10 nM pol β, 25 nM FEN1, and 5 nM LIG I. Lane 5 represent size markers of pol β DNA synthesis. In all the reactions, 25 nM 5'-radiolabeled substrates were used. All the substrates are illustrated schematically above the gel.

D

Substrate	Number of nucleotides deleted	Percentage of deletions
	3 - 60 nt	59.2% (29/49)

\*Number of the clones containing repair deletion products that were sequenced

◆Total number of the clones containing repair products that were sequenced



**Figure 1.14 Hairpin bypass synthesis of pol  $\beta$  and FEN1 flap cleavage coordinate with each other leading to CTG repeat deletion.**

(D) The number of nucleotides deleted and the percentage of CTG repeat deletion were listed. (E) The spectrum of repeat deletions from the substrate with a (CAG)<sub>25</sub> repeat-containing template hairpin was illustrated.

**FEN1 promotes CAG repeat deletion during BER**

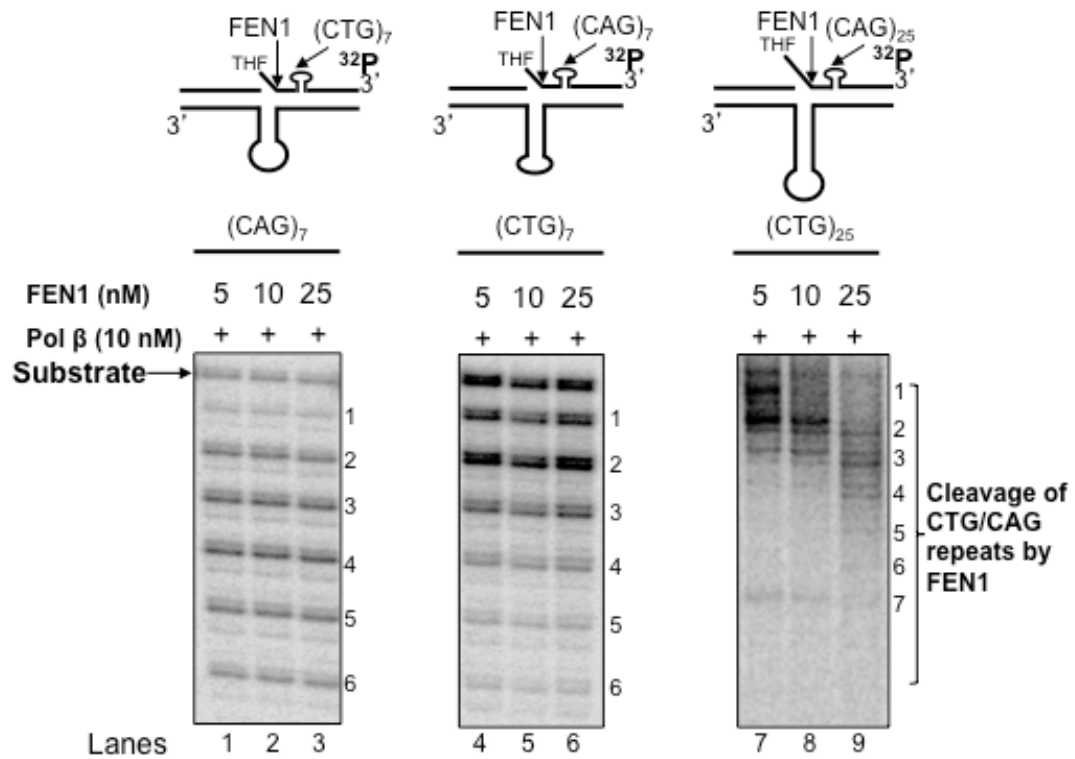
Flap endonuclease 1 has been found to play controversial roles in modulating TNR stability in replication and BER (112,166,192)(115). FEN1 cleavage on a 5'-TNR-flap during DNA replication is crucial for preventing TNR expansion (112,166,192). However, FEN1 alternate cleavage of a 5'-short flap attached to a hairpin during long-patch BER, promotes CAG repeat expansion (115). FEN1 flap cleavage may also be



involved in repeat deletion during BER by shortening repeat length. To test this possibility, we initially examined FEN1 cleavage during BER of a THF embedded in (CAG)<sub>20</sub> repeats. We found that with increasing concentrations of pol β (2.5, 5 and 10 nM), FEN1 (5, 10 and 25 nM) cleaved up to four CAG repeats in the presence of pol β (Figure 1.15, lanes 4–12). Pol β alone failed to give any cleavage products (Figure 1.15, lane 3) indicating the products were FEN1 specific. To further determine if the cleavage products may result from FEN1 cleavage in the presence of a template hairpin, we examined its cleavage on a 5'-THF-(CAG)<sub>7</sub> or 5'-THF-(CTG)<sub>7</sub> or 5'-THF-(CAG)<sub>25</sub> or 5'-THF-(CTG)<sub>25</sub> flap/hairpin in the presence of a template (CAG)<sub>7</sub> or (CTG)<sub>7</sub> or (CTG)<sub>25</sub> or (CAG)<sub>25</sub> hairpin (Figure 1.16 and Figure 1.17). We found that FEN1 removed up to five or six CAG or CTG repeats from the (CAG)<sub>7</sub> or (CTG)<sub>7</sub> flap/hairpin in the presence of pol β (5–25 nM) (Figure 1.16, lanes 1–6) which was similar to its cleavage during BER. Interestingly, FEN1 removed up to seven to eight repeats from the (CAG)<sub>25</sub> or (CTG)<sub>25</sub> flap/hairpin (Figure 1.16, lanes 7–9 and Figure 1.17) suggesting that a large hairpin allowed FEN1 to remove more repeats for larger deletion. To further determine if FEN1 cleavage is essential for repeat deletion, we examined repeat deletion in the absence and presence of FEN1 with a substrate containing a deoxyuridine that substituted the first guanine of (CAG)<sub>20</sub> repeats (Figure 1.18). We reasoned that removal of the base lesion by UDG would create an abasic site that results in a native 5'-sugar phosphate and a single-nucleotide gap, and these can be directly removed by pol β deoxyribose phosphate lyase and filled by pol β gap-filling synthesis without the need of FEN1 cleavage. Thus, no deletion products should be detected because the repeat length cannot be shortened by pol β. However, if FEN1 is included in the repair, it could remove short CAG repeats from

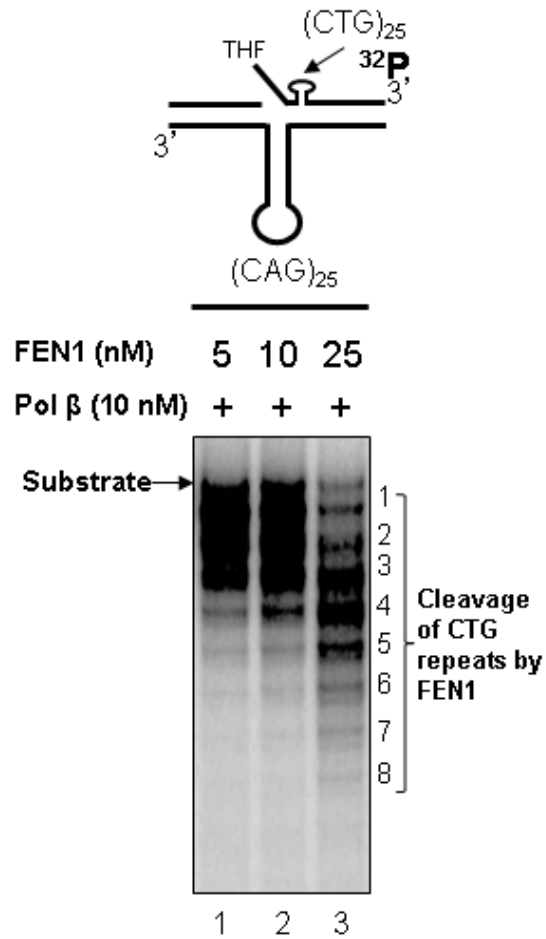
the 5'-end of a downstream hairpin, leading to shortening of repeats and deletions. In support of this idea, we identified a small amount of deletion product in the absence of FEN1 (Figure 1.18, panel b), but significant amounts of deletion products in the presence of 10 nM FEN1 (Figure 1.18, panel c). Quantitative analysis demonstrated that FEN1 increased the deletion by 3-fold (Figure 1.18, panel d). In contrast, reduction of FEN1 level in pol  $\beta^{+/+}$  MEF cell extracts by 5-fold through immuno-depletion (Figure 1.19) decreased the amount of deletion products only by 2-fold (Figure 1.18, panels f-i). The immuno-depletion of FEN1 did not affect pol  $\beta$  level in the cell extracts (Figure 1.19) indicating that reduction of deletion products was specifically from FEN1 deficiency. Thus, FEN1 is not essential for, but promotes TNR deletion. Our results indicate that FEN1 plays an important role in facilitating CAG repeat deletion by cooperating with pol  $\beta$  hairpin bypass synthesis during long-patch BER.



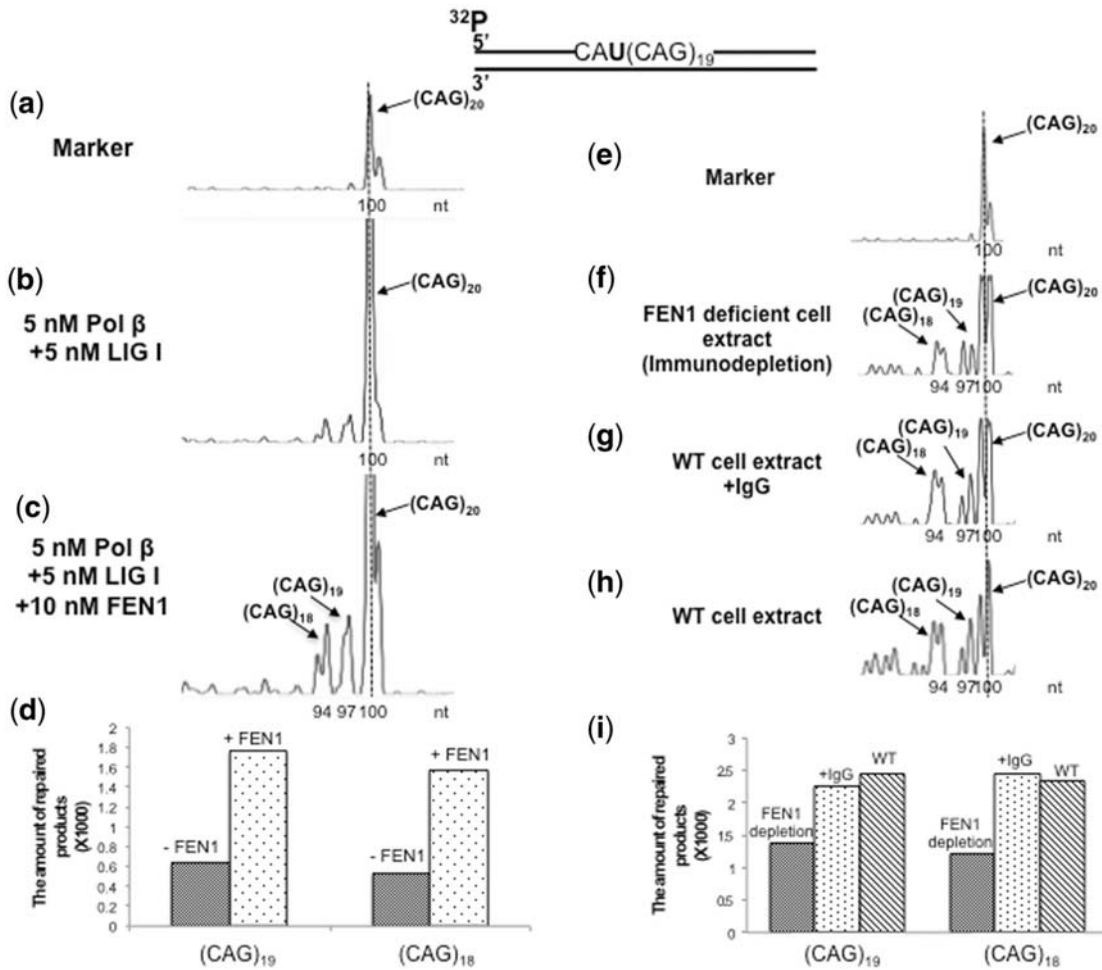


**Figure 1.16 FEN1 cleavage of repeats containing downstream flap during BER.**

FEN1 flap cleavage was examined in the context of a template hairpin with a substrate containing a downstream 5'-THF-(CAG)<sub>7</sub> or 5'-THF-(CTG)<sub>7</sub> or 5'-THF-(CAG)<sub>25</sub> hairpin/flap. Lanes 1–9 correspond to reaction mixtures with 25 nM substrate and 10 nM pol β in the presence of various concentrations of FEN1 (5, 10 and 25 nM). DNA substrate (25 nM) was used in all the reactions and FEN1 alternate flap cleavage sites are illustrated on the substrates shown above the gels.

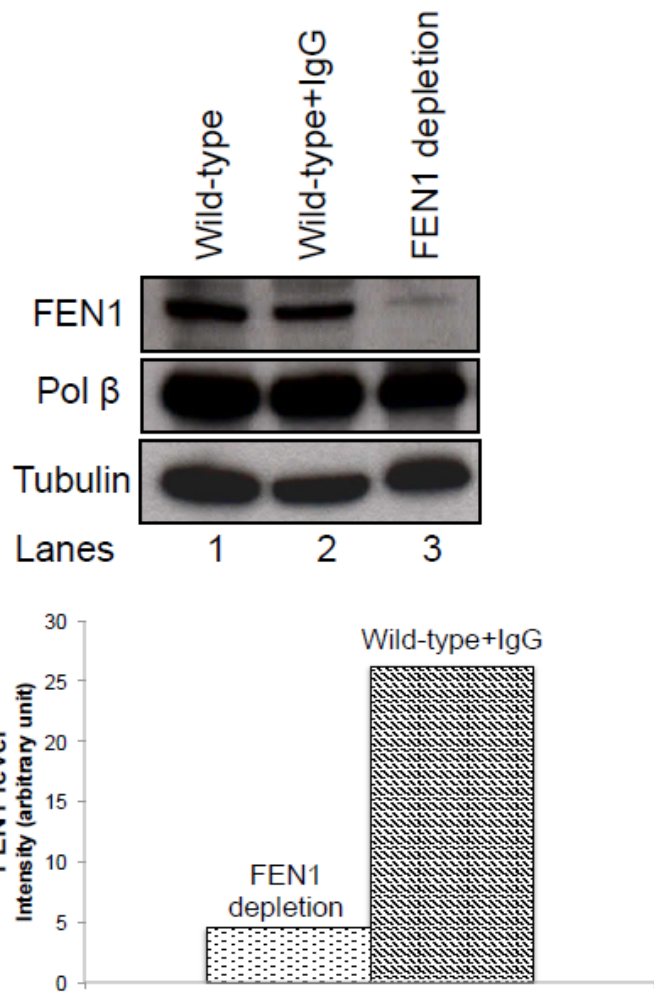


**Figure 1.17 FEN1 flap cleavage of a downstream 5'-THF-(CTG)<sub>25</sub> hairpin/flap.** FEN1 flap cleavage was examined in the context of a template hairpin with a substrate containing a downstream 5'-THF-(CTG)<sub>25</sub> hairpin/flap. Lanes 1–3 correspond to reaction mixtures with 25 nM substrate and 10 nM pol β in the presence of various concentrations of FEN1 (5 nM, 10 nM, and 25 nM). 25 nM 3'-radiolabeled DNA substrate was used in all the reactions.



**Figure 1.18 FEN1 promotes limited CAG repeat deletion.**

Repair of a deoxyuridine in the context of (CAG)<sub>20</sub> repeats was examined with an oligonucleotide substrate containing a deoxyuridine residue that substituted the guanine in the first CAG of (CAG)<sub>20</sub> repeats. BER reactions were reconstituted with 5 nM pol  $\beta$ , 2.5 nM UDG, 50 nM APE1 and 5 nM LIG I in the presence of 10 nM FEN1 or absence of FEN1 (panels b–d). BER reactions were also performed with cell extracts from pol  $\beta^{+/+}$  MEF cell extracts with FEN1 depletion, pol  $\beta^{+/+}$  MEF cell extracts with IgG and pol  $\beta^{+/+}$  MEF cell extracts (panels f–i). The sizes of the repair products were determined by DNA fragment analysis.



**Figure 1.19 Western-blot analysis of FEN1 depletion in MEF cell extract.**

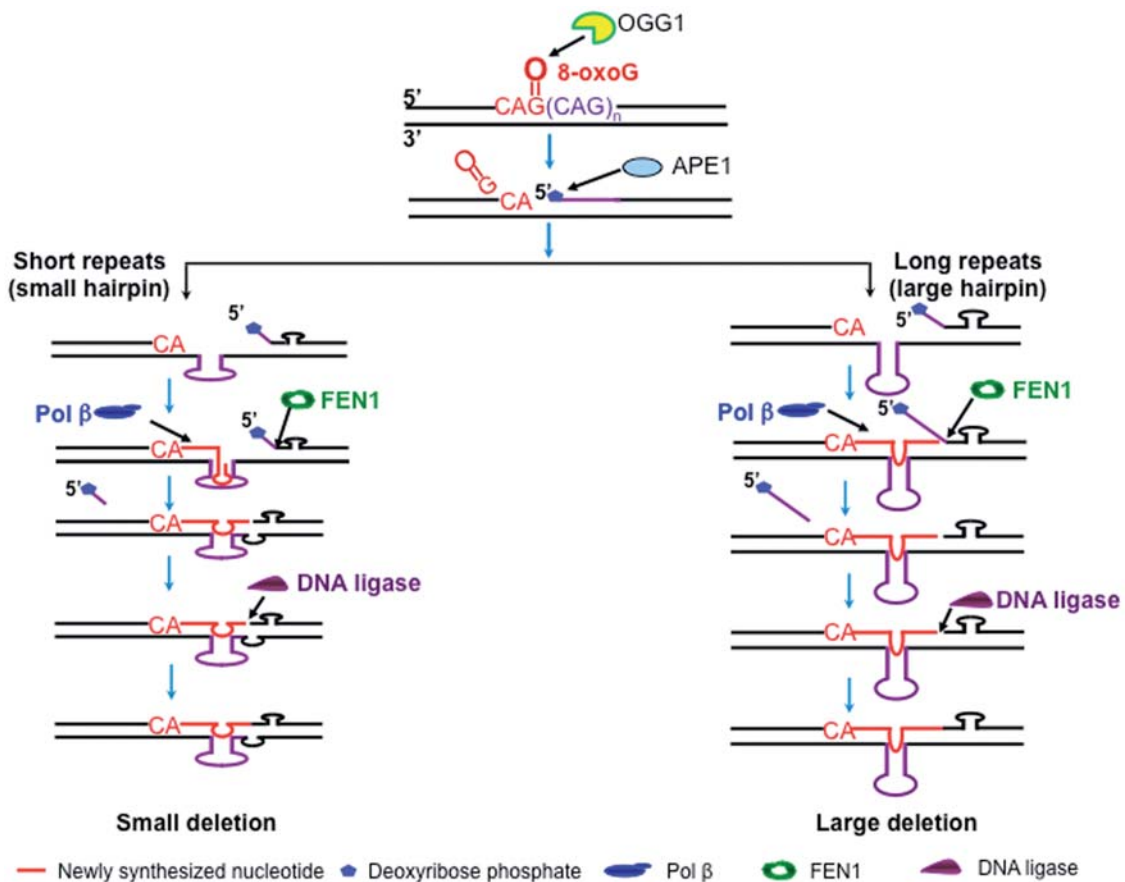
Mouse fibroblast cell extracts were subjected to immunodepletion using anti-FEN1 antibody as described in “Materials and methods.” Cell extracts from wild-type mouse cells were immunoprecipitated (IP) with normal mouse IgG as a negative control or anti-FEN1 antibody (lanes 2 and 3), and proteins were separated by SDS-PAGE and then blotted onto a membrane. The blot was probed for FEN1 (top panel) and then, after stripping, for Pol β (middle panel) or tubulin as a loading control (bottom panel). Lane 1 represents 60 μg of cell extract mixed directly with SDS sample buffer.

## DISCUSSION

In this study, for the first time, we discovered that CAG repeats can be deleted during BER (Figures 1.2, 1.3, 1.12, 1.13, 1.14 and 1.19). We provided direct evidence that repeat deletions are generated in both cell extract-based and reconstituted BER (Figure 1.2 and 1.3) and that the process is accomplished by pol  $\beta$  hairpin bypass synthesis in coordination with FEN1 alternate cleavage. We demonstrated that during BER, hairpin structures are generated on the damaged and template strands of (CAG)<sub>20</sub> repeats (Figure 1.4 and 1.5). These hairpins govern TNR deletion by cooperatively interacting with pol  $\beta$  hairpin bypass and FEN1 alternate flap cleavage. We discovered that pol  $\beta$  readily bypasses a small portion of a large template hairpin, directly resulting in loss of large size CAG repeats (Figure 1.7 and 1.8, lanes 16 and 17; Figure 1.12); however, pol  $\beta$  can pass through almost the entire length of a small template hairpin causing the loss of only one or two repeats (Figure 1.7 and 1.8, lanes 7, 8, 11 and 12; Figure 1.12). With the unique hairpin bypass of pol  $\beta$ , repeat deletions will have to be ultimately fulfilled through FEN1 cleavage. In supporting this notion, we found that FEN1 alternate flap cleavage removed up to four CAG repeats in the presence of a small template hairpin (Figure 1.15), while pol  $\beta$  hairpin bypass only inserted three or four repeats in the presence of FEN1 (Figure 1.8). Thus, FEN1 removes more repeats than those inserted by pol  $\beta$  leaving a gap that led to the formation of a loop containing one or two repeats on the template strand for small deletions. Consistent with this idea, we observed that FEN1 effectively cleaved six to seven repeat units from a pre-formed (CTG)<sub>7</sub> or (CAG)<sub>7</sub> flap/hairpins (Figure 1.16) and significantly stimulated small repeat



deletions (Figure 1.18). With longer repeats [i.e., (CAG)<sub>25</sub> or (CTG)<sub>25</sub>] that can form a large template hairpin, pol β skips over the hairpin by passing through a small portion of the hairpin allowing pol β to perform a strand-displacement synthesis to displace a longer repeat flap for FEN1 cleavage, thereby causing larger sizes of deletions (Figures 1.8, 1.12, 1.13 and 1.16). Our data support a model for TNR deletion during BER of oxidative DNA base lesions where ssDNA breakage is induced by OGG1 and APE1 (Figure 1.20), that in turn leads to DNA slippage and the formation of a 5'-downstream hairpin, template hairpin and multi-nucleotide gap. For short repeats, a small template hairpin allows pol β to pass through the large portion of the hairpin. FEN1 removes a short 5'-flap associated with a small downstream hairpin, leaving a gap, further allowing the formation of a loop with one or two repeats on the template leading to a ligatable nick, which upon nick sealing affords a repeat deletion (Figure 1.20, sub-pathway on the left). For a long repeat tract, ssDNA breakage leads to the formation of a large template hairpin that allows pol β to bypass only one small portion of the hairpin. The polymerase then skips over the hairpin to perform a strand displacement synthesis generating a long repeat flap that is removed by FEN1 leading to a large repeat deletion (Figure 1.20, right sub-pathway).



**Figure 1.20 Models for CAG repeat deletion during BER.**

An oxidized DNA base lesion in the context of CAG repeats is removed by OGG1 leaving an abasic site. APE1 incises the 5'-end of the abasic site resulting in ssDNA breaks. This allows the downstream strand to dissociate from its template forming various sizes of hairpins associated with short CAG repeat flaps. This results in multi-nucleotide gaps that further lead to the formation of a template hairpin. For short repeats, a small template hairpin allows pol b to pass through almost its entire length (left panel). FEN1 then removes one or two repeat units leaving a gap with one or two repeats. This then allows the template to loop out forming a ligatable nick that leads to small repeat deletion (left panel). For long repeats, a large template hairpin can form. This allows pol b to pass through one portion of the hairpin and displace downstream repeats for FEN1 to remove a large flap leading to large repeat deletion (right panel).

In this study, we also provide the first evidence that CAG repeat deletion during BER exhibits a length dependency with a relatively large deletion induced by long repeats that is similar to its expansion (71). The results suggest that TNR deletion and expansion share the same underlying molecular basis, i.e., the formation of hairpin structures. Our results indicate that the formation of a series of hairpin structures with varying sizes and stability dependent upon their length appear to govern TNR deletion. Thus, it is of interest to understand the length dependency of TNR instability mediated by BER, as this will further provide a basis for therapeutic purposes of TNR-related human diseases by manipulating the length of TNR through DNA repair.

Our observation of pol  $\beta$  passing through one part or an entire length of a template hairpin suggests that the polymerase can use its strand-displacement synthesis to pass through a hairpin. In this manner, pol  $\beta$  may destabilize hairpin structures, thereby preventing repeat deletion. It is possible that BER cofactors and the proteins from other repair pathways that can modulate pol  $\beta$  strand-displacement synthesis may help prevent or facilitate TNR deletion. For example, BER cofactor HMGB1 (193) that stimulates pol  $\beta$  strand-displacement synthesis may inhibit TNR deletion. However, proteins that can bind and stabilize hairpin structures such as MSH2/MSH3 (113,114) may inhibit pol  $\beta$  strand-displacement within a template hairpin and promote pol  $\beta$  skip-over of the hairpin facilitating repeat deletion. In contrast, DNA helicases that can destabilize hairpins, such as Bloom syndrome protein and Werner syndrome protein (194-196), may help pol  $\beta$  to pass through the entire length of a hairpin, thus preventing repeat deletion. The mechanisms by which BER proteins modulate TNR deletion through cooperating with the proteins from other repair pathways need to be further elucidated.

Cellular oxidative stress may induce single or multiple DNA base lesions at other locations, such as in the middle of repeat tracks or at the 3'-side of repeats, the locations of base lesions could result in the formation of different sizes of hairpins at specific sites of repeat tracks. This may modulate pol  $\beta$  hairpin bypass and FEN1 cleavage with different efficiency. It is also conceivable that multiple rounds of base lesion repair in the context of TNR could lead to the formation and accumulation of multiple template hairpins that could be processed by BER enzymes in a manner different from the one used for processing a single hairpin, resulting in large size deletion. The notion is implicated by a recent report showing that an abasic lesion located at various locations of CAG/CTG repeats is repaired by BER with different efficiency (197). The effects of the numbers and positions of DNA base lesions on TNR instability need to be further investigated.

Our study reveals that pol  $\beta$  hairpin bypass synthesis cooperates with FEN1 cleavage of a 5'-hairpin causing deletion of CAG repeats. Our previous work demonstrated that a robust pol  $\beta$  synthesis generates extra CAG repeats that can be ligated with a 5'-downstream hairpin shortened partially by FEN1 alternate flap cleavage promoting CAG repeat expansion (115). These findings suggest that repeat number synthesized by pol  $\beta$  and that cleaved by FEN1 governs the outcome of TNR instability as to expansion or deletion during BER. We suggest that expansion will occur if pol  $\beta$  synthesizes more repeats than those cleaved by FEN1, whereas deletion will occur if FEN1 removes more repeats than pol  $\beta$  synthesizes.

In summary, in this study, we identified an important role of oxidative DNA damage and BER in CAG repeat deletion for the first time. We provide the first evidence

that a base lesion-induced ssDNA breakage can lead to the formation of a hairpin in both the template and damaged strands of TNR tracks and that pol  $\beta$  employs its strand-displacement synthesis to pass through a portion or the entire length of the template hairpin determined by the size of hairpin structures. In contrast, FEN1 uses its alternate cleavage to process the hairpin on the damaged strand. Our results reveal that the cooperation between pol  $\beta$  hairpin bypass and FEN1 alternate cleavage results in small and large deletions. Our work further suggests that TNR deletion during DNA base damage and repair is governed by the interactions between TNR hairpins, pol  $\beta$  unique hairpin bypass and FEN1 alternate cleavage.

In CHAPTER 1, I have demonstrated that TNR deletion can be induced during BER of an oxidized DNA base lesion in the context of CAG repeats. For the first time we have discovered that CAG and CTG repeat deletion is caused by a template hairpin formed during BER. During the repair, pol  $\beta$  performed an efficient hairpin bypass synthesis to skip over a large hairpin by copying a portion of the hairpin, and FEN1 removed a TNR-containing flap via its alternate flap cleavage activity. In this manner, FEN1 removes more repeats than pol  $\beta$  synthesizes, thereby resulting in repeat deletion. Our results indicate that TNR deletion during BER is mediated by the coordination between pol  $\beta$  hairpin bypass and FEN1 alternate flap cleavage. In this study, we have identified a new mechanism underlying TNR deletion during BER of oxidative DNA damage. This will further allow us to explore the possibility for developing BER enzymes as a new target for prevention and treatment of TNR expansion-related neurodegenerative diseases.

Because expanded TNRs in the human genome usually exist in a form of hairpin or loop structures, and this allows TNR to be more susceptible to oxidative damage induced by ROS, repair of a base lesion in a hairpin loop should play a crucial role in preventing TNR expansion. Thus, a study on BER of an oxidized base lesion in the hairpin loop region is important for identifying BER as a target for prevention of TNR expansion. To explore a possible role of BER in preventing TNR expansion, we further examined if BER of an oxidized base lesion in the loop region of a TNR hairpin may disrupt the hairpin structure and if this could lead to the removal of the hairpin. I have demonstrated this part of work in CHAPTER 2.

**CHAPTER 2: BASE EXCISION REPAIR OF OXIDATIVE DNA DAMAGE  
COUPLED WITH REMOVAL OF A CAG REPEAT HAIRPIN ATTENUATES  
TRINUCLEOTIDE REPEAT EXPANSION**

**ABSTRACT**

Trinucleotide repeat (TNR) expansion is responsible for numerous human neurodegenerative diseases. However, the underlying mechanisms remain unclear. Recent studies have shown that DNA base excision repair (BER) can mediate TNR expansion and deletion by removing base lesions in different locations of a TNR tract, indicating that BER can promote or prevent TNR expansion in a damage location-dependent manner. In this study, we provide the first evidence that the repair of a DNA base lesion located in the loop region of a CAG repeat hairpin can remove the hairpin, attenuating repeat expansion. We found that an 8-oxoguanine located in the loop region of CAG hairpins of varying sizes was removed by OGG1 leaving an abasic site that was subsequently 5'-incised by AP endonuclease 1, introducing a single-strand breakage in the hairpin loop. This converted the hairpin into a double-flap intermediate with a 5'- and 3'-flap that was cleaved by flap endonuclease 1 and a 3'-5' endonuclease Mus81/Eme1, resulting in complete or partial removal of the CAG hairpin. This further resulted in prevention and attenuation of repeat expansion. Our results demonstrate that TNR expansion can be prevented via BER in hairpin loops that is coupled with the removal of TNR hairpins.

## INTRODUCTION

Trinucleotide repeat (TNR) expansion is associated with >40 neurodegenerative diseases such as Huntington's disease (75) [(CAG)<sub>n</sub>/(CTG)<sub>n</sub>], myotonic dystrophy [(CTG)<sub>n</sub>/(CAG)<sub>n</sub>], fragile X syndrome [(CGG)<sub>n</sub>/(CCG)<sub>n</sub>] and Friedreich's ataxia [(GAA)<sub>n</sub>/(TTC)<sub>n</sub>] (12,71,72). The molecular basis underlying TNR expansion is proposed to be the formation of non-B form DNA secondary structures during DNA replication, repair and recombination, including hairpins, tetraplexes and triplexes (85,140,183,198)(4,107). These structures can result in a series of effects that ultimately lead to TNR expansion. These include DNA polymerase stalling, which further causes DNA slippage and misalignment (4,108), replication fork stalling (86,98), inhibition of cleavage of hairpins by flap endonuclease 1 (FEN1) (110-112), trapping of mismatch repair proteins (113,114) and disruption of the coordination among repair enzymes (115). Thus, TNR expansion is a consequence of genome integration of unresolved DNA secondary structures formed during DNA replication, repair and recombination.

TNRs contain a high proportion of guanines, which are susceptible to loss of electrons and oxidation (149). This causes TNRs to be hotspots of oxidative DNA damage. Repair of oxidative DNA damage in the context of TNR tracts was previously shown to be associated with repeat expansion and deletion (103,199,200). An increased level of 8-oxoguanine (8-oxoG) was shown to be correlated with age-dependent CAG repeat expansions in the brain of HD transgenic mouse models (103,153). The oxidative DNA damage from H<sub>2</sub>O<sub>2</sub> caused small CAG repeat expansions in human cells (103,200), and the damage from potassium bromate increased the levels of 8-oxoG and CGG repeat expansions in the germ cells of fragile X syndrome premutation mice (187). H<sub>2</sub>O<sub>2</sub> can



also increase TNR deletions in bacteria and mouse kidney cells (26,177). We recently discovered that base excision repair (BER) of an abasic lesion can lead to CAG repeat deletion by inducing the formation of a template hairpin, promoting the coordination between DNA polymerase  $\beta$  (pol  $\beta$ ) hairpin bypass and FEN1 alternate flap cleavage (199). Furthermore, we have found that an abasic lesion located at different positions of a CTG repeat tract can result in either repeat expansion or deletion (200). Thus, oxidative DNA damage and BER play dual roles in modulating TNR instability.

Unpaired bases were found to be 10-fold more accessible to oxidative DNA damaging agents than paired bases in duplex DNA (201). Guanines located at the loop region of a TNR hairpin are hypersensitive to oxidative DNA damaging agents such as peroxynitrite (202). Moreover the oxidized base lesion 8-oxoG located in the stem region of a CAG repeat hairpin is preferentially relocated to the loop region via a dynamic intrastrand rearrangement of the repeats (203). However, it was shown that a base lesion in a hairpin loop that was detached from a CAG repeat tract was refractory to cleavage by 8-oxoG DNA glycosylase (OGG1) (204). This results in the accumulation of oxidative DNA damage in TNR hairpin loops and persistence of hairpin structures, leading to multiple rounds of 'toxic oxidation cycles' and causing TNR expansion (12,204). Thus, removal of an oxidized base lesion in the loop region of a TNR hairpin appears to be a critical step for preventing toxic oxidation cycles and TNR expansion. However, it is unknown how a base lesion in a TNR hairpin can be removed for prevention of TNR expansion.

In this study, we provide the first evidence that an 8-oxoG in the loop of a TNR hairpin in a CAG repeat tract can be simultaneously removed with the hairpin via BER.

This represents a novel pathway for preventing TNR expansion through BER coupled with removal of TNR hairpins.

## **MATERIALS AND METHODS**

### **Materials**

DNA oligonucleotides containing an 8-oxoG were synthesized by Eurofins MWG Operon (Huntsville, AL, USA). All other oligonucleotides were from Integrated DNA Technologies (IDT, Coralville, IA, USA). Deoxynucleoside 5'-triphosphates (dNTPs) and terminal deoxynucleotidyl transferase were from Fermentas (Glen Burnie, MD, USA). T4 polynucleotide kinase was from USB Corp. (Cleveland, OH, USA). All standard chemical reagents were from Sigma-Aldrich (St. Louis, MO, USA) and Thermo Fisher Scientific (Pittsburgh, PA, USA). The radionucleotides [ $\gamma$ - $^{32}\text{P}$ ] ATP (6000 mCi/mmol) and Cordycepin 5'-triphosphate 3'-[ $\alpha$ - $^{32}\text{P}$ ] (5000 mCi/mmol) were purchased from PerkinElmer Inc. (Boston, MA, USA). Micro Bio-Spin 6 chromatography columns were from Bio-Rad (Hercules, CA, USA). The Mus81/Eme1 expression vector pET21d-Mus81/HIS-Eme1 was a generous gift from Dr Stephen West at Clare Hall Laboratories, London Research Institute, Cancer Research UK, Hertfordshire, UK. Purified OGG1, AP endonuclease 1 (APE1), pol  $\beta$  and DNA ligase I (LIG I) were generous gifts from Dr Samuel Wilson at the National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC, or were expressed and purified according to the procedures described previously (205).

## **Oligonucleotide substrates**

Substrates containing a (CAG)<sub>7</sub> or (CAG)<sub>14</sub> hairpin with an 8-oxoG or a tetrahydrofuran (THF), an abasic site analog (used to represent an oxidized sugar in this study), located in the hairpin loop region were designed to mimic a TNR hairpin with a base lesion in the hairpin loop region. Substrates were constructed by annealing the damaged strands that contained (CAG)<sub>13</sub> or (CAG)<sub>20</sub> repeats to their template strand containing (CTG)<sub>7</sub> repeats at a molar ratio of 1:1.5. Substrates with a 3'- and 5'-flap were constructed to mimic double-flap intermediates generated by APE1 5'-incision of an abasic site in the loop region of CAG repeat hairpins of varying sizes. Substrates were constructed by annealing an upstream primer containing a (CAG)<sub>4</sub> or a (CAG)<sub>7</sub> flap/hairpin, and a downstream primer containing a 5'-phosphorylated THF residue attached to a (CAG)<sub>3</sub> or a (CAG)<sub>7</sub> flap/hairpin with the template strand containing (CTG)<sub>7</sub> repeats at a molar ratio of 1:1:1.5. For all the substrates, three CAG repeats flanked both the 5'- and 3'-side of the hairpins or flaps and base paired with three CTG repeats on the template strand. Substrates were radiolabeled at the 3'- or 5'-end of the damaged strand, the upstream primer or the downstream primer for measuring various types of enzymatic activities. The sequences of oligonucleotide substrates are listed in Table 2.1. DNA size markers were synthesized by IDT.

**Table 2.1 Oligonucleotide sequence**

Oligonucleotides	nt	Sequence (5' –3')
Downstream		
D1	40	pF- CAG CAGCAGCAGCAGCAG TA CGT AGA CTT ACT CAT TGC
D2	50	pF- CAG CAGCAGCAGCAGCAGCAGCAGCAG TA CGT AGA CTT ACT CAT TGC
D3	79	CGA GTC ATC TAG CAT CCG TA CAG CAGCAGCAGCAGCAG CA <b>8oxo-G</b> CAG CAGCAGCAGCAGCAG TA CGT AGA CTT ACT CAT TGC
D4	100	CGA GTC ATC TAG CAT CCG TA CAG CAGCAGCAGCAGCAGCAGCAGCAG CA <b>8oxo-G</b> CAG CAGCAGCAGCAGCAGCAGCAGCAG TA CGT AGA CTT ACT CAT TGC
D5	78	CGA GTC ATC TAG CAT CCG TA CAG CAGCAGCAGCAGCAG CA <b>F</b> CAG CAGCAGCAGCAGCAG TA CGT AGA CTT ACT CAT TGC
D6	100	CGA GTC ATC TAG CAT CCG TA CAG CAGCAGCAGCAGCAGCAGCAGCAG CA <b>F</b> CAG CAGCAGCAGCAGCAGCAGCAGCAG TA CGT AGA CTT ACT CAT TGC
Template		
T1	61	GCA ATG AGT AAG TCT ACG TAC TGC TGCTGCTGCTGCTGC TGT ACG GAT GCT AGA TGA CTC G
Upstream		
U1	40	CGA GTC ATC TAG CAT CCG TA CA GCA GCAGCAGCAGCAGCA
U2	49	CGA GTC ATC TAG CAT CCG TA CAG CAGCAGCAGCAGCAGCAGCAGCAG CA

F: tetrahydrofuran, THF 8-oxoG: 8-oxoguanine

## **Protein expression and purification**

FEN1 and Mus81/Eme1 were expressed in *E. coli* according to the procedures described previously (205,206). Briefly, cell pellets were lysed by a French press in lysis buffer containing 50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), pH 7.5, 30 mM NaCl, 1 mM dithiothreitol (DTT), 1 mM EDTA, 1 mM phenylmethylsulfonyl fluoride (PMSF), 1 mM benzamidine, 1 µg/ml pepstatin A and 1 µg/ml leupeptin. Cell lysates were then subjected to centrifugation at 12,000 rpm at 4 °C for 30 min to separate soluble proteins from cell debris. For FEN1 purification, the supernatant from cell lysates was initially subjected to purification by a 10-ml Sepharose Q column operated by the AKTA Fast Protein Liquid Chromatography system (FPLC; GE Healthcare, Piscataway, NJ, USA). The flow-through was collected and dialyzed into buffer containing 50 mM HEPES, pH 7.5, 30 mM NaCl and 1 mM PMSF, and subsequently loaded onto a 5-ml CM column (Bio-Rad, Hercules, CA, USA). Fractions were obtained with elution by a linear gradient of NaCl ranging from 30 mM to 2 M, and peak fractions were combined and dialyzed into buffer containing 50 mM HEPES, pH 7.5, 30 mM NaCl and 1 mM PMSF and loaded onto a 1-ml Mono-S column (GE Healthcare, Piscataway, NJ, USA). Peak fractions were obtained by elution with a linear gradient of NaCl from 30 mM to 2 M. The activity of purified FEN1 was measured by monitoring FEN1 cleavage of a 3-nt flap that was radiolabeled at the 5'-end of the flap. Purified FEN1 was aliquoted and frozen at -80°C until further use.

Mus81/Eme1 heterodimer was expressed in the *E. coli* BL21(DE3)-RP strain (Agilent Technologies, Santa Clara, CA, USA) transformed with the expression vector pET21d-Mus81/HIS-Eme1. Two hundred milliliters of lysogeny broth (LB) medium with

50 µg/ml ampicillin was inoculated with a newly transformed single colony and shaken at 225 rpm at 37 °C overnight. The overnight culture was then inoculated into three 1-liter LB cultures using 60 ml of overnight culture per liter. The newly inoculated culture was incubated at 30 °C with shaking until absorbance at 600 nm reached 1.0. Subsequently, the expression of Mus81/Eme1 was induced by 10 µM of isopropyl β-D-1-thiogalactopyranoside (IPTG) at 30 °C for 3 h. Cells were harvested by centrifugation at 3,000 rpm and 4 °C, for 40 min. Cell pellets were resuspended in lysis buffer containing 50 mM sodium phosphate, pH 7.0, 0.01% NP-40, 10% glycerol, 0.1 mM NaCl, 1 mM PMSF, 1 mM benzamide, 1 µg/ml pepstatin A and 1 µg/ml leupeptin, and lysed by a French press. The cell lysates were centrifuged at 12,000 rpm, and 4 °C, for 30 min. Supernatant was loaded onto a 10-ml P11 phospho-cellulose column (GE Healthcare, Piscataway, NJ, USA) using the AKTA FPLC system, and proteins were eluted with a linear gradient of NaCl ranging from 0.1 to 1 M. Peak fractions were collected and subjected to purification by Dynabeads coated with cobalt (Novex/Life Technologies Corp., Grand Island, NY, USA). Fractions (350 µl) were incubated with 50 µl of beads and 350 µl of binding buffer containing 50 mM sodium phosphate, pH 8.0, 300 mM NaCl and 0.001% Tween-20, and were rotated at 4 °C for 2 h. Protein-bead complexes were pelleted via a magnet at 4 °C for 2 min. Pellets were then washed with 300 µl of buffer containing 50 mM imidazole, three times. Mus81/Eme1 protein complexes were eluted with 100 µl of buffer containing 300 mM imidazole. The activity of Mus81/Eme1 was determined by measuring its cleavage of a 9 nt-3'-flap substrate that was radiolabeled at the 5'-end of the 3'-flap. Purified Mus81/Eme1 was aliquoted and frozen at -80°C until further use.

### ***In Vitro* BER reconstituted with purified enzymes**

*In vitro* BER of an 8-oxoG or oxidized abasic site (THF) in the loop region of CAG repeat hairpins with varying sizes was performed by incubating 25 nM substrates containing a (CAG)<sub>7</sub> or (CAG)<sub>14</sub> hairpin with an 8-oxoG or a THF residue in the hairpin loop with indicated concentrations of OGG1, APE1, pol β, Mus81/Eme1, FEN1 and LIG I. BER of single-strand DNA (ssDNA) break intermediates resulting from APE1 5'-incision of an abasic site in the hairpin loops were reconstituted by incubating purified Mus81/Eme1, pol β, FEN1 and LIG I with 25 nM substrates that contained different sizes of 3'- and 5'-flaps. Reactions were performed with BER reaction buffer containing 50 mM Tris-HCl, pH 7.5, 50 mM KCl, 0.1 mM EDTA, 0.1 mg/ml bovine serum albumin, 0.2 mM DTT and 0.01% Nonidet P-40 with 5 mM MgCl<sub>2</sub>, 2 mM ATP and 50 μM dNTPs. Reaction mixtures were incubated at 37 °C for 30 min, and terminated by incubation at 95 °C for 10 min in stopping buffer containing 95% formamide and 10 mM EDTA. Substrates and products were separated by 15% urea-denaturing polyacrylamide gel electrophoresis and detected by a Pharos FX Plus PhosphorImager from Bio-Rad. Substrates were <sup>32</sup>P-labeled at the 3'- or 5'-end of the strand with an 8-oxoG or a THF residue, the upstream primer and downstream primer.

### **Probing of hairpin structures and 3'- and 5'-flaps by Mung Bean Nuclease and S1 Nuclease digestion**

Formation of CAG repeat hairpins in hairpin-containing substrates or the 3'- and 5'-flaps in double-flap substrates was probed using Mung Bean Nuclease according to the procedure described by Xu et al. (199). Briefly, substrates (200 nM) were incubated with

0.1 or 0.15 U Mung Bean Nuclease in its reaction buffer at 37 °C for 1, 2, 3 and 5 min. Formation of CAG repeat hairpins in the substrates was also probed using 0.2 or 0.5 U S1 Nuclease and 200 nM substrates. Reactions were assembled in S1 Nuclease reaction buffer containing 50 mM sodium acetate (pH 4.5), 280 mM NaCl and 4.5 mM ZnSO<sub>4</sub> and incubated at 37 °C for 3, 5 and 10 min. Reaction mixtures were subsequently subjected to protease K digestion at 55 °C for 30 min to remove the nucleases. Substrates and products were separated by 15% urea-denaturing polyacrylamide gel electrophoresis and detected by a PhosphorImager.

## **RESULTS**

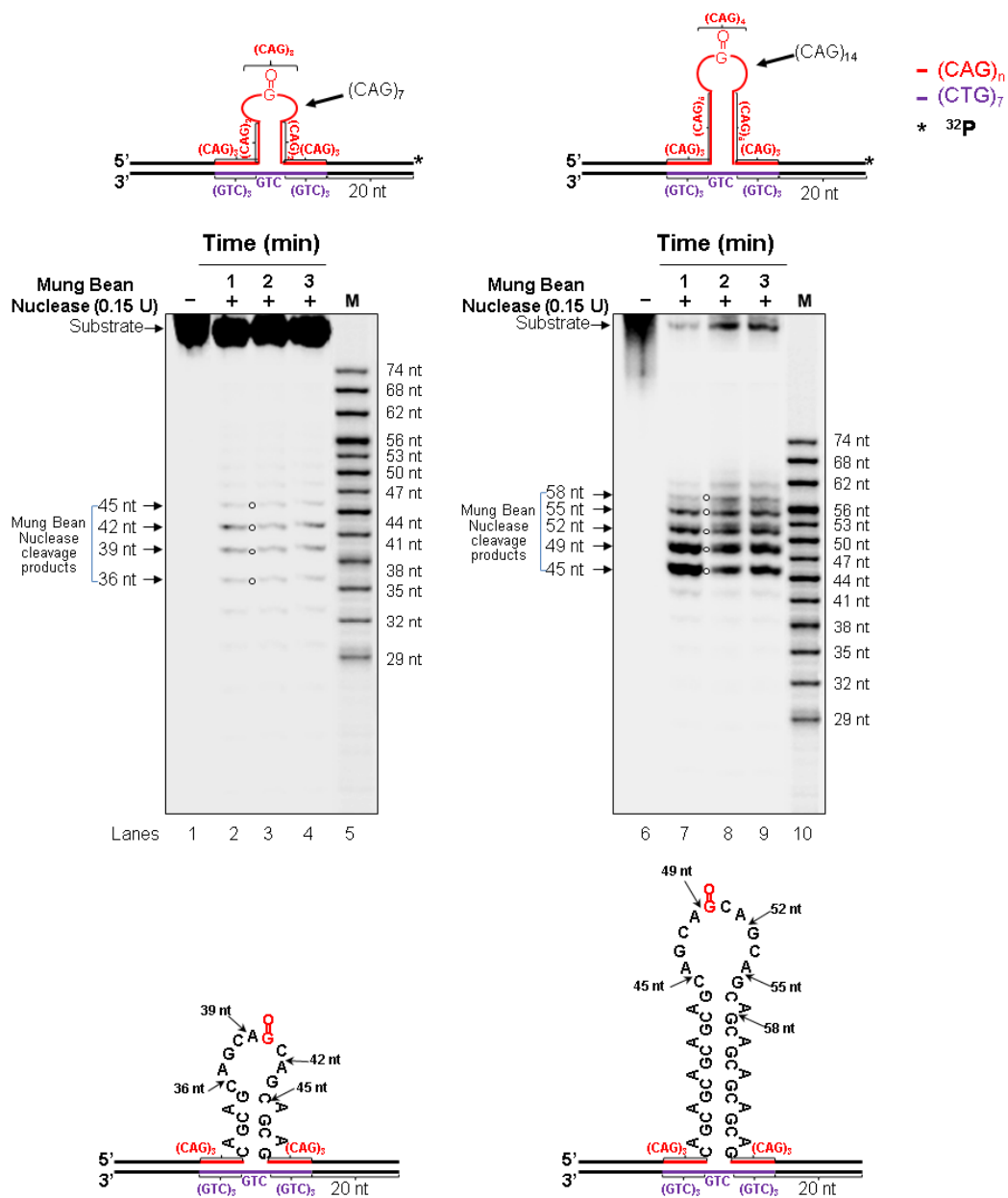
### **A DNA base lesion located in a CAG repeat hairpin loop can be removed by OGG1 and APE1**

OGG1 is a key enzyme that initiates BER in mammalian cells by removing an oxidized base lesion, 8-oxoG, leaving an abasic site that is incised by APE1 (11). The enzyme removes an 8-oxoG in a CAG hairpin loop that is detached from duplex DNA 700-fold slower than it does in duplex DNA (204). To determine if OGG1 can remove an 8-oxoG in a CAG hairpin loop that is located in duplex DNA, we examined OGG1 cleavage of an 8-oxoG in the loop region of a (CAG)<sub>7</sub> or (CAG)<sub>14</sub> hairpin in a CAG repeat tract. To verify the presence of an 8-oxoG and abasic site (THF residue) in the loop region of the substrates, we initially probed the formation of hairpins in the CAG repeat tract of the substrates using ssDNA-specific nucleases, Mung Bean Nuclease and S1 Nuclease. We found that Mung Bean Nuclease cleavage on the (CAG)<sub>7</sub>-8-oxoG



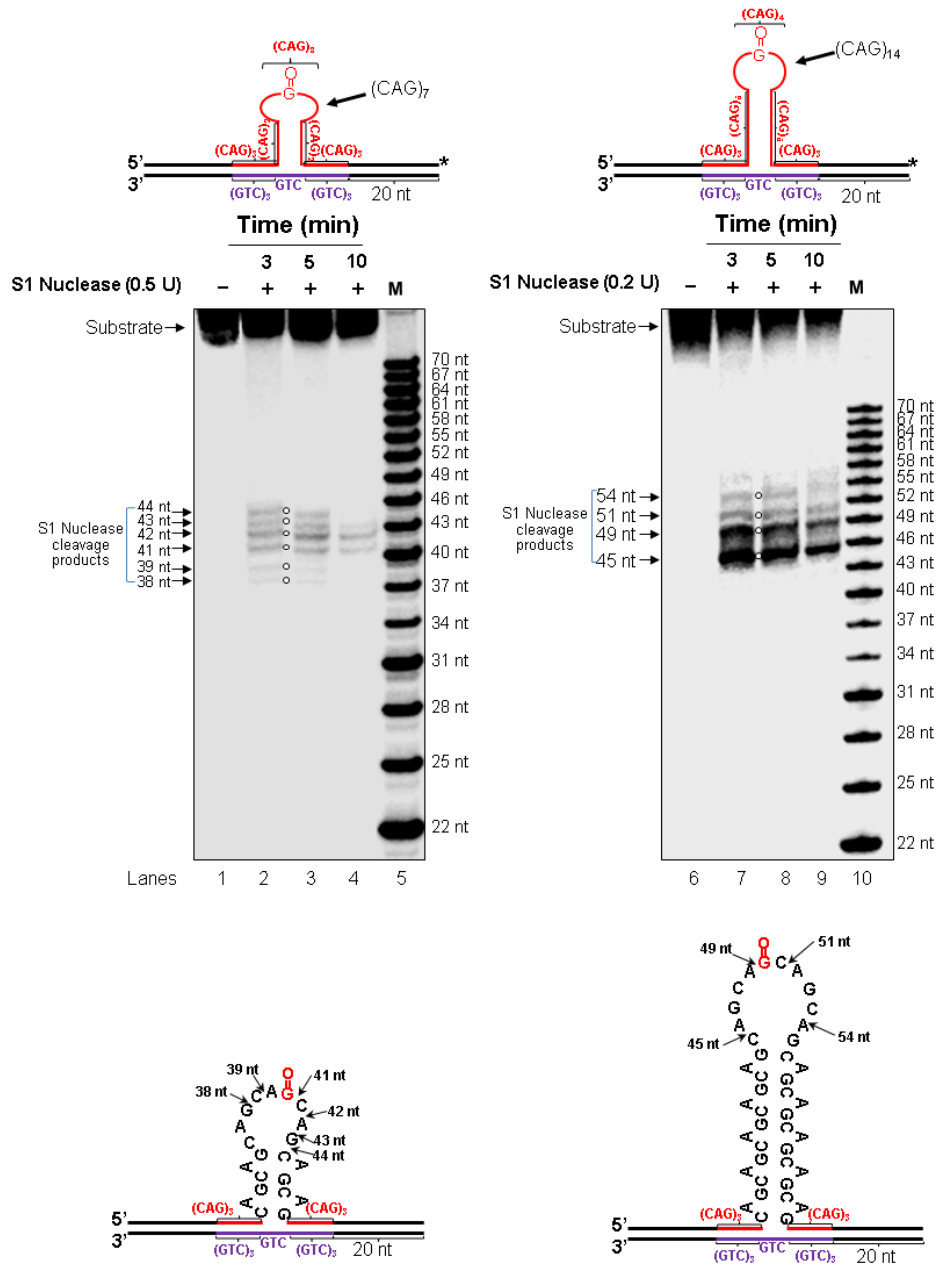
hairpin substrate mainly resulted in products with 36, 39, 42 and 45 nt (Figure 2.1, lanes 2–4), whereas its cleavage on the (CAG)<sub>7</sub>-THF substrate resulted in products with 35, 39, 41 and 44 nt (Figure 2.1, lanes 2–4). Consistent with the results of Mung Bean Nuclease cleavage, S1 Nuclease cleavage on the (CAG)<sub>7</sub>-8-oxoG and (CAG)<sub>7</sub>-THF hairpin substrates resulted in products with 38, 39, 41, 42, 43 and 44 nt (Figure 2.2 and 2.4, lanes 2–4). The results indicate the formation of a small hairpin that contained a single-strand (CAG)<sub>3</sub> loop with an 8-oxoG or THF and a (CAG)<sub>4</sub> stem (Figures 2.1 and 2.3) in the duplex of the substrates. Mung Bean Nuclease cleavage on the (CAG)<sub>14</sub>-8-oxoG substrate mainly resulted in products with 45, 49, 52, 55 and 58 nt (Figure 2.1, lanes 7–9). For the (CAG)<sub>14</sub>-THF hairpin substrate, the nuclease cleavage led to production of products with 45, 48 and 52 nt (Figure 2.3, lanes 7–9), whereas S1 Nuclease cleavage on the (CAG)<sub>14</sub>-8-oxoG and (CAG)<sub>14</sub>-THF hairpin substrates resulted in products with 45, 49, 51 and 54 nt (Figure 2.2 and 2.4, lanes 7–9). This indicates that the substrates contained a hairpin with a (CAG)<sub>4</sub> loop with an 8-oxoG or THF residue and 10 CAG repeats in the stem region. Thus, our results demonstrate that the 8-oxoG and THF residue (abasic site) were located in the loop region rather than in the stem of the hairpin substrates. This is consistent with a finding showing that a DNA base lesion such as an 8-oxoG is preferentially located in the loop region of a hairpin (203). We then determined the activities of OGG1 and APE1 on these substrates and discovered that 70% of 8-oxoG in the loop region of a small (CAG)<sub>7</sub> hairpin was removed (Figure 2.5, lane 2), leaving an abasic site that was efficiently cleaved by APE1 (Figure 2.5, lane 3; Figure 2, lane 2). About 40% of 8-oxoG in the loop region of a large (CAG)<sub>14</sub> hairpin was removed (Figure 2.5, lane 15), and the resulting abasic site was efficiently incised by APE1 (Figure 2.5,

lane 16; Figure 2.6, lane 13). These results indicate that OGG1 and APE1 can remove an 8-oxoG and abasic site in the loop region of a hairpin located in a CAG repeat tract, initiating BER. This further indicates that OGG1 and APE1 cleavage activity on an 8-oxoG and abasic site is sufficient to initiate BER in a CAG hairpin loop region, although their activities decrease as the size of the hairpin increases.



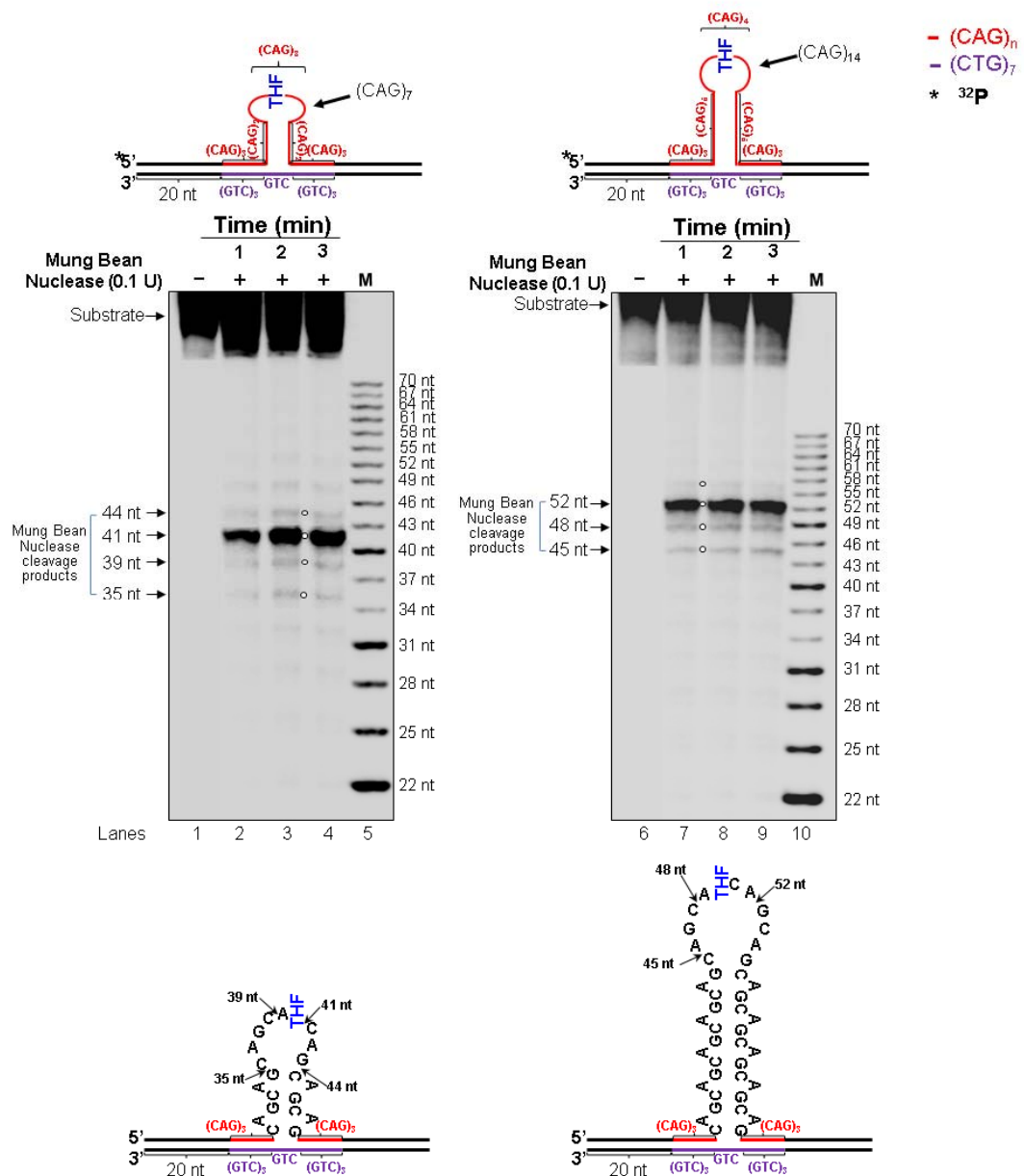
**Figure 2.1 Hairpin probing of (CAG)<sub>7</sub>- and (CAG)<sub>14</sub>-8-oxoG hairpin substrates by Mung Bean Nuclease.**

The damaged strand of the (CAG)<sub>7</sub>- or (CAG)<sub>14</sub>-8-oxoG hairpin substrate was radiolabeled at the 3'-end. Substrates were incubated with 0.15 U of Mung Bean Nuclease at 1-, 2- and 3-min time intervals (lanes 2-4, and lanes 7-9). Lanes 1 and 6 represent the undigested substrate. Lanes 5 and 10 represent 3'-radiolabeled synthesized size markers (M) with 29 nt, 32 nt, 35 nt, 38 nt, 41 nt, 44 nt, 47 nt, 50 nt, 53 nt, 56 nt, 62 nt, 68 nt and 74 nt, respectively. For all the experiments, 200 nM of substrate was used. Arrows and white dots indicate the major digestion products.



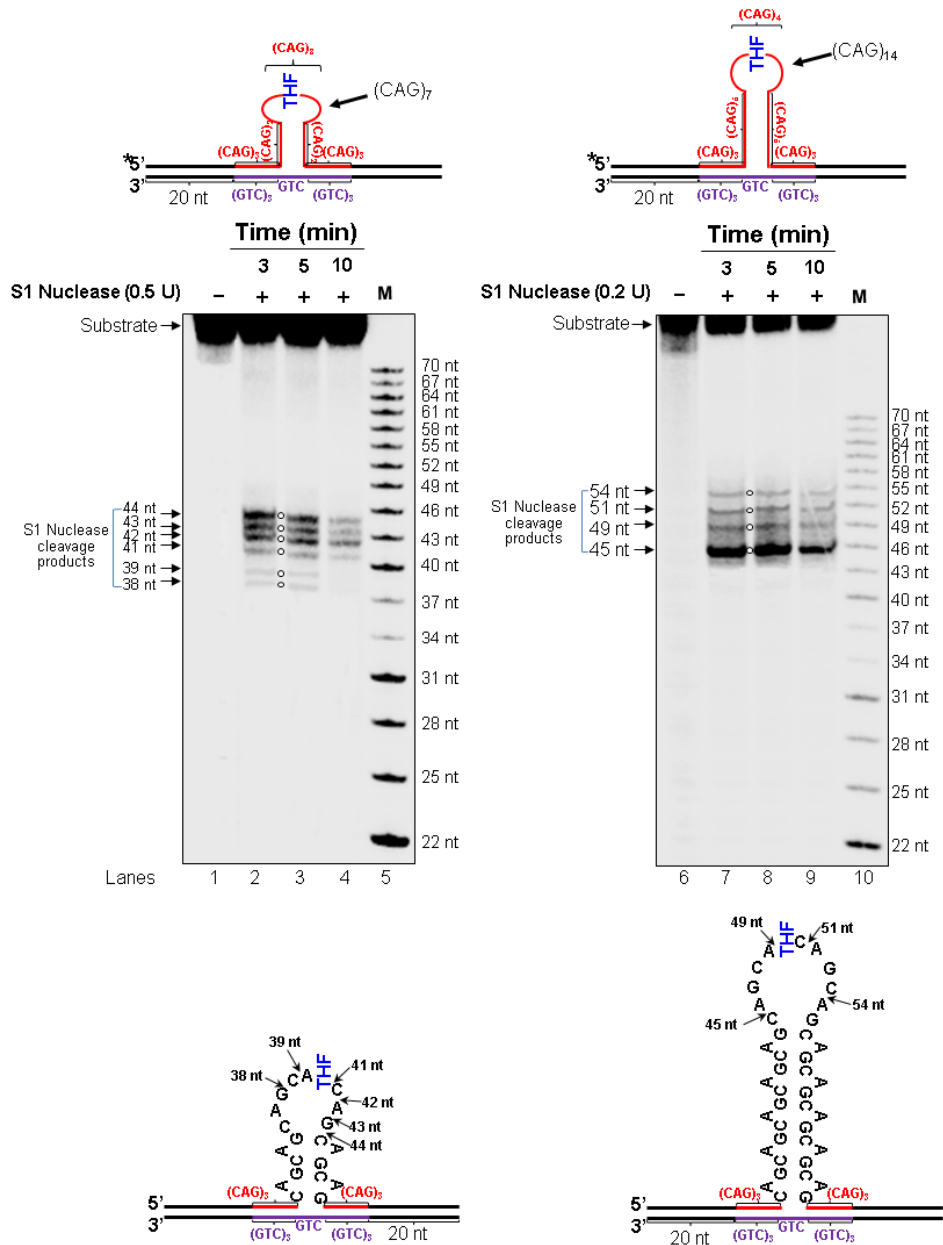
**Figure 2.2 Hairpin probing of (CAG)<sub>7</sub>- and (CAG)<sub>14</sub>-8-oxoG hairpin substrates by S1 Nuclease.**

The damaged strand of the (CAG)<sub>7</sub>- or (CAG)<sub>14</sub>-8-oxoG substrates was radiolabeled at 5'-end. Substrates were incubated with 0.2 U or 0.5 U of S1 Nuclease at 3-, 5- and 10-min time intervals (lanes 2-4, and lanes 7-9). Lanes 1 and 6 represent the undigested substrate. Lanes 5 and 10 represent 5'-radiolabeled synthesized size markers (M) with 22 nt, 25 nt, 28 nt, 31 nt, 34 nt, 37 nt, 40 nt, 43 nt, 46 nt, 49 nt, 52 nt, 55 nt, 58 nt, 61 nt, 64 nt, 67 nt, and 70 nt, respectively. For all the experiments, 200 nM of substrate was used. Arrows and white dots indicate the major digestion products.



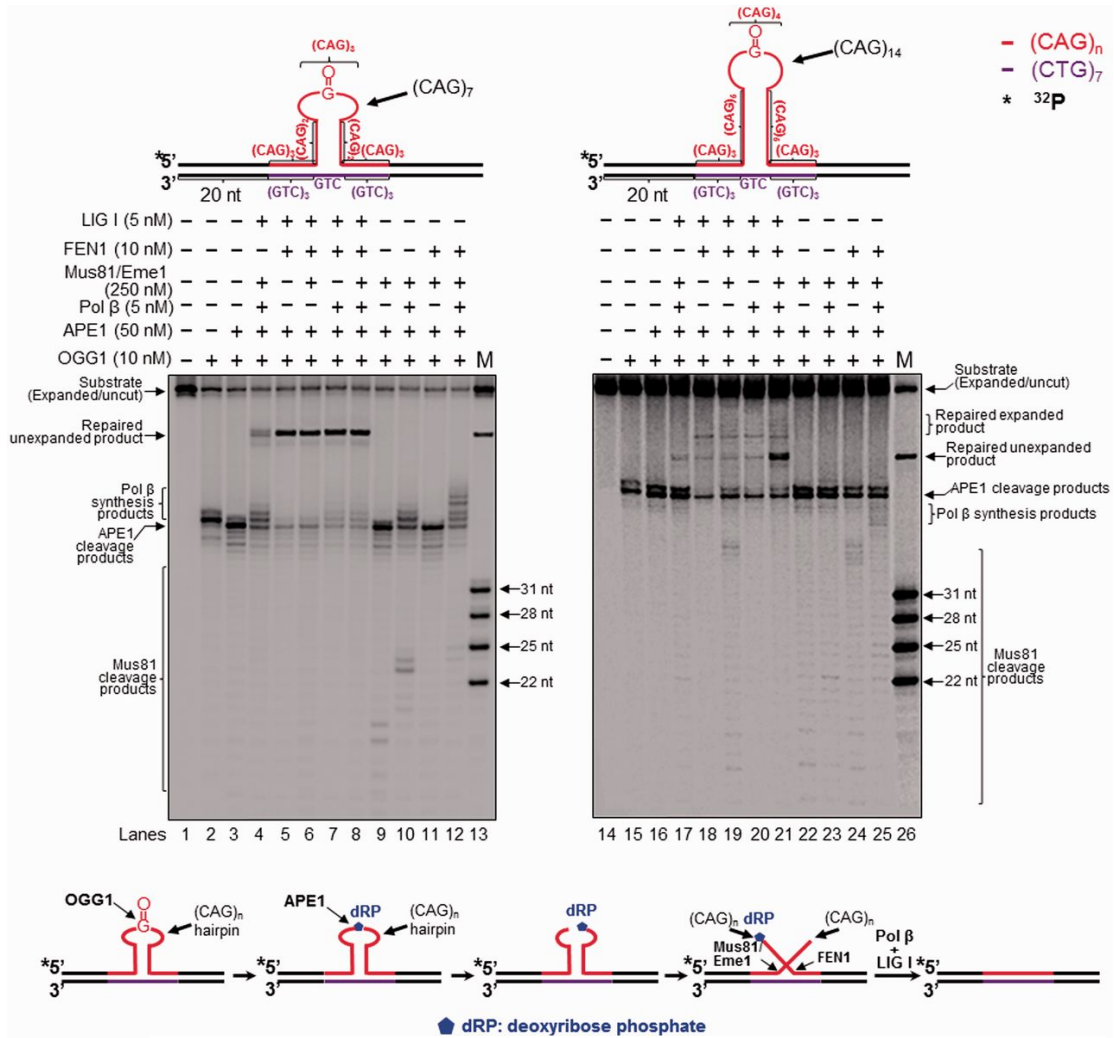
**Figure 2.3 Hairpin probing of (CAG)<sub>7</sub>- and (CAG)<sub>14</sub>-THF hairpin substrates by Mung Bean Nuclease.**

The (CAG)<sub>7</sub>- and (CAG)<sub>14</sub>-THF hairpin substrates were radiolabeled at the 5'-end of the damaged strand. Substrates (200 nM) were incubated with 0.1 U of Mung Bean Nuclease at 1-, 2- and 3-min time intervals (lanes 2-4 and lanes 7-9). Lanes 1 and 6 represent the undigested substrate. Lanes 5 and 10 represent 5'-radiolabeled synthesized size markers (M) with 22 nt, 25 nt, 28 nt, 31 nt, 34 nt, 37 nt, 40 nt, 43 nt, 46 nt, 49 nt, 52 nt, 55 nt, 58 nt, 61 nt, 64 nt, 67 nt, and 70 nt, respectively. For all the experiments, 200 nM of substrate was used. Arrows and white dots indicate the major digestion products.



**Figure 2.4 Hairpin probing of (CAG)<sub>7</sub>- and (CAG)<sub>14</sub>-THF hairpin substrates by S1 Nuclease.**

5'- radiolabeled substrates were incubated with 0.2 U or 0.5 U of S1 Nuclease at 3-, 5- and 10-min time intervals (lanes 2-4, and lanes 7-9). Lanes 1 and 6 represent the undigested substrate. Lanes 5 and 10 represent 5'-radiolabeled synthesized size markers (M) with 22 nt, 25 nt, 28 nt, 31 nt, 34 nt, 37 nt, 40 nt, 43 nt, 46 nt, 49 nt, 52 nt, 55 nt, 58 nt, 61 nt, 64 nt, 67 nt, and 70 nt, respectively. Arrows and white dots indicate the major nuclease digestion products.



**Figure 2.5 Attenuation of CAG repeat expansion through repair of 8-oxoG in the hairpin loop and Mus81/Eme1 cleavage.**

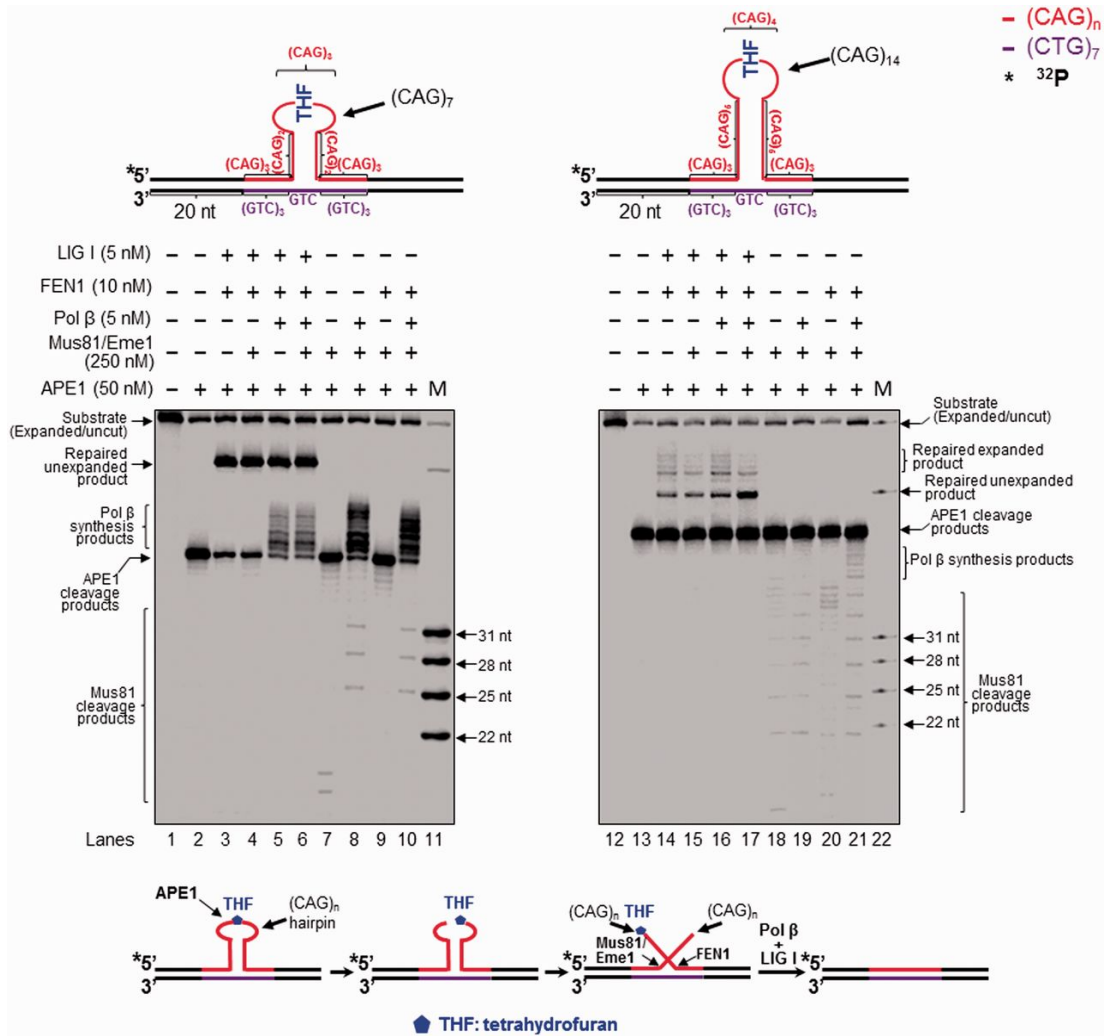
Lanes 1 and 14 correspond to substrates only. Lanes 2 and 15 correspond to reaction mixtures with 10 nM OGG1. Lanes 3 and 16 correspond to reaction mixtures with 10 nM OGG1 and 50 nM APE1. Lanes 4 and 17 correspond to reaction mixtures with 250 nM Mus81/Eme1, 5 nM pol  $\beta$ , 10 nM OGG1, 50 nM APE1 and 5 nM LIG I in the absence of FEN1. Lanes 5–6 and lanes 18–19 correspond to reaction mixtures with 10 nM OGG1, 50 nM APE1, 10 nM FEN1 and 5 nM LIG I in the absence or presence of 250 nM Mus81/Eme1. Lanes 7–8 and lanes 20–21 correspond to reaction mixtures with 10 nM OGG1, 50 nM APE1, 5 nM pol  $\beta$ , 10 nM FEN1, 5 nM LIG I in the absence or presence of 250 nM Mus81/Eme1. Lanes 9–10 and 22–23 correspond to reaction mixtures with 10 nM OGG1, 50 nM APE1, 250 nM Mus81/Eme1 in the absence and presence of 5 nM pol  $\beta$ . Lanes 11–12 and lanes 24–25 represent reaction mixtures with 10 nM OGG1, 50 nM APE1, 250 nM Mus81/Eme1 and 5 nM FEN1 in the absence and presence of pol  $\beta$ . Lanes 13 and 26 correspond to a series of synthesized size markers (M) for illustrating the size of repaired products and Mus81/Eme1 cleavage products.

## **BER of a base lesion in the loop region of a CAG repeat hairpin resulted in removal of the hairpin**

To determine if removal of an 8-oxoG and an abasic lesion located at a CAG repeat hairpin may result in any repaired product, we reconstituted BER for removing a base lesion in the loop region of a (CAG)<sub>7</sub> or (CAG)<sub>14</sub> hairpin substrate. Surprisingly, we found that BER in the small (CAG)<sub>7</sub> hairpin loop generated a product with the same length as the template strand. We designated this product as ‘unexpanded product’ (Figure 2.5, lanes 4–8; Figure 2.6, lanes 3–6). BER in the large (CAG)<sub>14</sub> hairpin loop resulted in the unexpanded product as well as the products that were shorter than the hairpin-containing substrate strand, but longer than the template strand. We named these products ‘expanded products’ (Figure 2.5, lanes 18–21; Figure 2.6, lanes 14–17). The results indicate that BER of a base lesion in the loop region of a CAG repeat hairpin removed the entire small (CAG)<sub>7</sub> and large (CAG)<sub>14</sub> hairpin, but also removed a part of a large (CAG)<sub>14</sub> hairpin. Interestingly, we found that production of the repair products required the presence of either a 3’-5’ endonuclease such as Mus81/Eme1, which was used to demonstrate the role of a 3’-flap endonuclease in removing a TNR hairpin, or FEN1 (Figure 2.5, lanes 4–5 and lanes 17–18) in BER reactions because repair reactions without flap endonucleases failed to convert the APE1 product into the repaired product even when base lesion repair was forced to short-patch BER (Figure 2.7, lanes 5 and 10). This suggests that APE1 resulted in a 5’-incised hairpin intermediate and flap intermediates that subsequently prevented DNA synthesis by pol β and DNA ligation (Figure 2.7, the scheme below the gels). Because BER enzymes cannot directly remove the APE1 incised hairpin, this further suggests that BER in the context of a hairpin loop

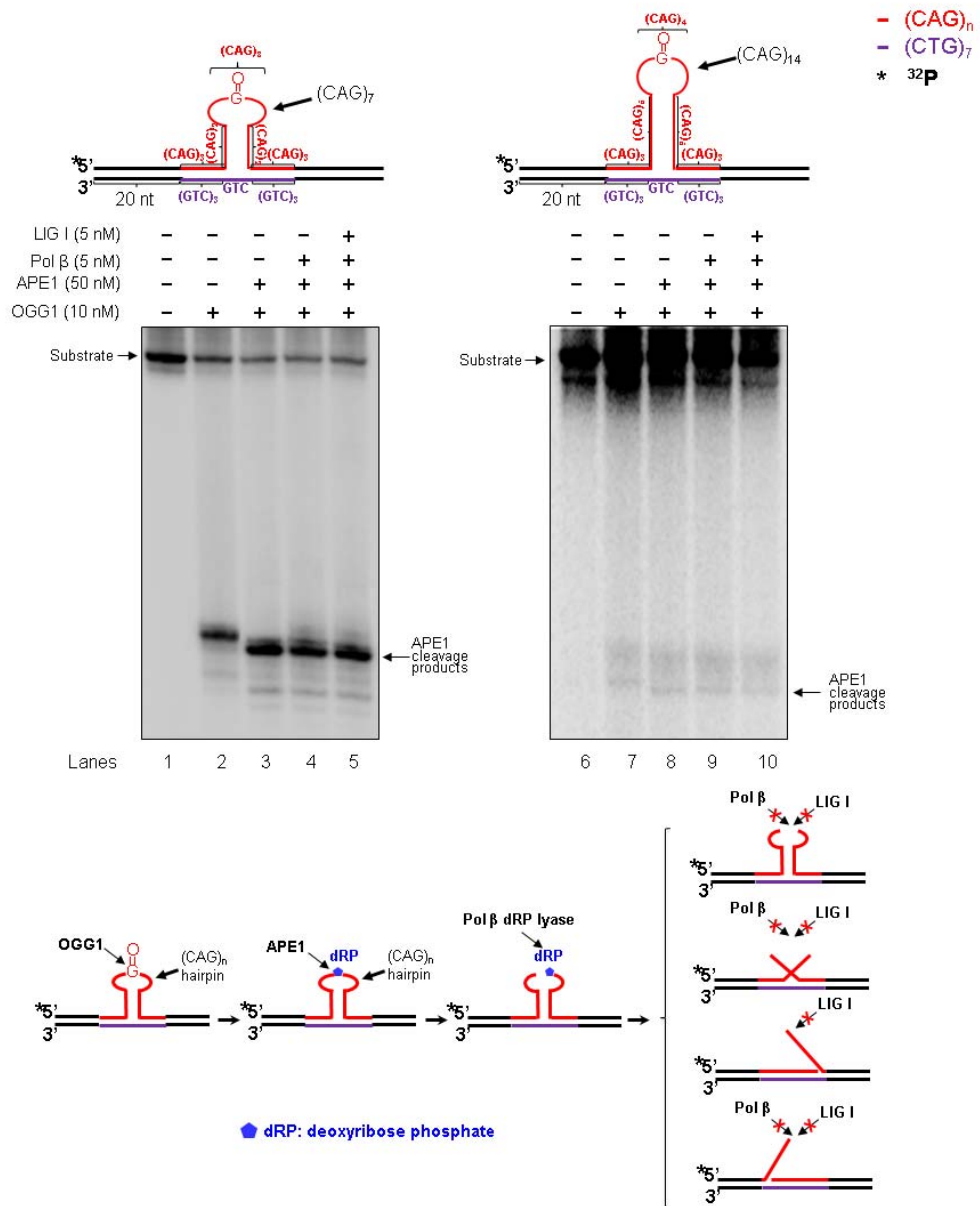


is accomplished through conversion of a hairpin stem into a double-flap intermediate. Thus, we suggest that incision of a hairpin loop by APE1 converted some of the hairpin stems into double-flap intermediates that can be subsequently processed by Mus81/Eme1 and FEN1. This further leads to removal of the hairpin and ultimately production of the repaired unexpanded product and shortened expansion products. Furthermore, we found that for 8-oxoG-containing hairpin substrates, the unexpanded product was generated in the presence of pol  $\beta$  and Mus81/Eme1 without the need of FEN1 (Figure 2.5, lanes 4 and 17). For THF-containing substrates, FEN1 was required for production of all repaired products (Figure 2.6, lanes 3–6, and lanes 14–17). This is because a THF residue that mimics an oxidized sugar cannot be removed through  $\beta$ -elimination by pol  $\beta$  dRP lyase via short-patch BER. Thus it has to be removed by long-patch BER through FEN1 cleavage of the nucleotide attached to the residue (203). Our results demonstrate that both a 3'-5' endonuclease and FEN1 play an important role in removing a hairpin during BER of a base lesion in a hairpin loop presumably by cleaving a 3'-flap or a 5'-flap.



**Figure 2.6 Attenuation of CAG repeat expansion through repair of an abasic site in the hairpin loop and Mus81/Eme1 cleavage.**

Repair of an abasic site in the loop region of a CAG hairpin and Mus81/Eme1 cleavage were examined with substrates containing a (CAG)<sub>7</sub> (left panel) or (CAG)<sub>14</sub> hairpin (right panel) with a THF in the hairpin loop region. Lanes 1 and 12 correspond to substrates only. Lanes 2 and 13 represent reaction mixtures with 50 nM APE1. Lanes 3, 5, 14 and 16 represent reaction mixtures with or without 5 nM pol β in the presence of 50 nM APE1, 10 nM FEN1 and 5 nM LIG I. Lanes 4, 6, 15 and 17 correspond to reaction mixtures with or without 5 nM pol β in the presence of 50 nM APE1, 5 nM FEN1, 250 nM Mus81/Eme1 and 5 nM LIG I. Lanes 7–8 and 18–19 correspond to reaction mixtures with or without 5 nM pol β in the presence of 250 nM Mus81/Eme1 and 50 nM APE1. Lanes 9–10 and 20–21 represent reaction mixtures with or without 5 nM pol β in the presence of 10 nM FEN1, 250 nM Mus81 and 50 nM APE1. Lanes 11 and 22 represent synthesized size markers (M). Substrates were <sup>32</sup>P-labeled at the 5'-end of their damaged strands and are illustrated schematically above the gels.



**Figure 2.7 Short-patch BER of (CAG)<sub>7</sub>- and (CAG)<sub>14</sub>-8-oxoG hairpin containing substrates.**

Short-patch BER of an 8-oxoG in the loop region of a CAG hairpin was examined by incubating the (CAG)<sub>7</sub>- and (CAG)<sub>14</sub>-8-oxoG substrates (25 nM) with OGG1 (10 nM) alone (lanes 2 and 7), or OGG1 (10 nM) along with APE1 (50 nM) (lanes 3 and 8) or OGG1 (10 nM) along with both APE1 (50 nM) and pol β (5 nM) (lanes 4 and 9) or OGG1 (10 nM) along with APE1 (50 nM), pol β (5 nM) and LIG (5 nM) (lanes 5 and 10) at 37 °C for 30 min. Lanes 1 and 6 correspond to substrates only. Substrates were <sup>32</sup>P-labeled at the 5'-end of their damaged strands and are illustrated schematically above the gels. The scheme of short-patch BER of an 8-oxoG in the hairpin substrates is illustrated below the gels.

**The 3'-5' endonuclease Mus81/Eme1 removed the entirety of a CAG repeat hairpin by cleaving an upstream 3'-flap during BER, preventing CAG repeat expansion**

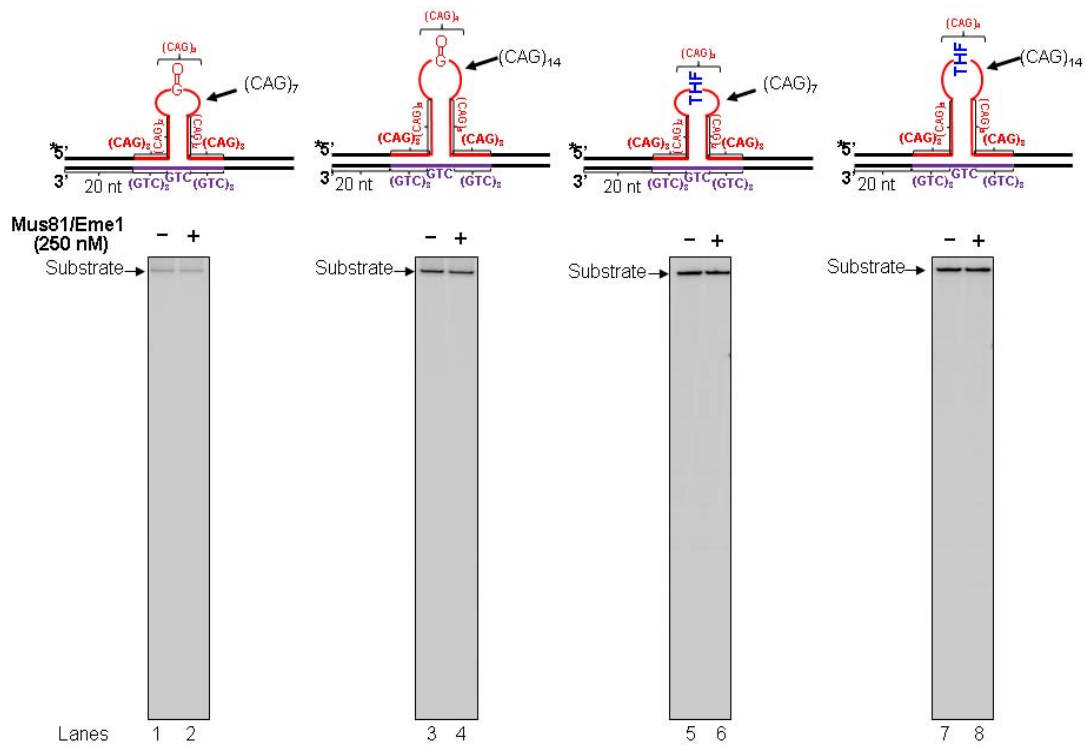
Because removal of an 8-oxoG and 5'-incision of an abasic lesion in a hairpin loop may further convert the hairpin into an ssDNA break intermediate with an upstream 3'-flap and a downstream 5'-flap, cleavage of the flaps by a 3'-5' flap endonuclease such as Mus81/Eme1 and 5'-flap endonuclease FEN1 may remove the hairpin and produce the unexpanded product. To test this possibility, we initially characterized the activity of the 3'-5' flap endonuclease Mus81/Eme1 (206-211) in the context of a hairpin loop during BER. We found that Mus81/Eme1 endonucleolytically cleaved the (CAG)<sub>7</sub> (Figure 2.5, lanes 9–12; Figure 2.6, lanes 7–10) and (CAG)<sub>14</sub> hairpin substrates (Figure 2.5, lanes 22–25; Figure 2.6, lanes 18–21), resulting in a series of 3'-flap cleavage products. These products were not from the cleavage of the entire hairpins because no Mus81/Eme1 cleavage products were detected from a (CAG)<sub>7</sub> or (CAG)<sub>14</sub> hairpin substrate in the absence of OGG1 and APE1 (Figure 2.8, lanes 2,4,6 and 8). This indicates that BER in the (CAG)<sub>7</sub> and (CAG)<sub>14</sub> hairpin loop regions converted the hairpins into intermediates with a 3'-flap that was subsequently cleaved by Mus81/Eme1. Because the Mus81/Eme1 cleavage products from (CAG)<sub>7</sub> and (CAG)<sub>14</sub> hairpin substrates correspond to fragments containing a part of the 20-nt random sequence flanking the (CAG)<sub>7</sub> (Figure 2.5, lane 9; Figure 2.6, lane 7) and (CAG)<sub>14</sub> hairpin (Figure 2.5, lane 22; Figure 2.6, lane 18); this suggests that a downstream 5'-(CAG)<sub>4</sub> or (CAG)<sub>7</sub> repeat flap annealed to the template and displaced the upstream strand, creating a long flap that was subsequently removed by Mus81/Eme1. This further resulted in removal of the hairpins and production of the unexpanded product. To further confirm this, we determined if Mus81/Eme1 cleavage on

a double-flap substrate with an upstream 3'-(CAG)<sub>4</sub> - and a downstream 5'-(CAG)<sub>3</sub>-flap or a substrate with a 3'-(CAG)<sub>7</sub>-flap and 5'-(CAG)<sub>7</sub>-flap, could also result in the unexpanded product. The substrates mimic double-flap intermediates converted from APE1-incised hairpins with varying sizes. To verify the presence of the 3'- and 5'-flaps in the substrates, we used Mung Bean Nuclease to probe the formation of 3'- and 5'-flaps of the double-flap substrates. The results showed that the nuclease cleavage of the upstream strand of the (CAG)<sub>3</sub>/(CAG)<sub>4</sub> double-flap substrate generated products with 29–38 nt indicative of cleavage of 1–4 CAG repeats (Figure 2.9, lanes 2–4). The nuclease cleavage of the downstream strand of the substrate resulted in products with 30–36 nt indicative of cleavage of 1–3 CAG repeats (Figure 2.9, lanes 7–9). This indicates that the substrate contained an upstream (CAG)<sub>4</sub> flap and a downstream 5'-(CAG)<sub>3</sub>-THF flap (Figure 2.9, the flap substrates illustrated at the bottom). Mung Bean Nuclease cleavage on the upstream strand of the (CAG)<sub>7</sub> double-flap substrate resulted in products with 29–46 nt (Figure 2.10, lanes 2–5). The nuclease cleavage on the downstream strand of the substrate generated products with 30–47 nt (Figure 2.10, lanes 8–11). The results indicate that the nuclease cleaved 1–7 CAG repeats in both the upstream and downstream strands demonstrating the formation of a (CAG)<sub>7</sub>-flap in both of the strands (Figure 2.10, the flap substrates illustrated at the bottom). We found that the Mus81/Eme1 cleavage pattern on the double-flap substrates was similar to that on the (CAG)<sub>7</sub> and (CAG)<sub>14</sub> hairpin substrates (Figure 2.11, lanes 6–7 and lanes 16–17), indicating that the hairpins were indeed converted to the double-flap intermediates that were processed by Mus81/Eme1 during BER.

Interestingly, we found that for the large (CAG)<sub>14</sub> hairpin and its corresponding double-flap intermediate, Mus81/Eme1 cleavage in the presence of FEN1 mainly resulted in relatively larger products compared with its cleavage in the absence of FEN1 (Figure 2.5, compare lane 24 with lane 22; Figure 2.6, compare lane 20 with lane 18; Figure 2.11, compare lane 18 with lane 16). This indicates that Mus81/Eme1 cleaved a shorter 3'-flap in the presence of FEN1 than it did in the absence of FEN1. This suggests that removal of a 5'-flap by FEN1 allowed the upstream 3'-flap to anneal to the template, resulting in a short 3'-flap that was cleaved by Mus81/Eme1. To further determine the number of Mus81/Eme1 cleavage products, the double-flap substrates were radiolabeled at the 3'-end of the upstream strand and incubated with Mus81/Eme1 in the absence and presence of FEN1 or/and pol  $\beta$  (Figure 2.12). The results showed that in the absence of FEN1, Mus81/Eme1 flap cleavage on the (CAG)<sub>3</sub>/(CAG)<sub>4</sub> double-flap substrate resulted in two products with 21 and 22 nt, respectively (Figure 2.12, lanes 2–3), indicating that the 3'-flap endonuclease cleaved within the random sequence region that flanked the repeated sequence and removed the entirety of the 3'-repeat-containing flap. In the presence of FEN1, Mus81/Eme1 cleavage resulted in a 1-nt product (Figure 2.12, lanes 4–5). This further confirmed that FEN1 removed a 5'-(CAG)<sub>3</sub>-THF flap before Mus81/Eme1 removal of a 3'-(CAG)<sub>4</sub>-flap allowing the 3'-flap to anneal to the template to create a 1-nt 3'-flap that was subsequently removed by Mus81/Eme1. For the (CAG)<sub>7</sub> double-flap substrate, Mus81/Eme1 cleavage resulted in multiple products with 9–31 nt (Figure 2.12, lanes 8–11), indicating that the 3'-(CAG)<sub>7</sub> flap folded into a series of intermediates with a small hairpin attached to a short flap that was captured and cleaved by Mus81/Eme1. In the presence of FEN1, Mus81/Eme1 cleavage mainly resulted in products with 9–10 nt

(Figure 2.12, lanes 10–11). Again, this confirmed that FEN1 removed the downstream 5'-CAG repeat flap allowing the upstream flap to anneal to the template strand creating a short flap with 9 or 10 nt that was then cleaved by the 3'-endonuclease. In conclusion, our results demonstrate that BER in a CAG repeat hairpin loop can convert the hairpin into a double-flap intermediate with a 3'-flap and 5'-flap, which can be cleaved by a 3'-5' and 5'-3' flap endonuclease, such as Mus81/Eme1 and FEN1, in a cooperative manner

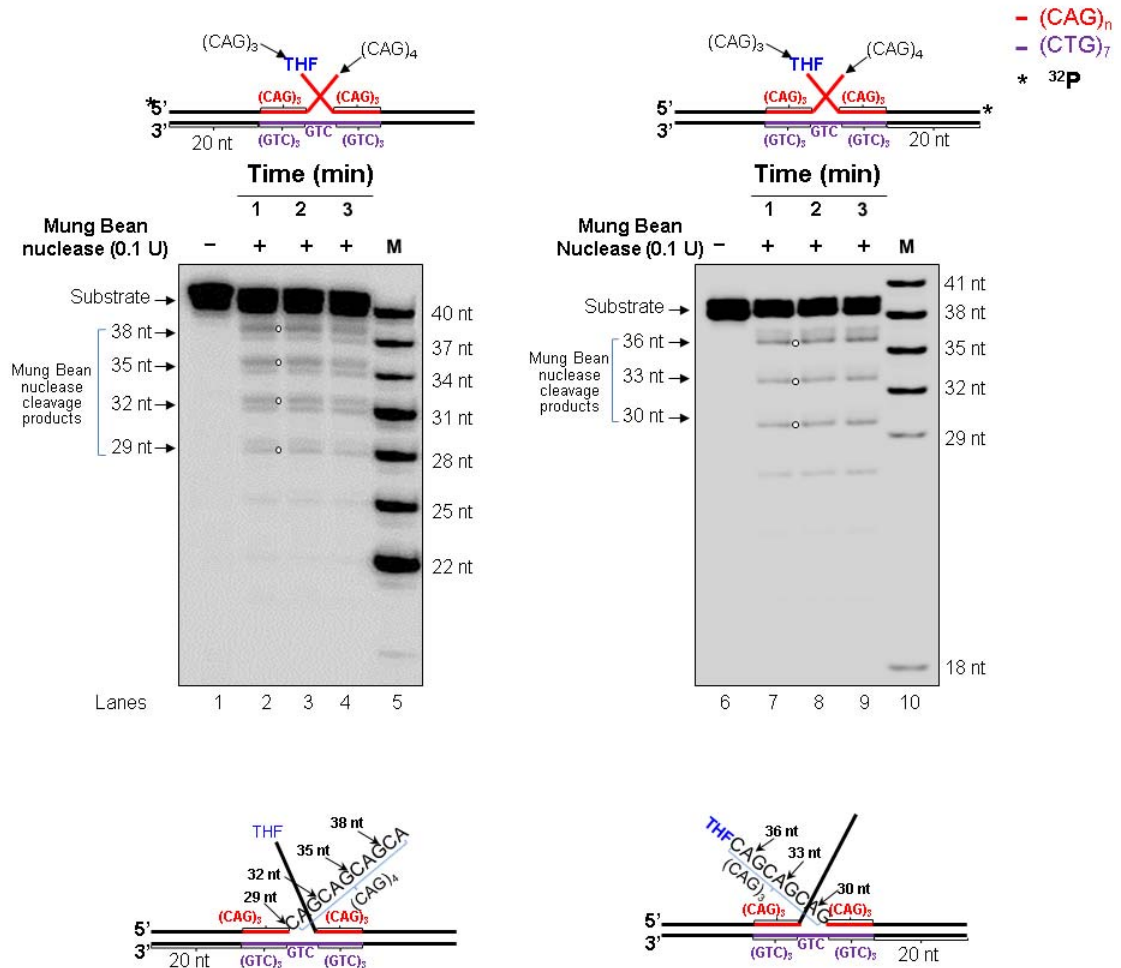
\* 32P



**Figure 2.8 Mus81/Eme1 cleavage of a hairpin.**

The (CAG)<sub>7-8</sub>-oxoG/THF and (CAG)<sub>14</sub>-8-oxoG/THF substrates were radiolabeled at the 5'-end of the hairpin containing strand. Substrates (25 nM) were incubated with 250 nM purified Mus81/Eme1 at 37°C for 30 min (lanes 2, 4, 6 and 8). Lanes 1, 3, 5 and 7 represent substrate alone. Substrates are illustrated schematically above the gels.

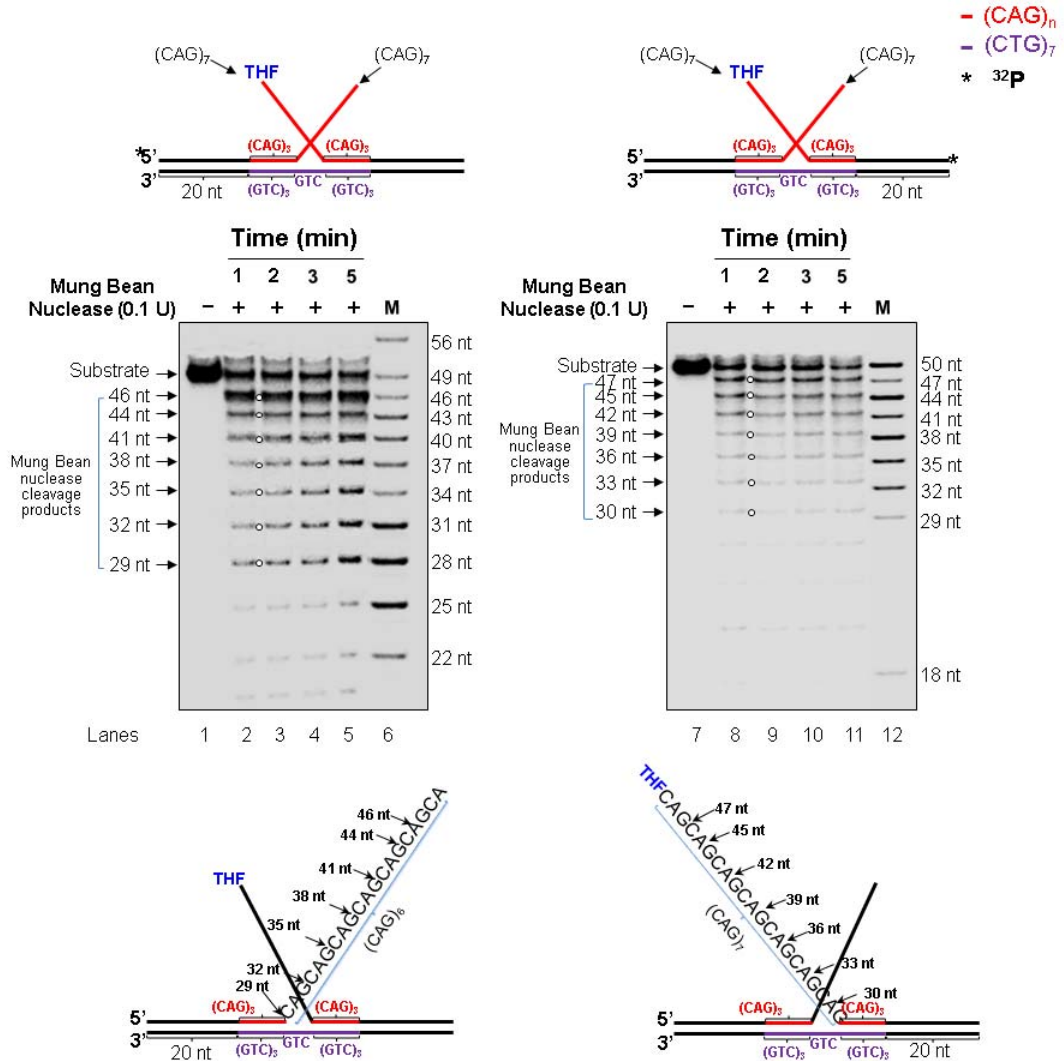
resulting in removal of the hairpin and prevention of repeat expansion.



**Figure 2.9 Flap probing of (CAG)<sub>3</sub>/(CAG)<sub>4</sub> double-flap substrate.**

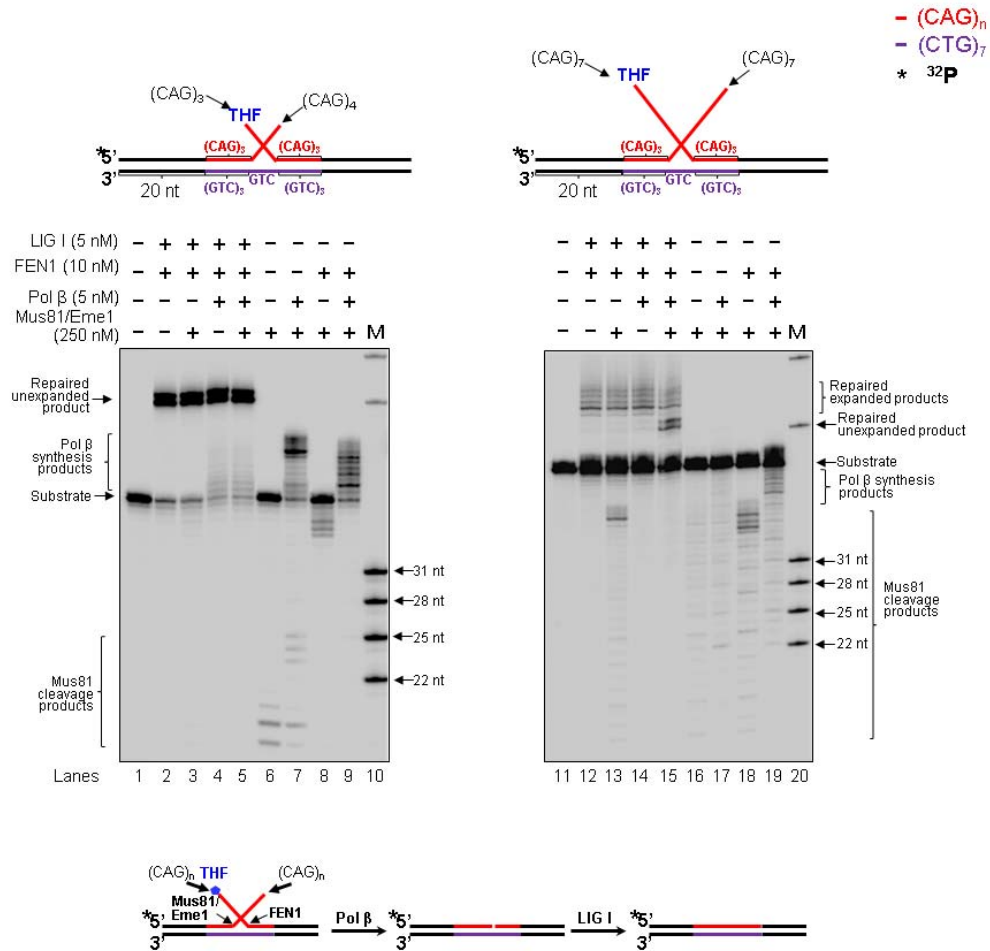
A substrate with a 3'-(CAG)<sub>4</sub> flap and a 5'-(CAG)<sub>3</sub>-THF flap was radiolabeled at the 5'-end (left panel) of the upstream strand or the 3'-end (right panel) of the downstream strand for probing the formation of an upstream 3'-flap and a downstream 5'-flap. The substrate (200 nM) was incubated with 0.1 U of Mung Bean Nuclease at 1-, 2- and 3-min time intervals (lanes 2-4 and lanes 7-9). Lanes 1 and 6 represent an undigested substrate. Lanes 5 and 10 represent 5'-radiolabeled synthesized size markers (M) with 22 nt, 25 nt, 28 nt, 31 nt, 34 nt, 37 nt, and 40 nt, respectively (left panel) and 3'-radiolabeled size markers (M) with 18 nt, 29 nt, 32 nt, 35 nt, 38 nt, and 41 nt, respectively. Arrows and white dots indicate the major nuclease digestion products. The substrates are illustrated schematically above the gels. Double-flaps deduced by a specific nuclease cleavage pattern and the nuclease digestion sites are illustrated schematically below the gels.





**Figure 2.10 Flap probing of (CAG)<sub>7</sub> double-flap substrate.**

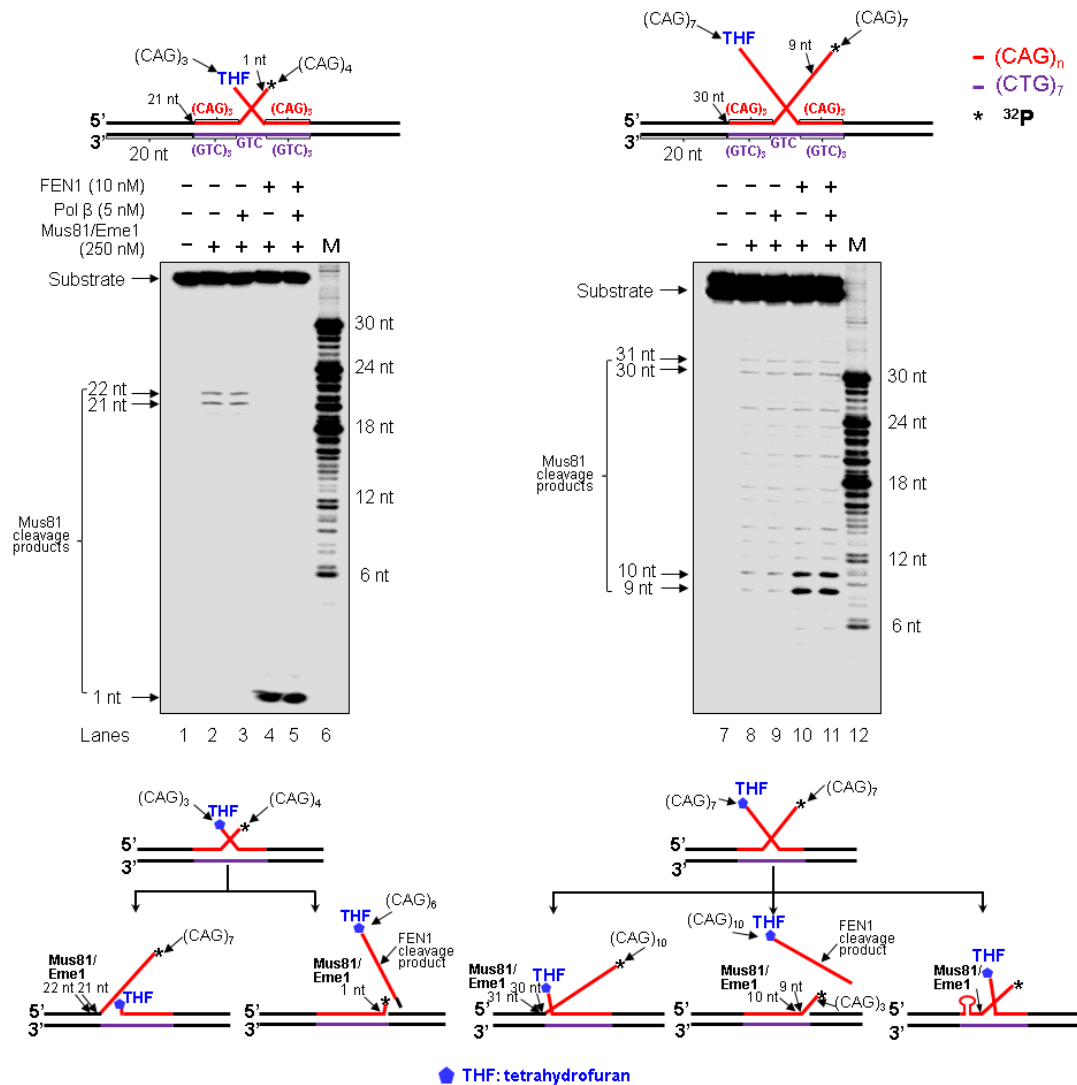
A substrate with a 3'-(CAG)<sub>7</sub>- and 5'-(CAG)<sub>7</sub>-THF flap was radiolabeled at the 5'-end of the upstream, strand (left panel) or the 3'-end of the downstream strand (right panel). The substrate (200 nM) was incubated with 0.1 U of Mung Bean Nuclease at 1-, 2-, 3-, and 5-min time intervals (lanes 2-5 and lanes 8-11). Lanes 1 and 7 represent an undigested substrate. Lane 6 represents 5'-radiolabeled synthesized size markers (M) with 22 nt, 25 nt, 28 nt, 31 nt, 34 nt, 37 nt, 40 nt, 43 nt, 46 nt, 49 nt and 56 nt for probing the formation of a 3'-flap. Lane 12 represents 3'-radiolabeled size markers (M) with 18 nt, 29 nt, 32 nt, 35 nt, 38 nt, 41 nt, 44 nt, 47 nt and 50 nt for probing the formation of a 5'-flap. Arrows and white dots indicate the major nuclease digestion products. Substrates are illustrated schematically above the gels. Double-flaps deduced by a specific nuclease cleavage pattern and the nuclease digestion sites are illustrated schematically below the gels.



◆ THF: tetrahydrofuran

### Figure 2.11 Processing of a CAG repeat double-flap intermediate during BER prevented or attenuated CAG repeat expansion.

Processing of a CAG repeat double-flap substrate and its effects on repeat stability during BER was examined with substrates containing an upstream 3'-(CAG)<sub>4</sub> and a downstream 5'-THF-(CAG)<sub>3</sub> (left panel) or an upstream 3'-(CAG)<sub>7</sub> flap and a downstream 5'-THF-(CAG)<sub>7</sub> flap (right panel). Lanes 1 and 11 represent substrates only. Lanes 2, 4, 12 and 14 correspond to reaction mixtures with or without 5 nM pol β in the presence of 10 nM FEN1 and 5 nM LIG I. Lanes 3, 5, 13 and 15 correspond to reaction mixtures with or without 5 nM pol β in the presence of 250 nM Mus81/Eme1, 10 nM FEN1 and 5 nM LIG I. Lanes 6–7 and 16–17 represent reaction mixtures with or without 5 nM pol β in the presence of 250 nM Mus81/Eme1. Lanes 8–9 and 18–19 correspond to reaction mixtures with or without 5 nM pol β in the presence of 250 nM Mus81/Eme1 and 10 nM FEN1. Lanes 10 and 20 correspond to synthesized size markers (M). Substrates were <sup>32</sup>P-labeled at the 5'-end of their upstream strands.



**Figure 2.12 Mus81/Eme1 cleavage of the upstream flap of a CAG repeat double-flap intermediate.**

Mus81/Eme1 cleavage of the upstream flap of a CAG repeat double-flap intermediate was examined using a double-flap substrate with an upstream 3'-(CAG)<sub>4</sub> and a downstream 5'-(CAG)<sub>3</sub>-THF flap (left panel) or a substrate with an upstream 3'-(CAG)<sub>7</sub> flap and a downstream 5'-(CAG)<sub>7</sub>-THF flap (right panel). The substrates were radiolabeled at the 3'-end of the upstream strand. Lanes 1 and 7 represent a substrate alone. Lanes 2 and 8 represent reaction mixtures with 250 nM Mus81/Eme1. Lanes 3 and 9 correspond to reaction mixtures with 250 nM Mus81/Eme1 and 5 nM pol β. Lanes 4 and 10 represent reaction mixtures with 250 nM Mus81/Eme1 and 10 nM FEN1. Lanes 5 and 11 correspond to reaction mixtures with 250 nM Mus81/Eme1, 5 nM pol β and 10 nM FEN1. Lanes 6 and 12 represent synthesized size markers (M). The substrates are illustrated schematically above the gels. A scheme that indicates Mus81/Eme1 cleavage of the upstream 3'-flap of the double-flap substrates is illustrated below the gels.

## **FEN1 processed a CAG repeat hairpin by removing a 5'-CAG repeat flap resulting in attenuation of repeat expansion during BER**

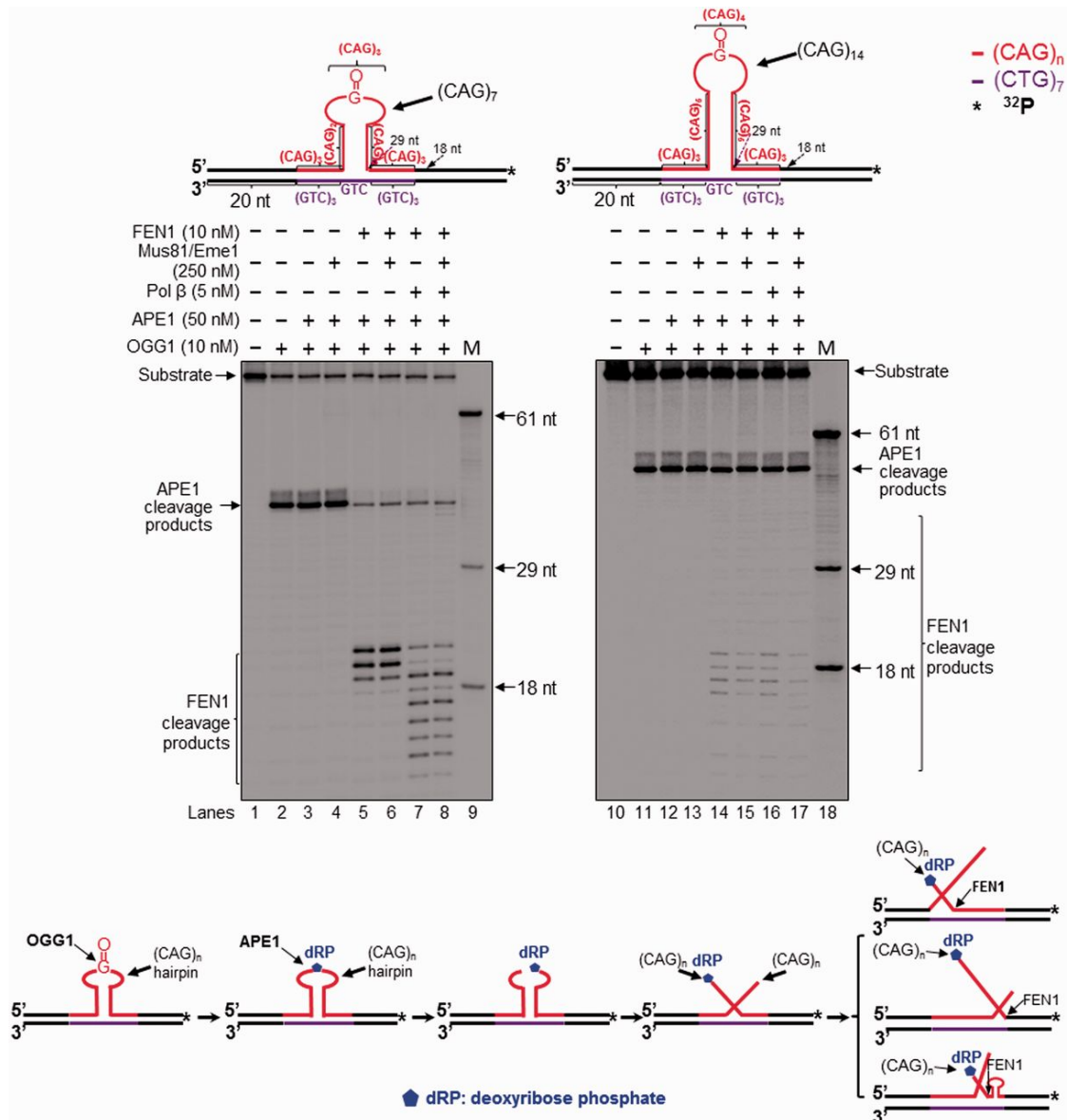
FEN1 plays a dual role in modulating TNR stability by processing a 5'-TNR flap/hairpin (115,166,192,199). Coordination between FEN1 and pol  $\beta$  leads to both TNR expansion and deletion during BER (199,200). It is possible that FEN1 may facilitate removal of a hairpin by processing a 5'-CAG repeat flap during BER in a hairpin loop, preventing CAG repeat expansion. To test this possibility, we examined FEN1 cleavage during BER in a CAG repeat hairpin loop and a double-flap intermediate. We found that FEN1 cleavage on the (CAG)<sub>7</sub> hairpin substrate resulted in the unexpanded product (Figure 2.5, lanes 5 and 7; Figure 2.6, lanes 3 and 5), indicating that the enzyme removed the entire hairpin. Interestingly, FEN1 cleavage on the (CAG)<sub>14</sub> hairpin resulted in both the unexpanded product and expanded products that were shorter than the expanded/uncut substrate strand (Figure 2.5, lanes 18–21; Figure 2.6, lanes 14–17). This indicates that FEN1 cleavage completely removed a small and large CAG repeat hairpin, but also removed a part of a large hairpin during BER.

To further determine how FEN1 can remove a hairpin by flap cleavage on the hairpin substrates, we examined FEN1 cleavage using the substrates that were radiolabeled at the 3'-end of the hairpin-containing expanded/uncut strand. The results showed that in the absence of pol  $\beta$ , FEN1 efficiently cleaved a 5'-(CAG)<sub>6</sub> flap from the (CAG)<sub>7</sub> hairpin substrate (Figure 2.13, lane 5; Figure 2.14, lane 4), but removed a 5'-(CAG)<sub>10</sub> flap and other shorter 5'-flaps from (CAG)<sub>14</sub> hairpin substrates with a low efficiency (Figure 2.13, lane 14; Figure 2.14, lane 12). This suggests that BER in the loop of a small hairpin generated a short 5'-flap that was efficiently cleaved by FEN1 resulting

in removal of the entirety of a small hairpin. On the other hand, BER in a large hairpin generated a long 5'-flap that inhibited FEN1 cleavage. This forced FEN1 to use its alternate flap cleavage activity to capture and cleave a short 5'-flap, thereby resulting in partial removal of a large hairpin. This was further supported by the fact that FEN1 alternate cleavage resulted in multiple products (The right panel of Figures 2.13–2.15), indicating that the enzyme captured and cleaved a series of short flaps. This suggests that a long (CAG)<sub>7</sub> flap folded into multiple small alternate hairpins attached to a short flap of varying sizes. FEN1 then loaded from the 5'-end of the short flap and tracked down to the bottom of the flaps and cleaved them. To determine if all FEN1 products resulted from its flap cleavage activity, we examined its cleavage on a (CAG)<sub>7</sub> or (CAG)<sub>14</sub> hairpin substrate in the absence of OGG1 and APE1. We failed to observe any FEN1 cleavage products from the substrates (Figure 2.16 lanes 2, 4, 6, and 8), indicating that FEN1 cannot directly cleave a hairpin, and that the products were from the flap cleavage activity of the enzyme. This further demonstrates a critical role of FEN1 flap cleavage in removing hairpin structures during BER. In the presence of pol β, FEN1 cleavage on the small (CAG)<sub>7</sub> hairpin substrate was stimulated (Figure 2.13, compare lane 5 with lane 7; Figure 2.14, compare lane 4 with lane 6), presumably by pol β strand-displacement synthesis that generated a 5'-flap. However, FEN1 cleavage on the (CAG)<sub>14</sub> hairpin substrate was not affected by pol β synthesis (Figure 2.13, compare lane 14 with lane 16; Figure 2.14, compare lane 12 with lane 14), suggesting that a long 3'-flap was formed during BER, and this inhibited pol β DNA synthesis. Consistent with this, FEN1 cleavage on the double-flap substrates that mimic flap intermediates from an incised hairpin loop resulted in the same products as those from its cleavage on the hairpin substrates (Figure

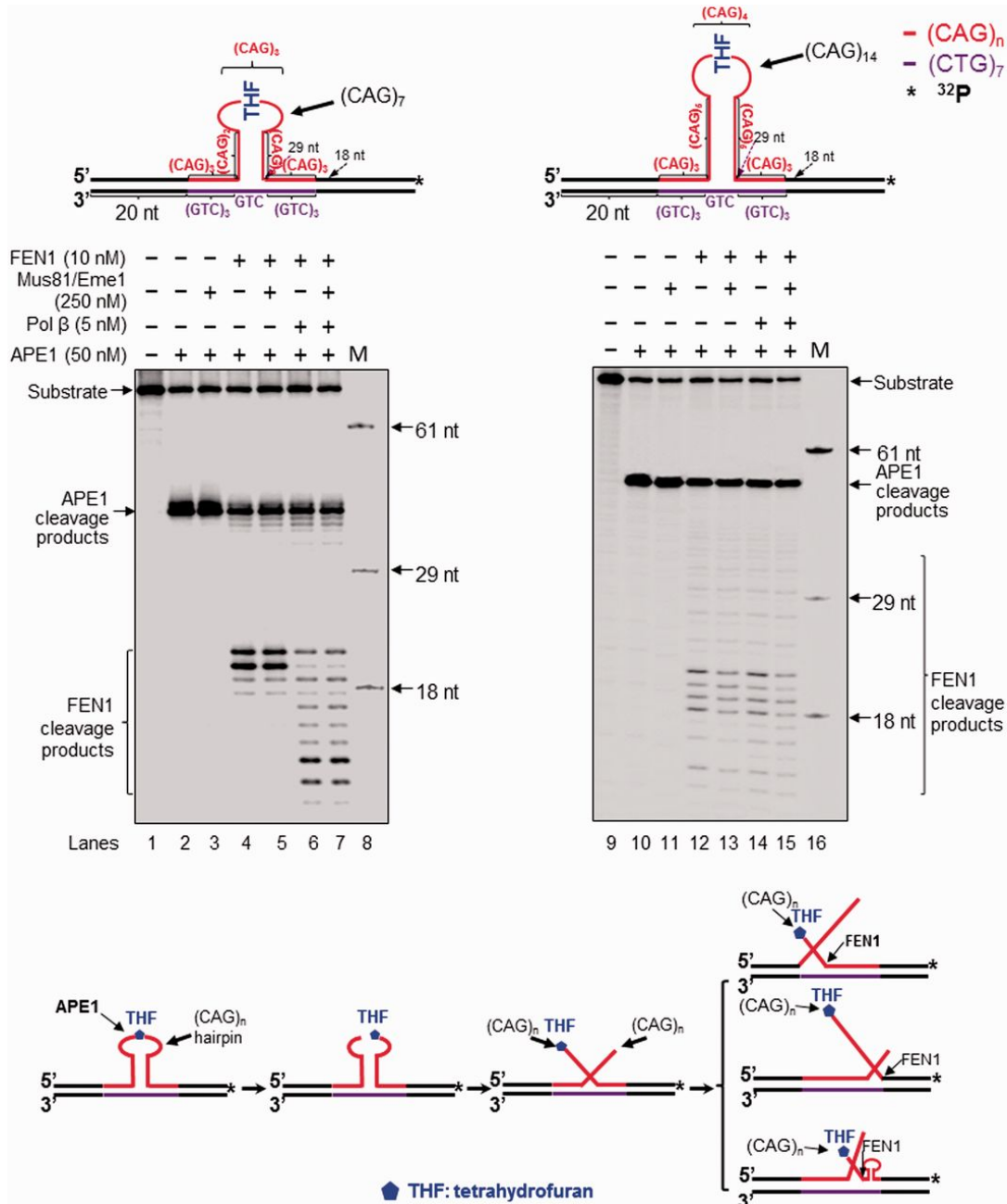
2.15, lanes 3–6 and 10–13). This confirmed that BER in a hairpin loop generated a 5'-CAG-flap that was cleaved by FEN1.

Because Mus81/Eme1 can stimulate FEN1 flap cleavage (212), we then asked if Mus81/Eme1 can stimulate FEN1 cleavage on the (CAG)<sub>7</sub> and (CAG)<sub>14</sub> hairpin substrates. We found that Mus81/Eme1 failed to stimulate FEN1 cleavage on the substrates (Figure 2.13, compare lane 7 with lane 8, lane 16 with lane 17; Figure 2.14, compare lane 6 with lane 7, lane 14 with lane 15; Figure 2.15, compare lane 5 with lane 6, lane 12 with lane 13). This suggests that Mus81/Eme1 cleavage on a 3'-flap does not affect FEN1 processing of a CAG repeat hairpin.



**Figure 2.13 FEN1 processing of a hairpin during repair of an 8-oxoG in the hairpin loop.**

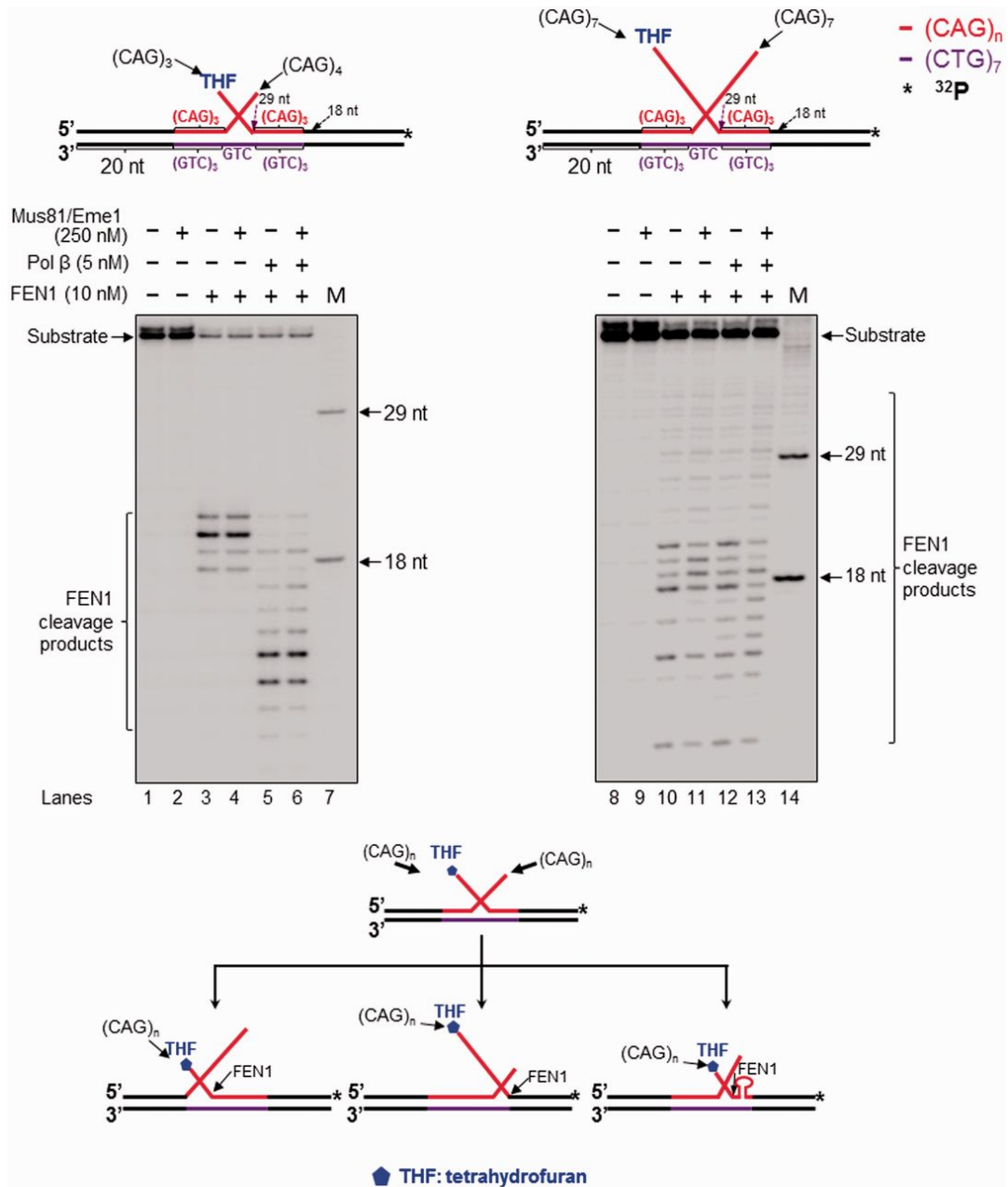
Lanes 1 and 10 indicate substrates only. Lanes 2 and 11 correspond to reaction mixtures with 10 nM OGG1. Lanes 3 and 12 represent reaction mixtures with 10 nM OGG1 and 50 nM APE1. Lanes 4 and 13 represent reaction mixtures with 10 nM OGG1, 50 nM APE1 and 250 nM Mus81/Eme1. Lanes 5–6 and 14–15 correspond to reaction mixtures with and without 250 nM Mus81/Eme1 in the presence of 10 nM OGG1, 50 nM APE1 and 10 nM FEN1. Lanes 7–8 and 16–17 represent reaction mixtures with or without 250 nM Mus81/Eme1 in the presence of 10 nM OGG1, 50 nM APE1, 10 nM FEN1 and 5 nM pol β. Lanes 9 and 18 correspond to a series of synthesized size markers (M) for illustrating the size of FEN1 cleavage products.



**Figure 2.14 FEN1 processing of a hairpin during repair of an abasic site in the hairpin loop.**

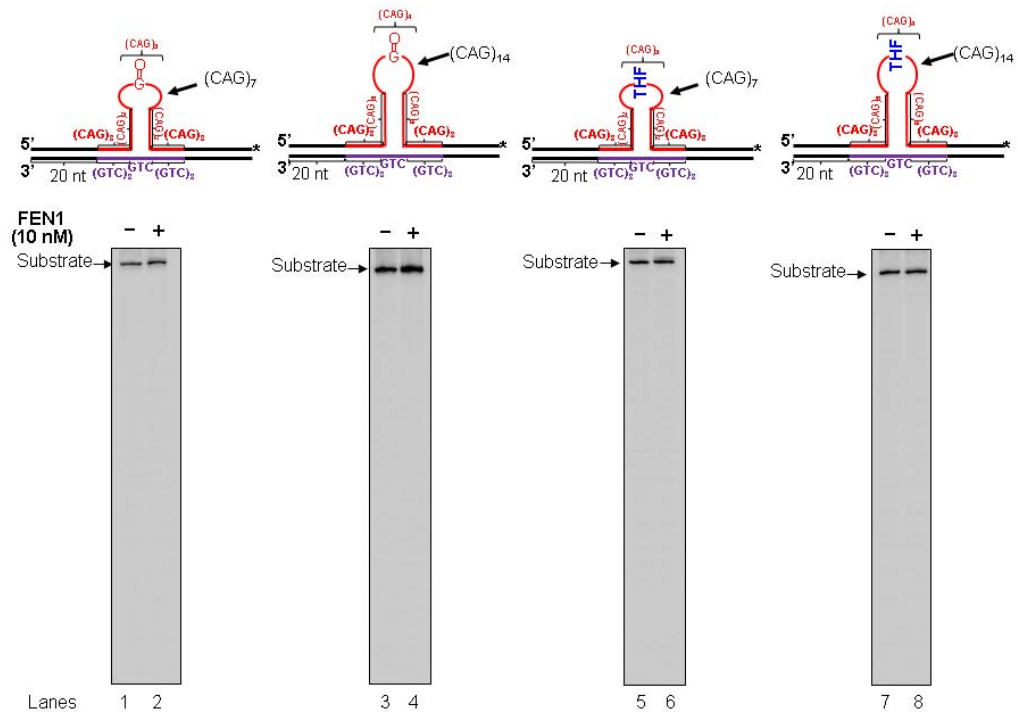
Lanes 1 and 9 represent substrates only. Lanes 2 and 10 correspond to reaction mixtures with 50 nM APE1. Lanes 3 and 11 correspond to reaction mixtures with 50 nM APE1 and 250 nM Mus81/Eme1. Lanes 4–5 and 12–13 correspond to reaction mixtures with 50 nM APE1 and 10 nM FEN1 in the absence or presence of 250 nM Mus81/Eme1. Lanes 6–7 and 14–15 correspond to reaction mixtures with or without 250 nM Mus81/Eme1 in the presence of 50 nM APE1, 10 nM FEN1 and 5 nM pol β. Lanes 8 and 16 correspond to a series of synthesized size markers (M).





**Figure 2.15 FEN1 cleavage of CAG repeat double-flap intermediates.**

Lanes 1 and 8 correspond to substrates only. Lanes 2 and 9 correspond to reaction mixtures with 250 nM Mus81/Eme1. Lanes 3–4 and 10–11 correspond to reaction mixtures with or without 250 nM Mus81/Eme1 in the presence of 10 nM FEN1. Lanes 5–6 and 12–13 correspond to reaction mixtures with and without 250 nM Mus81/Eme1 in the presence of 10 nM FEN1 and 5 nM pol β. Lanes 7 and 14 correspond to synthesized size markers (M). Substrates were <sup>32</sup>P-labeled at the 3'-end of their downstream strands. A scheme that indicates FEN1 cleavage activity in removing a double-flap with a 5'-THF is illustrated below the gels.



**Figure 2.16 FEN1 cleavage of a hairpin.**

The (CAG)<sub>7</sub>-8-oxoG/THF and (CAG)<sub>14</sub>-8-oxoG/THF hairpin substrates were labeled at the 3'-end of the hairpin-containing strand. Substrates (25 nM) were incubated with 10 nM purified FEN1 at 37°C for 30 min (lanes 2, 4, 6 and 8). Lanes 1, 3, 5 and 7 represent a substrate alone. Substrates are illustrated schematically above the gels.

## **Pol $\beta$ gap-filling synthesis facilitated the production of the unexpanded product during BER of a base lesion on a large CAG hairpin loop**

As a central component of BER, pol  $\beta$  fills in gaps, performs strand-displacement synthesis and modulates TNR stability during BER (11). Thus it may also play a critical role during BER in a hairpin loop. We found that pol  $\beta$  stimulated the production of the unexpanded product from the large (CAG)<sub>14</sub> hairpin (Figure 2.5, lanes 20 and 21; Figure 2.6, lanes 16 and 17), but not that from the small (CAG)<sub>7</sub> hairpin (Figure 2.5, lanes 7 and 8; Figure 2.6, lanes 5 and 6). The absence of pol  $\beta$  reduced the production of the unexpanded product from the large hairpin (Figure 2.5, lanes 18 and 19; Figure 2.6, lanes 14 and 15) and the long double-flap substrates (Figure 2.11, lanes 12 and 13), indicating that pol  $\beta$  is critical for removing the entirety of a large hairpin. In addition, pol  $\beta$  performed more efficient DNA synthesis during BER in the small hairpin loop (Figure 2.5, lanes 10 and 12; Figure 2.6, lanes 8 and 10) and the short double-flap (Figure 2.11, lanes 7 and 9) than in the large hairpin loop (Figure 2.5, lanes 23 and 25; Figure 2.4, lanes 19 and 21) and the long double-flap (Figure 2.11, lanes 17 and 19). This indicates that a short 3'-flap from APE1 5'-incision of a small hairpin readily annealed to its template strand and was efficiently extended by pol  $\beta$ , whereas a long 5'-flap from a large hairpin partially annealed to its template by folding into an intermediate with a small hairpin, thus inhibiting pol  $\beta$  DNA synthesis. Our results further demonstrate that pol  $\beta$  was capable of performing DNA synthesis with the (CAG)<sub>3</sub>/(CAG)<sub>4</sub> double-flap substrate in the absence of Mus81/Eme1 (Figure 2.11, lane 4). This indicates that a small (CAG)<sub>7</sub> hairpin was readily converted into a short flap during BER and was subsequently fully annealed to the template strand leaving a base-paired 3'-terminus for pol  $\beta$  DNA

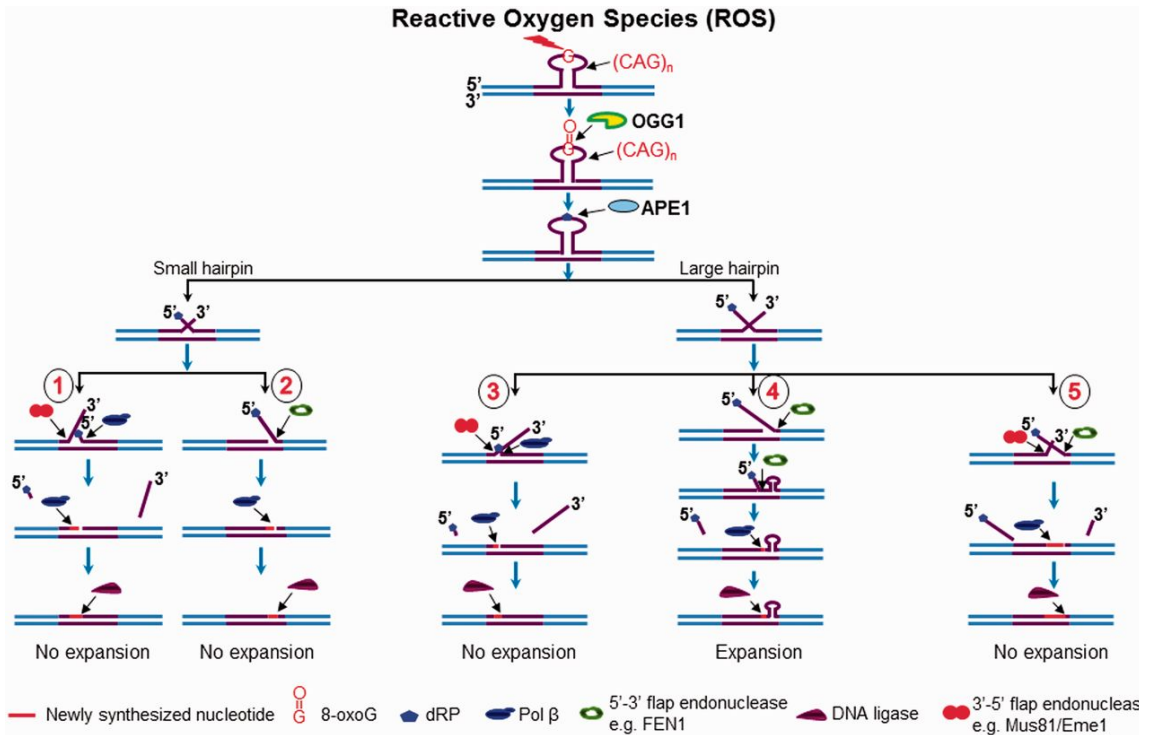
synthesis without need of Mus81/Eme1 cleavage. However, for a large (CAG)<sub>14</sub> hairpin, pol  $\beta$  failed to synthesize DNA in the absence of Mus81/Eme1 (Figure 2.11, lane 14), indicating that the long flap formed from a large hairpin failed to completely anneal to the template strand, leaving a short 3'-flap that prevented pol  $\beta$  DNA synthesis and needs to be removed by Mus81/Eme1. This further demonstrates that removal of a 3'- or 5'-flap by Mus81/Eme1 and FEN1 plays a crucial role in facilitating pol  $\beta$  DNA synthesis (Figure 2.11, lane 9 and lane 19). Thus, we suggest that pol  $\beta$  DNA synthesis can be hindered by a 3'-flap rather than by the Mus81/Eme1 protein complex during BER in hairpin loops. Removal of a 3'-flap by Mus81/Eme1 is essential in facilitating pol  $\beta$  DNA synthesis during BER in a large hairpin.

## **DISCUSSION**

In this study, we provide the first evidence that 8-oxoG in the loop region of a CAG repeat hairpin in a CAG repeat tract can be removed by OGG1, leaving an abasic site that can be subsequently incised by APE1 (Figures 2.5 and 2.6). We demonstrate that OGG1 removes a base lesion located in the loop region of a small (CAG)<sub>7</sub> hairpin more efficiently than that in a large (CAG)<sub>14</sub> hairpin loop (Figure 2.5, compare lanes 2–3 with lanes 15–16). Furthermore, we found that 5'-incision of an abasic site in the loop region of a hairpin resulted in an ssDNA break in the loop that was repaired by BER, leading to partial or complete removal of the hairpin (Figures 2.5 and 2.6). This indicates that BER in the loop of a hairpin can be coupled with the removal of the hairpin. We show that BER in a CAG hairpin loop results in removal of the hairpin, leading to attenuation of CAG repeat expansion. This is accomplished by the coordination among FEN1 cleavage

of a 5'-flap, cleavage of a 3'-flap by a 3'-5' endonuclease such as Mus81/Eme1 and pol  $\beta$  DNA synthesis (Figures 2.5, 2.6 and 2.11). We further demonstrate that as a 3'-5' flap endonuclease, Mus81/Eme1 removed both a small (CAG)<sub>7</sub> hairpin and a large (CAG)<sub>14</sub> hairpin by cleaving a 3'-CAG repeat flap (Figures 2.5, 2.6, 2.11 and 2.12). FEN1 5'-flap cleavage resulted in complete removal of both a small and large hairpin, as well as partial removal of a large hairpin (Figures 2.5, 2.6 and 2.11). This was demonstrated by the fact that FEN1 cleavage resulted in the unexpanded product (Figures 2.5, 2.6 and 2.11) as well as a series of shortened expanded products (The right panel of Figures 2.5, 2.6 and 2.11). These results indicate that a single-strand break in the loop region of a CAG repeat hairpin converts the hairpin into an intermediate with a 3'- and 5'-flap that can be cleaved by a 3'-5' endonuclease or FEN1. This further suggests that a long 5'-flap can fold into a small hairpin with a short 5'-flap that is cleaved by FEN1 alternate flap cleavage activity. Finally, we found that the presence of both Mus81/Eme1 and FEN1, along with pol  $\beta$ , significantly stimulated the production of the unexpanded product, indicating that a 3'-5' endonuclease such as Mus81/Eme1 can cooperate with FEN1 and pol  $\beta$  to remove a hairpin during BER, thereby preventing CAG repeat expansion. Our results support a model that is shown in Figure 2.17, in which an 8-oxoG in the loop region of a CAG repeat hairpin is induced by oxidative DNA damage. OGG1 removes the base lesion, leaving an abasic site that can be 5'-incised by APE1. This results in a single-strand break in the loop region of the hairpin, converting the cleaved hairpin into a double-flap intermediate with a 3'-flap and a 5'-flap. For a small hairpin, a short 3'-flap and a 5'-flap are generated and cleaved by the 3'-5' endonuclease Mus81/Eme1 and the 5'-3' endonuclease FEN1. This leads to a complete removal of the hairpin via BER of an 8-

oxoG in the hairpin loop, thereby preventing repeat expansion (Figure 2.17, sub-pathways 1 and 2). Repair of an 8-oxoG in a large hairpin loop results in a double-flap



**Figure 2.17 Prevention or attenuation of CAG repeat expansion by BER coupled with removal of a hairpin.**

Oxidative stress can induce an oxidized DNA base lesion, 8-oxoG in the loop region of a CAG hairpin. OGG1 removes 8-oxoG leaving an abasic site that is subsequently 5'-incised by APE1 generating ssDNA break in the hairpin loop. This further results in the formation of a double-flap intermediate with a 3'- and 5'-CAG repeat flap. For a small hairpin, APE1 5'-incision leads to the formation of a double-flap intermediate with a short 3'- and 5'- flap that can be completely cleaved by a 3'-5' endonuclease, such as Mus81/Eme1, or by FEN1 flap cleavage. The flap cleavage can result in a gap that is filled by pol β DNA synthesis and this completes damage repair. This results in prevention of repeat expansion (sub-pathway 1 and 2). Repair of a base lesion located in the loop region of a large CAG repeat hairpin results in the formation of a relatively long 3'- and 5'-flap. The 5'-flap can anneal to the template strand to create a long 3'-flap that is cleaved by a 3'-5' endonuclease such as Mus81/Eme1, thereby preventing repeat expansion (sub-pathway 3). A 3'-flap can also anneal to its template to create a long 5'-flap for FEN1 cleavage. FEN1 cleaves the 5'-flap either by its conventional flap cleavage to remove the entire hairpin or by its alternate flap cleavage to partially remove the hairpin. This results in prevention or attenuation of repeat expansion (sub-pathway 3 and 4). A double-flap intermediate can also be processed simultaneously by Mus81/Eme1 and FEN1 leading to the prevention of repeat expansion (sub-pathway 5).

intermediate with a long 3'-flap and 5'-flap. The 3'-5' endonuclease Mus81/Eme1 removes the entirety of the hairpin by cleaving the 3'-flap, preventing repeat expansion (Figure 2.17, sub-pathway 3). FEN1 processes the 5'-flap either by its conventional flap cleavage to completely remove the hairpin or by its alternate cleavage activities to partially remove the hairpin. This results in prevention or attenuation of repeat expansion (Figure 2.17, sub-pathway 4). Simultaneous cleavage of a 3'-flap and a 5'-flap by Mus81/Eme1 and FEN1, in cooperation with pol  $\beta$  gap-filling synthesis, prevents FEN1 alternate flap cleavage, thereby promoting the prevention of repeat expansion (Figure 2.17, sub-pathway 5).

Removal of hairpin structures has been shown to be a challenge to various biological systems, in particular to eukaryotes. In *E. coli*, a structure-specific endo/exonuclease complex SbcCD has been found to endonucleolytically incise the loop region of a CAG/CTG hairpin by using its endonuclease activity, and then exonucleolytically cleave the hairpin stem with its dsDNA exonuclease activity (213,214). In bacteriophage T7, endonuclease I can cleave a looped-out CAG/CTG repeat bubble structure, raising the possibility that this enzyme may also cleave a CAG/CTG repeat hairpin (215). However, in yeast, CTG and CAG repeat hairpin structures can readily escape cellular repair mechanisms, leading to repeat expansion (128). In human cell extracts, three different paths for resolving a hairpin have been proposed, depending on the location of the hairpin. A CTG hairpin or bubble (named slipped repeats or slip-out) formed in a nascent DNA strand with a 5'-nick can be partially removed, leading to repeat expansion (216,217). A hairpin opposite a 5'-nick in the template strand results in complete removal of the hairpin and maintenance of repeat length (216,217). This may be

mediated by destabilization of a template hairpin by DNA polymerase passing-through of the hairpin via DNA synthesis (199) or by a DNA helicase (218). A hairpin opposite a 3'-nick in the template strand can lead to repeat deletion (216,217). It has been found that the mismatch repair protein complexes, MSH2/MSH3 (MutS $\beta$ ), MutL $\alpha$  and PMS2 may mediate the removal of small (CAG)<sub>1</sub>-(CAG)<sub>3</sub> hairpins/slip-outs (119), but not for large hairpins (219,220), suggesting an important role of mismatch repair in removing small TNR hairpins during DNA replication and repair. In this study, for the first time, we identified BER as a new pathway for removing CAG repeat hairpins in a CAG repeat tract via repair of a DNA base lesion in the hairpin loop region. Our results indicate that BER in CAG hairpin loops incises the hairpin loop region, converting the hairpins into intermediates with a 5'- and a 3'-flap that are cleaved by FEN1 and the 3'-5' flap endonuclease, Mus81/Eme1. This ultimately results in removal of a CAG repeat hairpin and prevention or attenuation of TNR expansion. Thus, here we suggest a unique pathway for removing TNR hairpins, which is induced by a base lesion and mediated by BER at the hairpin loop region.

Our study demonstrates an important role of the flap cleavage activity of both 3'-5' and 5'-3' endonucleases such as Mus81/Eme1 and FEN1 in preventing or attenuating TNR expansion by removing hairpin structures through coupling with BER. Our discovery is supported by a finding that the nucleotide excision repair protein XPG, which is also a 5'-3' flap endonuclease (221), can promote removal of a hairpin through a 5'-incision (222). It is also supported by the fact that the 3'-5' endonuclease/exonuclease Mre11-Rad50-Nbs1 complex can facilitate the removal of CAG/CTG hairpins *in vitro* and *in vivo* (116). The 3'-5' exonuclease activity of Mre11 was initially identified as an



enzymatic activity that resects the 3'-end of a double-stranded DNA for repairing double-strand DNA breaks (223,224). It is possible that this activity of Mre11 may also cleave TNRs from the 3'-end of an upstream TNR-containing strand, removing an extra number of the repeats. This further allows the downstream repeat flap to anneal to its template, thereby promoting removal of a hairpin and prevention of repeat expansion. It is conceivable that other 3'-5' and 5'-3' endo/exonucleases may also use the same mechanisms as the ones exemplified by Mus81/Eme1 and FEN1 to prevent TNR expansion.

Our results showed that more APE1 incision products than OGG1 products were produced (Figure 2.3, compare lanes 2 and 15 with lanes 3 and 16), indicating that OGG1 removed an 8-oxoG in a hairpin loop more efficiently in the presence of APE1 than in the absence of the enzyme. This is consistent with the notion that APE1 can facilitate OGG1 incision of an 8-oxoG by kicking OGG1 off an abasic site, releasing OGG1 from its product, and thereby increasing OGG1 recycling and its efficiency for removing oxidative DNA damage (225,226). This indicates that APE1 can also stimulate OGG1 incision of 8-oxoG by dislodging the enzyme from an abasic site located in a hairpin loop. Thus, the coordination between OGG1 and APE1 in removing DNA base lesions is conserved in both duplex DNA and hairpins. BER cofactors that can stimulate OGG1 and APE1 activity may stimulate removal of hairpins during BER, thereby facilitating prevention or attenuation of repeat expansion. For example, BER cofactor X-ray repair cross-complementing protein 1 (XRCC1) physically interacts with OGG1 and stimulates its activity by 2- to 3-fold, presumably by recruiting OGG1 to an 8-oxoG (191,227). High mobility group box 1B (HMGB1) can stimulate APE1 5'-incision of an abasic site (193).

Thus these cofactors may promote removal of hairpins and prevention of repeat expansions by stimulating OGG1 and APE1 activity. The coordination among OGG1, XRCC1, APE1 and HMGB1 in repairing a base lesion in the context of a TNR hairpin that is coupled with removal of a hairpin remains to be elucidated.

In this study, we showed that BER enzymes, pol  $\beta$  and FEN1 can coordinate with a 3'-5' flap endonuclease represented by Mus81/Eme1 to remove a relatively large CAG repeat hairpin during BER in the hairpin loop region. The repair process is accomplished by the coordination among Mus81/Eme1 cleavage on a 3'-flap, FEN1 removal of a 5'-flap and pol  $\beta$  gap-filling synthesis. We found that FEN1 alone resulted in the unexpanded product as well as a series of shortened expansion products, indicating that FEN1 can cleave a part of a long CAG flap using its alternate flap cleavage activity. This activity of FEN1 can be inhibited by its cooperation with a 3'-5' flap endonuclease such as Mus81/Eme1, thereby promoting the complete removal of the CAG repeat hairpin and prevention of repeat expansion. Thus, the cooperation between FEN1 and a 3'-5' flap endonuclease plays a critical role in preventing CAG repeat expansion during BER. Because the production of the unexpanded product was also stimulated by pol  $\beta$  (Figure 2.3, compare lane 19 with lane 21; Figure 2.4, compare lane 15 with lane 17; Figure 2.9A, compare lane 13 with lane 15), this further indicates that pol  $\beta$  fills in a gap resulting from the 3'- and 5'-flap cleavage by Mus81/Eme1 and FEN1, creating a ligatable nick for DNA ligase to complete the repair. Thus, it appears that the coordination among FEN1, the 3'-5' flap endonuclease Mus81/Eme1 and pol  $\beta$  plays a crucial role in removing a large CAG repeat hairpin and prevention of CAG repeat expansion.

Interestingly, we found that for a small hairpin with a base lesion in the loop region, FEN1 alone can completely remove the downstream flap to produce the non-expansion product, thereby preventing repeat expansion (Figure 2.3, lane 5; Figure 2.4, lane 3; Figure 2.9A, lane 2). This is because a short upstream flap can anneal back to the template and displace the downstream strand, creating a downstream 5'-flap. This is consistent with the fact that FEN1 can efficiently remove a small hairpin during BER by removing a short flap via flap equilibration, as demonstrated previously (228). For a large hairpin that can be converted to an intermediate with long double-flaps, FEN1 can use its conventional or alternate flap cleavage to completely or partially remove the hairpin. However, the efficiency of FEN1 in removing a large hairpin may be modulated by interacting with other replication and repair proteins. For example, the Bambara group has demonstrated that replication protein A (RPA) can inhibit FEN1 cleavage of a long flap with 30 nt but not a short flap (229,230). This may compromise the efficiency of FEN1 to remove a large hairpin. Yet, the inhibitory effect from RPA may be alleviated by the cooperation between FEN1 and Dna2, a replication helicase and nuclease (231). This is because binding of RPA to the long flap can stimulate Dna2 cleavage on the flap, which shortens it. This allows FEN1 to readily load onto the short flap and remove it completely (231). It is conceivable that during BER of a base lesion in a large hairpin, a long double-flap can be initially bound by RPA. This then stimulates Dna2 cleavage activity, which shortens a long 5'-flap. FEN1 subsequently cleaves the shortened 5'-flap efficiently, leading to complete removal of the large hairpin. Moreover, because BER cofactors and enzymes of other repair pathways such as proliferating cell nuclear antigen (PCNA) and the 5'-3' exonuclease Exo I can facilitate FEN1 flap cleavage (232,233), it is

possible that these proteins may also facilitate removal of a TNR hairpin. The cooperation among FEN1 and other BER proteins and cofactors, as well as the enzymes of other repair pathways in removing hairpin structures during BER needs to be elucidated.

In summary, in this study we discovered a unique pathway for removal of a TNR hairpin via BER in the loop region of the hairpin. We have demonstrated that OGG1 and APE1 can remove an 8-oxoG located at the hairpin loop region, leaving an abasic site that can be further incised by APE1, and resulting in incision of the hairpin and the conversion of the hairpin into an intermediate with a 3'-flap and a 5'-flap. For a small hairpin, FEN1 or a 3'-5' endonuclease such as Mus81/Eme1 alone can completely remove the hairpin by cleaving a 5'-flap or a 3'-flap, whereas removal of a large hairpin via BER is mediated by the cooperation among the cleavage of a 5'-flap by FEN1, the cleavage of a 3'-flap by a 3'-5' flap endonuclease and pol  $\beta$  gap-filling synthesis. Our study indicates that BER is coupled with the removal of a TNR hairpin to prevent or attenuate TNR expansion, and we suggest a new role of oxidative DNA damage and BER in attenuating TNR expansion by removing a CAG repeat hairpin.

In CHAPTER 1 and CHAPTER 2, I have demonstrated that BER of an oxidative base lesion in the context of TNR can result in reduction of TNR units. The results shown in CHAPTER 1 and CHAPTER 2 indicate that BER can perform efficient repair of the damage that is coupled with removal of TNR units whether the oxidative DNA lesion is located in a duplex DNA or in a hairpin loop region. Thus, we conclude that BER of oxidative DNA damage in a TNR repeat tract can prevent repeat expansion by promoting shortening of a TNR tract. However, some repair resistant oxidative damage such as

cyclodeoxypurines can accumulate in a TNR tract conferring a challenge for DNA repair and have to be bypassed by DNA polymerase lesion bypass synthesis. Because pol  $\beta$  can perform lesion bypass synthesis at an AP site, this raises a new question regarding the roles of BER enzymes, especially a role of pol  $\beta$  in modulating TNR instability during its lesion bypass. To explore pol  $\beta$ 's role in modulating TNR instability via its bypass of an oxidized DNA base lesion, in CHAPTER 3 we examined lesion bypass synthesis of pol  $\beta$  with a 5',8-cyclodeoxypurine and its effects on TNR instability during BER and replication by employing a series of substrates that contain a helix distorting base lesion, 5',8- cyclo-2'-deoxyadenosine located in a CAG repeat tract.

**CHAPTER 3: A 5',8-CYCLODEOXYPURINE LESION INDUCES  
TRINUCLEOTIDE REPEAT DELETION VIA A UNIQUE LESION BYPASS BY  
DNA POLYMERASE  $\beta$**

**ABSTRACT**

5',8-cyclo-deoxypurines are common forms of oxidized DNA lesions resulting from endogenous and environmental oxidative stress such as ionizing radiation. The lesions can only be repaired by nucleotide excision repair with a low efficiency. This results in their accumulation in the genome that leads to stall of replication DNA polymerases and poor lesion bypass by translesion DNA polymerases. Trinucleotide repeats (TNR) consist of tandem repeats of Gs and As, and are therefore susceptible to generation of 5',8-cyclo-deoxypurines. In this study, we provided the first evidence that 5',8-cyclo-dA in a CAG repeat tract induced the formation of a small CAG loop. This further promoted DNA polymerase  $\beta$  (pol  $\beta$ ) skip-over of the loop structure, thereby exclusively resulting in a CTG repeat deletion during DNA lagging strand synthesis and base excision repair. We demonstrated that pol  $\beta$  bypassed a 5',8-cyclo-dA resulting in deletion with varying sizes. This was caused by formation of various sizes of CAG repeat loops that was bypassed by pol  $\beta$ . This resulted in a long flap that was subsequently cleaved by flap endonuclease 1 (FEN1). Our study indicates that accumulation of 5',8-cyclo-deoxypurines in the human genome can lead to TNR instability via a unique lesion bypass by pol  $\beta$ .

## INTRODUCTION

Reactive oxygen species (ROS) generated from environmental stresses such as ionizing radiation and endogenous oxidative stresses from energy metabolism, can result in more than 30 types of oxidative DNA base lesions in the human genome(24,234). Among ROS, hydroxyl radical ( $\text{HO}^\bullet$ ) is the one that leads to production of most of oxidative base lesions because of its high reactivity with DNA bases(235). Among oxidative DNA base lesions induced by  $\text{HO}^\bullet$ , 5',8-cyclo-2'-deoxypurines (cdPu) including 5',8-cyclo-2'-deoxyadenosine (5',8-cdA) and 5',8-cyclo-2'-deoxyguanosine (5',8-cdG) are DNA base lesions that can induce DNA structure abnormality(234,236-238). These lesions usually cause more severe biological adverse effects than 8-oxoguanine, the most frequently occurred oxidized base lesion. Cyclodeoxypurine lesions contain the glycosidic bond along with an extra covalent bond formed between the C5' position of the 2'-deoxyribose and the C8 position of the purine. Cyclodeoxypurine lesions occur in two diastereomeric forms with a 5'R or 5'S configuration. It has been found that similar amount of the 5'S and 5'R diastereomers of cyclodeoxynucleoside lesions are generated in double-stranded DNA by  $\text{HO}^\bullet$  *in vitro*.

Cyclodeoxypurine base lesions are frequently detected in the genomic DNA of mammalian cells and cells of other organisms such as English sole. It is estimated that in fetal and postnatal rat liver, 180-320 lesion/cell of cdPus can be produced (235). Although most of oxidative DNA base lesions are removed by DNA glycosylases such as 8-oxoguanine DNA glycosylase (OGG1) and Endonuclease III-like protein 1 (NTH1) through base excision repair (BER) pathway (35,42), the CPU base lesion can only be

repaired by nucleotide excision repair (NER) (35,239-241). This is because the extra covalent bond of cdPu stabilizes the glycosidic bond and prevents the lesion from being removed by DNA glycosylases and direct hydrolysis of damage (42,242,243). However, cellular efficiency of removing a 5'R- and 5'S-cdA lesion by NER is 2- to 4-fold less than that for NER to repair a cis-B[ $\alpha$ ]P-N<sup>2</sup>-dG adduct (239). Similar to that for repairing a cyclo-dA, the efficiency for NER to repair a 5'R-cdG is only 70% of that in removing a cis-B[ $\alpha$ ]P-N<sup>2</sup>-dG adduct. In addition, the efficiency for NER to repair a 5'S-cdG lesion is 2-fold less than that for its repair of a 5'R-cdG (239). Thus, it appears that a low efficiency for repairing cdPu lesions in cells can lead to accumulation of a high level of the lesions, especially the 5'S configuration lesion in the genomic DNA. This then results in a series of adverse effects on cellular function. These include replication fork stalling, inhibition of DNA synthesis by replication DNA polymerases such as human pol  $\delta$  and pol  $\epsilon$  and T7 DNA polymerase (240), inhibition of translesion DNA synthesis as well as inhibition of gene transcription. It has been found that a 5'S-cdA in a TATA box of a promoter region can inhibit promoter binding of TATA binding protein, thus preventing recruitment of RNA polymerase II to a transcription start site(244). Moreover, a cdPu can directly inhibit RNA synthesis by RNA polymerase II (35,245). Several studies showed that human and bacteria Y family translesion synthesis polymerases such as pol  $\eta$ , pol  $\iota$  and pol  $\xi$  are strongly blocked by a template cdPu, especially a 5'S-cdA and 5'S-cdG during a lesion bypass synthesis (246-248). This may leads to accumulation of DNA strand breaks that can ultimately cause cell death. In addition, the translesion synthesis to bypass a cdPu by human Y family polymerases is highly mutagenic. For example, bypass of a 5'-S-cdA by pol  $\zeta$  and pol  $\kappa$  causes an A to T transversion, and bypass of a 5'S-cdG



results in a G to A transition and a G to T transversion (248). Thus, accumulation of cdPus in the genome can result in mutagenesis and cell death that are associated with development of human diseases including cancer, neurodegeneration and aging (75,241,249-252).

Trinucleotide repeats (TNRs) are highly polymorphic in human genome and can be expanded or deleted during DNA metabolisms including DNA replication, repair and transcription (71,132,133,253-257). TNR expansion is associated with more than 40 neurodegenerative diseases such as Huntington's disease, myotonic dystrophy, fragile X syndrome, and Friedreich's ataxia (74,253,258). TNR deletion in androgen receptor is related to ovarian and prostate cancers (82,179,259). It has been found that formation of non-B form DNA secondary structures such as hairpins, triplexes and tetraplexes is responsible for TNR expansion and deletion during DNA replication, repair, recombination and gene transcription (4,85,86,140,198,260). The secondary structures formed in a newly synthesized strand or damaged strand of the genome usually result in replication fork stalling (4,107), pausing of DNA polymerases and DNA slippage (4,108), trapping of mismatch repair proteins (113,114), inhibition of flap endonuclease 1 (FEN1) processing of a hairpin (111,167,261) and disruption of coordination among repair enzymes (115), and this ultimately leads to TNR expansion. In contrast, TNR secondary structures formed at the template strand during DNA repair promote DNA polymerases to skip over the structures, and this causes TNR deletion.

TNR contains a series of adenosines and guanines. This makes it as a hotspot for generating oxidative DNA damage including cdPu lesions. Previous studies have shown that oxidative DNA damage and its repair in a TNR tract are associated with TNR

expansion and deletion (200,257,262). An oxidative DNA damaging agent, H<sub>2</sub>O<sub>2</sub> can induce small CAG repeat expansions in human cells (200,262) and TNR deletion in bacteria and mouse kidney cells (26,177) as well. This appears to be associated with production of 8-oxoguanine (8-oxoG). This is supported by the finding that an increased level of 8-oxoG in the brain of HD transgenic mouse models was correlated with CAG repeat expansion (103,263). This is also supported by the fact that an environmental oxidative DNA damaging agent, potassium bromate induced production of 8-oxoG in the germ cells of a fragile X syndrome mouse model and caused CGG repeat expansion (154). This is further supported by our recent discovery showing that potassium bromate can induce both expansion and deletion of CTG repeats in human kidney cells (200). We have identified a molecular basis underlying TNR instability induced by oxidative DNA damage, i.e. the formation of a hairpin in the damaged strand or the template strand during BER in a TNR track (115,200,257). This subsequently led to synthesis of extra repeat units by pol  $\beta$  (200,264) or pol  $\beta$  skip-over of a template hairpin (200,257). By coordinating with FEN1 alternate flap cleavage, this in turn resulted in TNR expansion or deletion (115,200,257). Interestingly, we have recently discovered that BER of an 8-oxoG or an abasic site located in a hairpin loop region can also lead to a partial or complete removal of a hairpin (265), thereby preventing repeat expansion. All the findings indicate that an oxidative base lesion can modulate TNR instability through its repair on the damaged strand. However, it is unknown if and how an oxidative DNA base lesion located on the template strand of a TNR tract can induce TNR instability.

Although a significant progress has been made in exploring bypass of a cdPu lesion by translesion synthesis polymerases during DNA replication with a relatively low

efficiency, it remains unknown that whether other repair DNA polymerases can bypass cdPu lesions during DNA replication and repair to help cells to survive the base lesions that are resistant to DNA repair. Similar to the Y family translesion synthesis polymerases, pol  $\beta$  can bypass an apurinic/apyrimidinic site (AP site) by inserting an adenosine opposite to the damage (266). In addition, pol  $\beta$  can switch with pol  $\delta$  and pol  $\epsilon$  to bypass an AP site (267,268). During Okazaki fragment maturation, pol  $\beta$  can bypass an AP site that was annealed with a downstream primer by switching with pol  $\delta$  in a strand-displacement synthesis dependent manner (267). Moreover, because pol  $\beta$  tends to skip over a TNR hairpin structure on the template (256,257), we further hypothesize that pol  $\beta$  may also skip over a cdPu lesion located at a TNR tract to bypass the lesion during DNA replication and BER, and this may subsequently cause TNR instability. To test this hypothesis, in this study, we examined pol  $\beta$  DNA synthesis in bypassing a 5',8-cdA in the CAG repeat-containing template strand and determined the effects of the base lesion on CTG repeat instability. We provide the first evidence that a 5',8-cdA in a CAG repeat tract of the template strand can be bypassed by pol  $\beta$  skip-over of a small bubble that contains the base lesion. This preferentially results in CTG repeat deletion.

## **MATERIALS AND METHODS**

### **Materials**

DNA oligonucleotides containing a 5'S-5',8-cdA or 5'R-5',8-cdA were synthesized and HPLC purified according to the procedures described previously(239). All other oligonucleotides were purchased from Integrated DNA Technologies (IDT Inc.,

Coralville, IA). The radionucleotides [ $\gamma$ - $^{32}$ P] ATP (6000 mCi/mmol) and Cordycepin 5'-triphosphate 3'-[ $\alpha$ - $^{32}$ P] (5000 mCi/mmol) were purchased from PerkinElmer Inc. (Boston, MA). Micro Bio-Spin 6 chromatography columns were from Bio-Rad (Hercules, CA). Deoxynucleoside 5'-triphosphates (dNTPs) were from Fermentas (Glen Burnie, MD). S1 Nuclease was from Promega (Madison, WI). All other standard chemical reagents were purchased from Sigma-Aldrich (St. Louis, MO) and Thermo Fisher Scientific (Pittsburgh, PA). Purified pol  $\beta$ , FEN1, and DNA ligase I (LIG I) were generous gifts from Dr. Samuel H. Wilson at the National Institute of Environmental Health Sciences (NIEHS)/National Institutes of Health, Research Triangle Park, NC.

### **Oligonucleotide Substrates**

Substrates containing a template strand that has (CAG)<sub>10</sub> repeats with a normal A, or 5'S-cdA, or 5'R-cdA located at the fifth repeat units counted from the 5'-end were designed to mimic DNA replication or repair intermediates without or with a template 5',8-cdA lesion. The downstream primers of the substrates with (CTG)<sub>9</sub>, (CTG)<sub>5</sub> repeats or (CTG)<sub>1</sub>, contained a 5'-phosphate or a tetrahydrofuran (THF) residue with a 5'-phosphate, an abasic site analog. The corresponding upstream primers contained a 3'-(CTG)<sub>1</sub>, (CTG)<sub>5</sub> or (CTG)<sub>9</sub> repeats, respectively. The substrates were used to represent DNA lagging strand synthesis or BER intermediates with a 1-nt or 6-nt gap opposite to a template 5',8-cdA lesion at different positions of the (CAG)<sub>10</sub> repeat tract. Substrates representing intermediates with a 1-nt gap were constructed by annealing the upstream and downstream primers to their template strands at a molar ratio of 1:1:1.5. Substrates representing replication intermediates with a 6-nt gap opposite to a template 5',8-cdA

lesion at different positions of the (CAG)<sub>10</sub> repeat tract were constructed by annealing the upstream primers with a 3'-(CTG)<sub>4</sub> or 3'-(CTG)<sub>5</sub> repeats and downstream primers contained a 5'-phosphate with (CTG)<sub>4</sub> or (CTG)<sub>3</sub> repeats with the template strands with or without a 5',8-cdA damage at a molar ratio of 1:1:1.5. Substrates representing DNA leading strand synthesis intermediates were constructed by annealing the upstream primer with (CTG)<sub>5</sub> repeats with the template strands with or without 5',8-cdA damage at a molar ratio of 1:1.5. Substrates representing intact double strand DNA with or without a template 5',8-cdA damage in a (CAG)<sub>10</sub> repeat tract were constructed by annealing the template strands with their complimentary strand at a molar ratio of 1:1.5. Substrates were radiolabeled at the 5'-end of the upstream primers, or the 5'-end of the template strands, or at the 3'-end of the downstream primers for measuring various types of enzymatic activities. The sequences of oligonucleotide substrates are listed in Table 3.1. DNA size markers that correspond to the repaired or ligated products after BER or lagging strand maturation, FEN1 cleavage products and S1 Nuclease cleavage products were synthesized and purified by Urea denaturing polyacrylamide gel electrophoresis (191)

**Table 3.1 Oligonucleotide sequence**

Oligonucleotides	nt	Sequence (5' –3')
<b>Downstream</b>		
D1	27	pF- GCTGCTGCTGCTGTACGGATGCTAGAT
D2	41	pF- CTGCTGCTGCTGCTGCTGCTGCTGCTGTACGGATGCTAGAT
D3	14	pF- TACGGATGCTAGAT
D4	27	p- GCTGCTGCTGCTGTACGGATGCTAGAT
D5	41	p- CTGCTGCTGCTGCTGCTGCTGCTGCTGTACGGATGCTAGAT
D6	14	p- TACGGATGCTAGAT
D7	22	p- TGCTGCTGTACGGATGCTAGAT
D8	26	p- CTGCTGCTGCTGTACGGATGCTAGAT
<b>Template</b>		
T1	58	ATCTAGCATCCGTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTACGTAGACTTAC T
T2	58	ATCTAGCATCCGTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTACGTAGA CTTACT
T3	58	ATCTAGCATCCGTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTACGTAGA CTTACT
<b>Upstream</b>		
U1	30	AGTAAGTCTACGTA CTGCTGCTGCTGCTGC
U2	16	AGTAAGTCTACGTA CT
U3	43	AGTAAGTCTACGTA CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
U4	26	AGTAAGTCTACGTA CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT

F: tetrahydrofuran, THF

## ***In Vitro* BER and ligation reaction through lesion bypass reconstituted with purified enzymes**

*In vitro* BER of an intermediate with a 1-nt gap at different locations opposite to a template 5',8-cdA lesion in the context of (CAG)<sub>10</sub> repeats was performed with purified pol β, FEN1, LIG I, and substrates with a template normal A or 5',8-cdA and a 5'-phosphate THF residue in the downstream primers. Ligation reactions through lesion bypass reactions that represent steps of Okazaki fragment maturation were performed with purified pol β, FEN1, LIG I, and substrates containing a 5'-phosphate in the downstream primers with a normal A or 5',8-cdA. Ten microliters of reaction was reconstituted with the indicated concentrations of substrates and BER enzymes in buffer that contained Tris pH 7.5, 5 mM MgCl<sub>2</sub>, 2 mM ATP and 50 μM dNTPs. Reaction mixtures were assembled on ice and incubated at 37°C for 30 min. Reactions were then terminated by transfer to 95°C for 10 min in buffer containing 95% formamide and 10 mM EDTA. Substrates and products were separated by 15% urea-denaturing PAGE and detected by PhosphorImager. Substrates were <sup>32</sup>P-labeled at the 5'-end of the upstream primers. BER and lagging strand maturation reactions mediated by pol β lesion bypass synthesis was reconstituted by incubating purified pol β, FEN1, and LIG I with 25 nM substrates that contained a 5'R-cdA or a 5'S-cdA in the template strand with or without a downstream primer.

## **Enzymatic activity assays**

Pol β DNA synthesis was measured using substrates that mimic replication and BER intermediates with a template strand that contained (CAG)<sub>10</sub> repeats with or without

a 5',8-cdA. DNA synthesis was measured at 37°C in a 10- $\mu$ l reaction mixture that contained BER reaction buffer containing Tris pH 7.5, 5 mM MgCl<sub>2</sub> and 50  $\mu$ M dNTPs. FEN1 flap cleavage activity was examined in the same buffer with 5 mM MgCl<sub>2</sub> and 50  $\mu$ M dNTPs in the absence or presence of increasing concentrations of pol  $\beta$  at 37°C for 30 min. Substrates and products were separated by 15% urea-denaturing PAGE and detected by a PhosphorImager.

### **Probing of secondary structures by S1 Nuclease digestion**

The formation of CAG repeat hairpin or bubble structures in the template strand was probed by incubating 0.15 U or 0.05 U S1 Nuclease with 25 nM substrates that contained (CAG)<sub>10</sub> repeats with or without a 5',8-cdA embedded in the fifth CAG in the template strands in the presence or absence of pol  $\beta$ . Substrates were pre-incubated with 5 nM pol  $\beta$  at 37°C for 30 min to generate a pol  $\beta$  lesion bypass synthesis intermediate. The 10  $\mu$ l reaction mixture was assembled by incubating 0.15 U or 0.05 U of S1 Nuclease with 25 nM substrates in reaction buffer containing 50 mM sodium acetate (pH 4.5), 280 mM NaCl and 4.5 mM ZnSO<sub>4</sub>. The reaction was incubated at 37°C for 3, 5, 10, 15, and 30 min, and subsequently subjected to protease K digestion at 55°C for 30 min to remove S1 Nuclease. Substrates and products were separated by 15% urea-denaturing PAGE and detected by a PhosphorImager.



## RESULTS

### **A 5',8-cdA in the (CAG)<sub>10</sub> repeat tract of the template strand can be bypassed by pol $\beta$ during BER and DNA lagging strand maturation**

Pol  $\beta$  is an essential BER enzyme that performs a single-nucleotide gap-filling synthesis in short-patch BER and multi-nucleotide gap-filling and strand displacement synthesis during long-patch BER (11). It modulates TNR expansion or deletion by inserting extra repeat units or skipping over a part of a template TNR hairpin during BER (115,257). Previous studies also showed that pol  $\beta$  can efficiently bypass a template AP site by following the “A rule” (266). Moreover, pol  $\beta$  can switch with replication polymerases to bypass a template AP site during DNA leading and lagging strand synthesis (267,268). It has been shown that pol  $\beta$  coordinates with FEN1 and performs an efficient AP site bypass synthesis to facilitate Okazaki fragment maturation (267). Therefore, pol  $\beta$  could also bypass a 5',8-cdA damage in a CAG repeat track in the template strand, and this may alter the stability of CTG repeats. To test this possibility, we initially examined pol  $\beta$  lesion bypass synthesis activity with substrates containing a 1-nt gap opposite to a normal A (Figure 3.1 and 3.2, lanes 2-3) or 5'S-cdA (Figure 3.1 and 3.2, lanes 7-8) or 5'R-cdA (Figure 3.1 and 3.2, lanes 12-13) embedded in the fifth CAG unit of a (CAG)<sub>10</sub> repeat tract in the template strand. Substrates with a downstream 5'-phosphorylated THF mimic BER intermediates resulting from APE1 5'-incision of an abasic site that is opposite to a template 5',8-cdA. Substrates containing a downstream primer with a 5'-phosphate residue mimic intermediates of DNA lagging strand synthesis. We found that pol  $\beta$  performed a multi-nucleotide insertion with all the

substrates (Figure 3.1 and 3.2, lanes 2-3, 7-8, and 12-13) indicating that pol  $\beta$  bypassed a 5',8-cdA lesion located in the context of (CAG)<sub>10</sub> repeats during BER and DNA lagging strand synthesis.

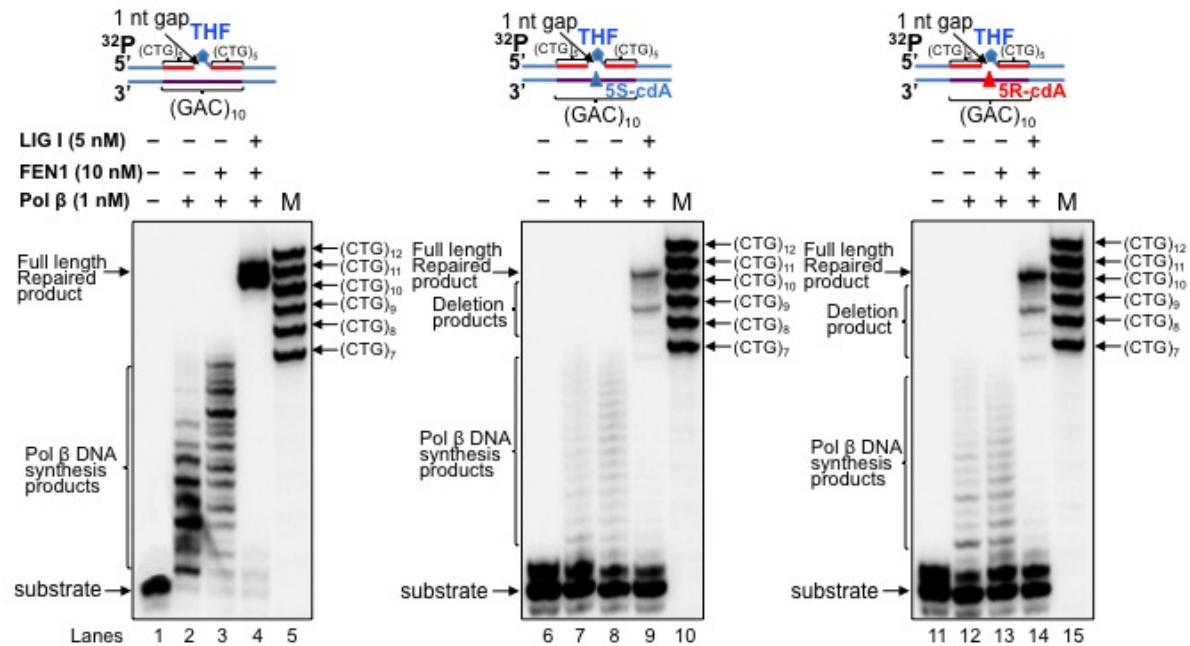
However, pol  $\beta$  DNA synthesis activity in bypassing a 5'S-cdA or 5'R-cdA was much weaker than its activity to insert a correct nucleotide to base-pair with a normal A (Figure 3.1 and 3.2, compare lanes 7-8, lanes 12-13 with lanes 12-13). Pol  $\beta$  synthesis mainly inserted 1 nt to bypass the 5',8-cdA lesion (Figure 3.1 and 3.2, lanes 7-8, 12-13) in comparison with its predominant multi-nucleotide insertion to bypass a normal A (Figure 3.1 and 3.2, lanes 2-3). In the presence of FEN1, pol  $\beta$  DNA synthesis was increased with all the substrates (Figure 3.1 and 3.2, compare lanes 3, 8 and 13 with lanes 2, 7 and 12). The results indicate that pol  $\beta$  bypass of a 5',8-cdA was enhanced by FEN1 cleavage of a downstream flap.

Because previous studies showed that pol  $\beta$  DNA synthesis efficiency varied during BER of an oxidized base lesion that was located at different locations in a TNR tract (197,200). We further examined whether the position of a 1-nt gap relative to a template 5',8-cdA in a (CAG)<sub>10</sub> tract can modulate the efficiency of pol  $\beta$  bypass synthesis. Pol  $\beta$  DNA synthesis with the substrates that contain a 1-nt gap located at the upstream or downstream of a template 5',8-cdA during BER and DNA lagging strand maturation was determined (Figure 3.3-3.6). The results showed that with the substrates containing a 1-nt gap that was located at the upstream of the 5',8-cdA, pol  $\beta$  initially performed efficient DNA synthesis reaching to 1-nt prior to the damage, and then inserted 7-15 nucleotides to bypass the lesion with a low efficiency (Figure 3.3 and 3.4, lanes 7-8 and 12-13). However, with the substrates containing a 1-nt gap located at the

downstream of a 5',8-cdA, pol  $\beta$  performed a relatively more processive DNA synthesis than its synthesis with the substrates containing a 1-nt gap at the upstream side of the damage (Figure 3.5 and 3.6, lanes 2-3). This indicates that pol  $\beta$  bypassed a 5',8-cdA with a lower efficiency, whereas it performed an efficient DNA synthesis in the regions surround the cdA lesion during BER and DNA lagging strand synthesis. This further suggests that an upstream template 5',8-cdA damage induced a DNA strand distortion, thereby destabilizing the base pairing of upstream strand with the template and inhibiting pol  $\beta$  DNA synthesis. For all the substrates, pol  $\beta$  DNA synthesis was not affected by the presence of FEN1 (Figure 3.3-3.6, compare lanes 3, 8 and 13 with lanes 2, 7 and 12) indicating that FEN1 cleavage failed to facilitate pol  $\beta$  bypass of a 5',8-cdA (Figure 3.3 and 3.4, lanes 3, 8 and 13) if a gap located at the upstream or downstream of the lesion.

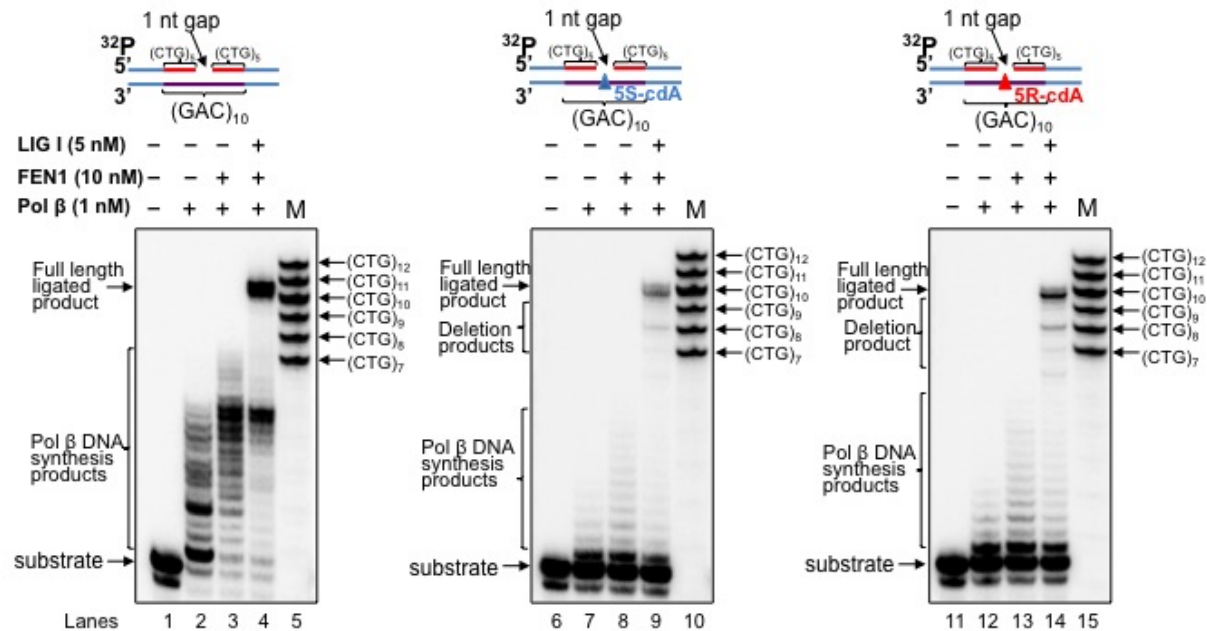
To further examine if pol  $\beta$  bypasses a 5',8-cdA lesion in a large gap and an open template DNA during BER and DNA lagging strand synthesis in the same manner as it did with a single-nucleotide gap, we examined pol  $\beta$  lesion bypass with a substrate containing a 5',8-cdA that was opposite to the middle of a 6-nt gap or a open template substrate that mimic a lagging strand synthesis intermediate (Figure 3.7 and 3.8). The results showed that pol  $\beta$  efficiently bypassed the 5',8-cdA in a large gap by performing multi-nucleotide gap-filling synthesis (Figure 3.7, lanes 7-8 and 12-13). With the substrate containing a dA, pol  $\beta$  performed a processive gap-filling synthesis without pausing (Figure 3.7, lanes 2-3). In the presence of FEN1, pol  $\beta$  lesion bypass synthesis activity was stimulated (Figure 3.7, compare lanes 3, 8 and 13 with lanes 2, 7 and 12). Similarly, pol  $\beta$  bypassed a 5',8-cdA in an open template with a distributive multi-nucleotide insertion (Figure 3.8, lanes 5-9), whereas it performed an efficient processive

DNA synthesis with the open template substrate containing a dA (Figure 3.8, lanes 2-3). The results indicate that pol  $\beta$  can bypass a 5',8-cdA in a large gap and an open template during DNA repair and replication.



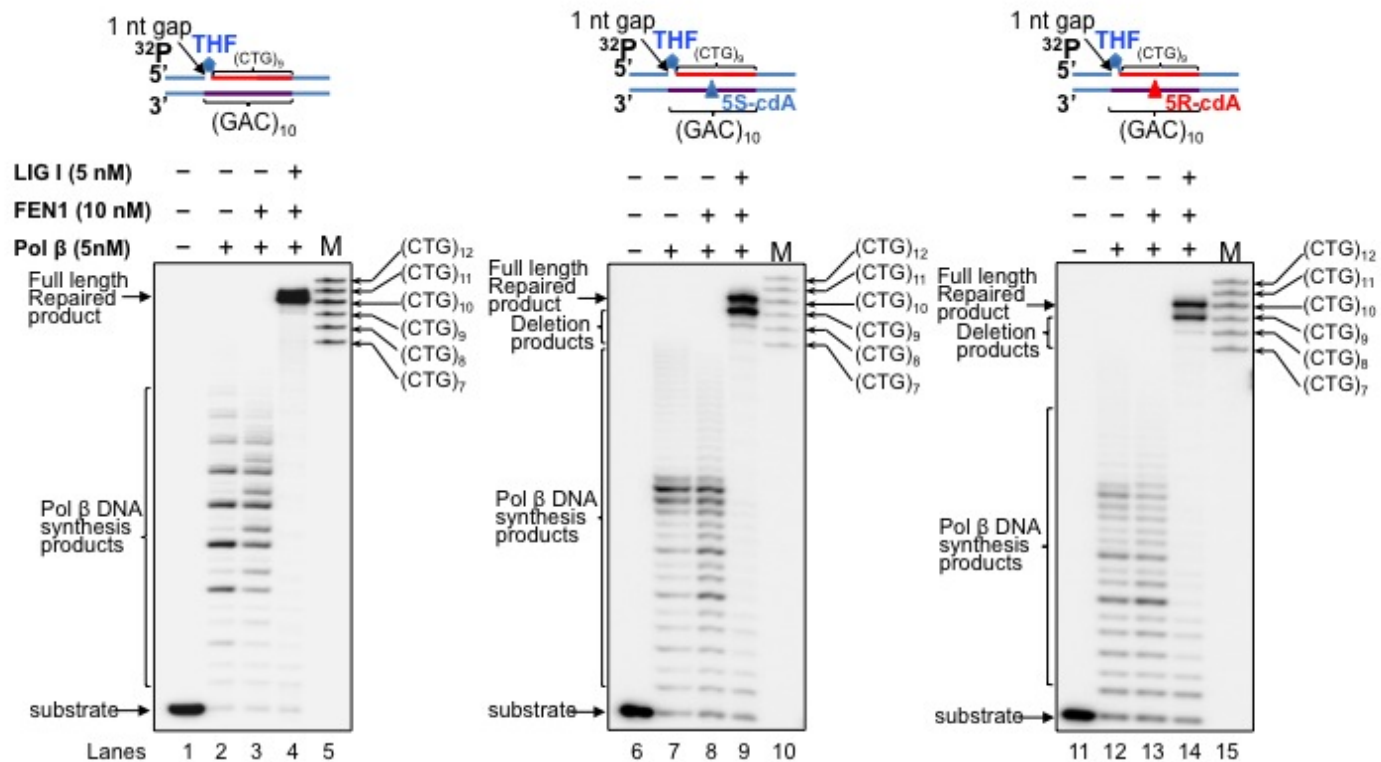
**Figure 3.1 CTG repeat deletion via pol β bypass of a 5',8-cdA in a (CAG)<sub>10</sub> tract that is opposite to a 1-nt gap during BER.**

Substrates with a 5'-phosphorylated THF residue and 1-nt gap opposite to a template dA, or 5'S-cdA, or 5'R-cdA located in the fifth CAG of a (CAG)<sub>10</sub> repeat template were incubated with pol β in the absence or presence of FEN1 or incubated with pol β and FEN1 along with LIG I. Pol β DNA synthesis products and repair products were separated from substrates by urea denaturing gel electrophoresis. Lanes 1, 6 and 11 correspond to substrates only. Lanes 2-3, 7-8 and 12-13 correspond to reaction mixtures with 1 nM pol β in the absence or presence of 10 nM FEN1. Lanes 4, 9 and 14 correspond to reaction mixtures with 1 nM pol β, 10 nM FEN1 and 5 nM LIG I. Lanes 5, 10 and 15 correspond to a series of synthesized size markers (M) for illustrating the size of repaired products. Substrates were <sup>32</sup>P-labeled at the 5'-end of the upstream primer. Substrates are illustrated schematically above the gel.



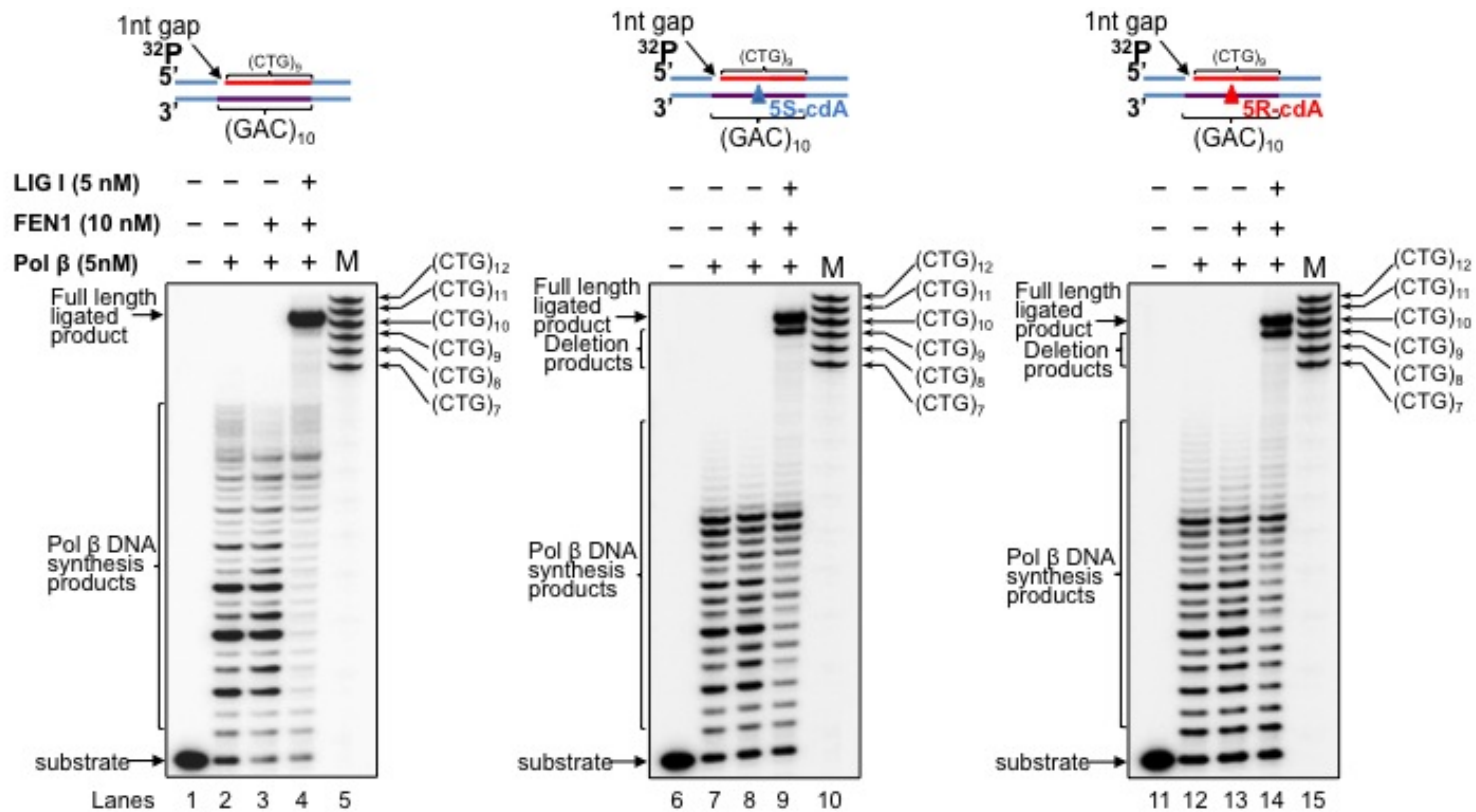
**Figure 3.2 CTG repeat deletion via pol  $\beta$  bypass of a 5',8-cdA in a (CAG)<sub>10</sub> tract that is opposite to a 1-nt gap during DNA replication.**

Pol  $\beta$  bypass of a 1-nt gap opposite to a 5',8-cdA in a (CAG)<sub>10</sub> repeat tract and ligation reaction during Okazaki fragment maturation were examined with the substrates containing a 1-nt gap opposite to a template dA, or 5'S-cdA or 5'R-cdA located in the fifth CAG of (CAG)<sub>10</sub> repeat and a 5'-phosphate in the downstream primer. Lanes 1, 6 and 11 correspond to substrates only. Lanes 2-3, 7-8 and 12-13 correspond to reaction mixtures with 1 nM pol  $\beta$  in the presence or absence of 10 nM FEN1. Lanes 4, 9 and 14 correspond to reaction mixtures with 1 nM pol  $\beta$ , 10 nM FEN1 and 5 nM LIG I. Lanes 5, 10 and 15 correspond to a series of synthesized size markers (M) for illustrating the size of ligated products. Substrates were <sup>32</sup>P-labeled at the 5'-end of the upstream primer. Substrates are illustrated schematically above the gel.



**Figure 3.3 Pol β bypass of 5',8-cdA in a (CAG)<sub>10</sub> tract at downstream of a 1-nt gap during BER.**

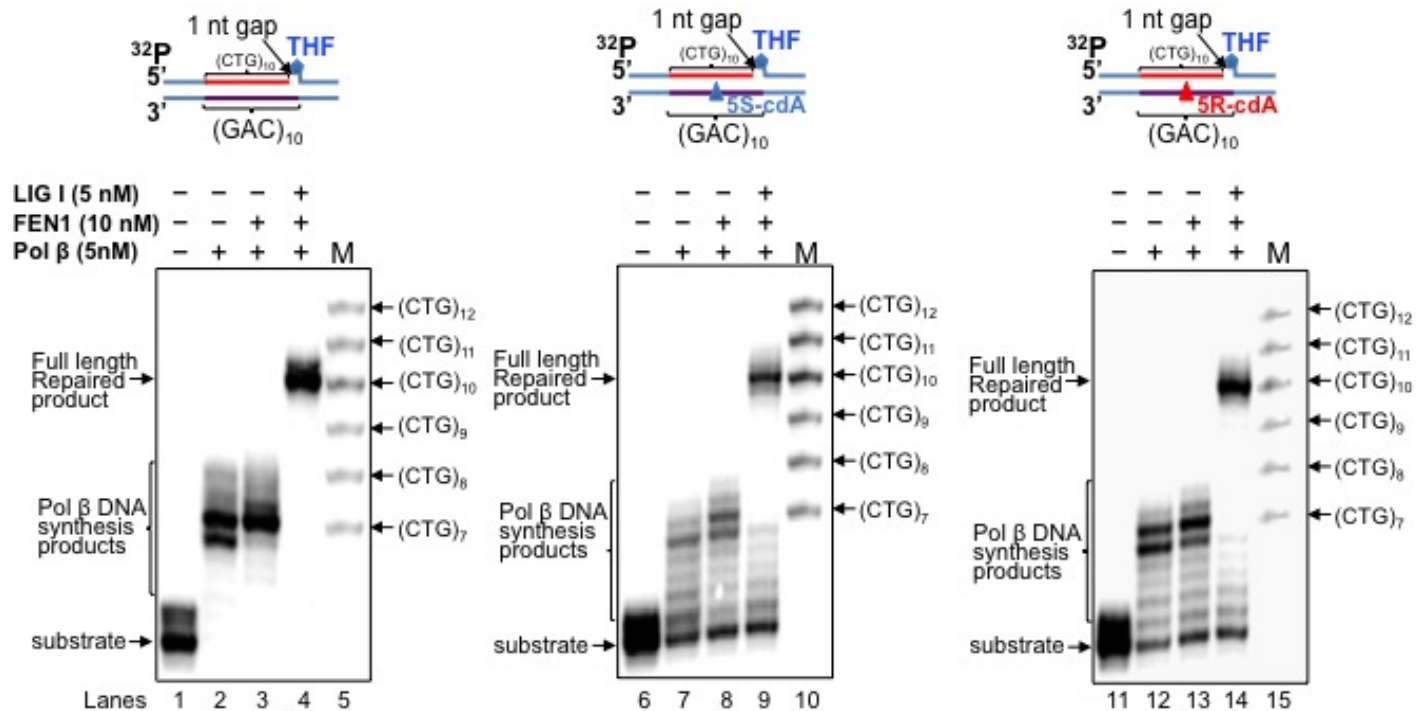
Pol β lesion bypass synthesis and BER of an abasic site located at the upstream of a template 5',8-cdA in (CAG)<sub>10</sub> repeats were examined with the substrates containing a 1-nt gap upstream of a template dA, or 5'S-cdA, or 5'R-cdA located in the fifth CAG of (CAG)<sub>10</sub> repeats, and a 5'-phosphate THF in the downstream primer. Lanes 1, 6 and 11 correspond to substrate only. Lanes 2-3, 7-8 and 12-13 correspond to reaction mixtures with 1 nM pol β in the presence or absence of 10 nM FEN1. Lanes 4, 9 and 14 correspond to reaction mixtures with 1 nM pol β, 10 nM FEN1 and 5 nM LIG I. Lanes 5, 10 and 15 correspond to a series of synthesized size markers (M) for illustrating the size of repaired products. Substrates were <sup>32</sup>P-labeled at the 5'-end of their damaged strands and are illustrated schematically above the gel.



**Figure 3.4 Pol β bypass of 5',8-cdA in a (CAG)<sub>10</sub> tract at downstream of a 1-nt gap during DNA replication.**

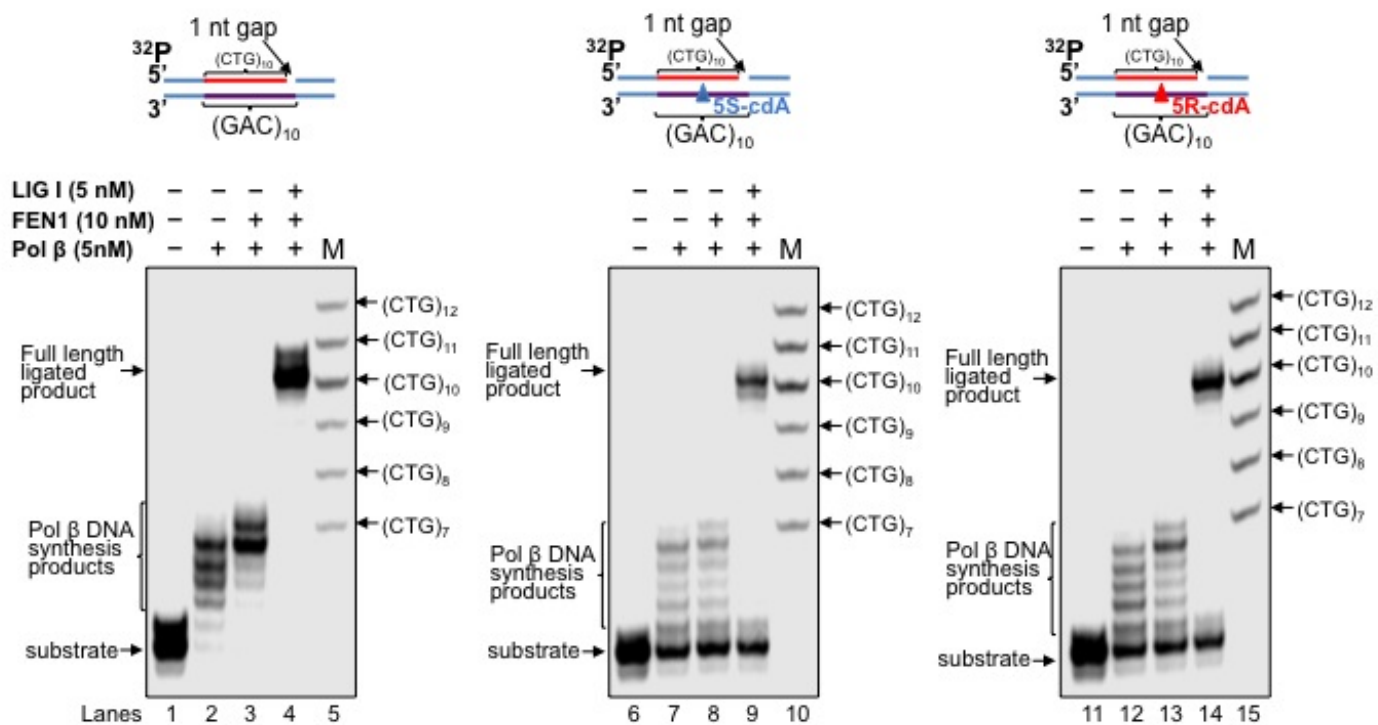
Pol β bypass of a 1-nt gap located at the upstream of a template 5',8-cdA in a (CAG)<sub>10</sub> repeat tract and ligation reaction during Okazaki fragment maturation was examined with the substrates containing a 1-nt gap upstream of a dA, or 5'S-cdA, or 5'R-cdA located at the fifth CAG of the repeat template and a 5'-phosphate in the downstream primer. Lanes 1, 6 and 11 correspond to substrates only. Lanes 2-3, 7-8 and 12-13 correspond to reaction mixtures with 1 nM pol β in the absence or presence of 10 nM FEN1. Lanes 4, 9 and 14 correspond to reaction mixtures with 1 nM pol β, 10 nM FEN1 and 5 nM LIG I. Lanes 5, 10 and 15 correspond to a series of synthesized size markers (M) for illustrating the size of ligated products. Substrates were <sup>32</sup>P-labeled at the 5'-end of their damaged strands and are illustrated schematically above the gel.





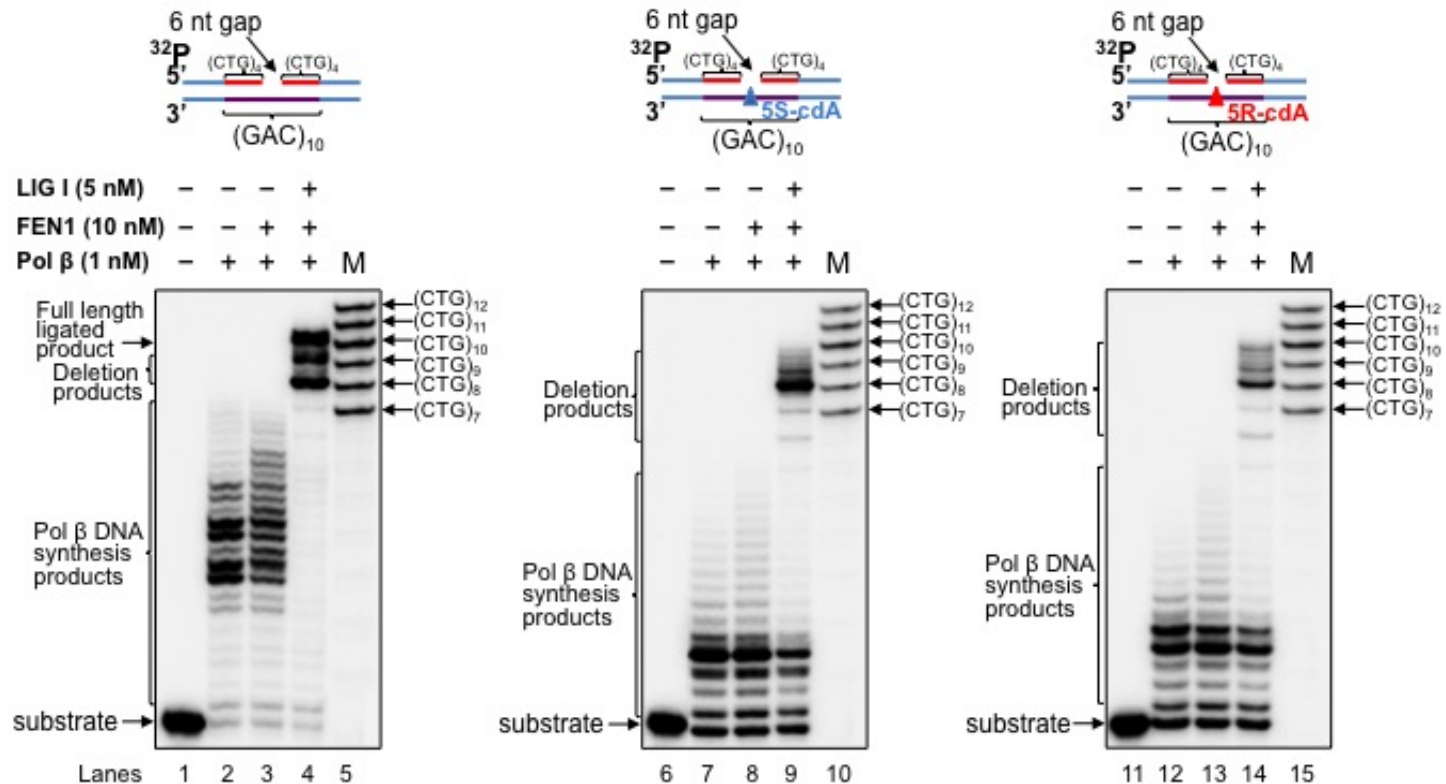
**Figure 3.5 Pol  $\beta$  bypass of 5',8-cdA in a (CAG)<sub>10</sub> tract at upstream of a 1-nt gap during BER.**

Pol  $\beta$  lesion bypass synthesis and BER of an abasic site located at the downstream of a template 5',8-cdA in (CAG)<sub>10</sub> repeats were examined with the substrates containing a 1-nt gap downstream of a template dA, or 5'S-cdA, or 5'R-cdA located in the fifth CAG of the CAG repeat template and a 5'-phosphate THF in the downstream primer. Lanes 1, 6 and 11 correspond to substrate only. Lanes 2-3, 7-8 and 12-13 correspond to reaction mixtures with 1 nM pol  $\beta$  in the absence or presence of 10 nM FEN1. Lanes 4, 9 and 14 correspond to reaction mixtures with 1 nM pol  $\beta$ , 10 nM FEN1 and 5 nM LIG I. Lanes 5, 10 and 15 correspond to a series of synthesized size markers (M) for illustrating the size of repaired products. Substrates were <sup>32</sup>P-labeled at the 5'-end of their damaged strands and are illustrated schematically above the gel.



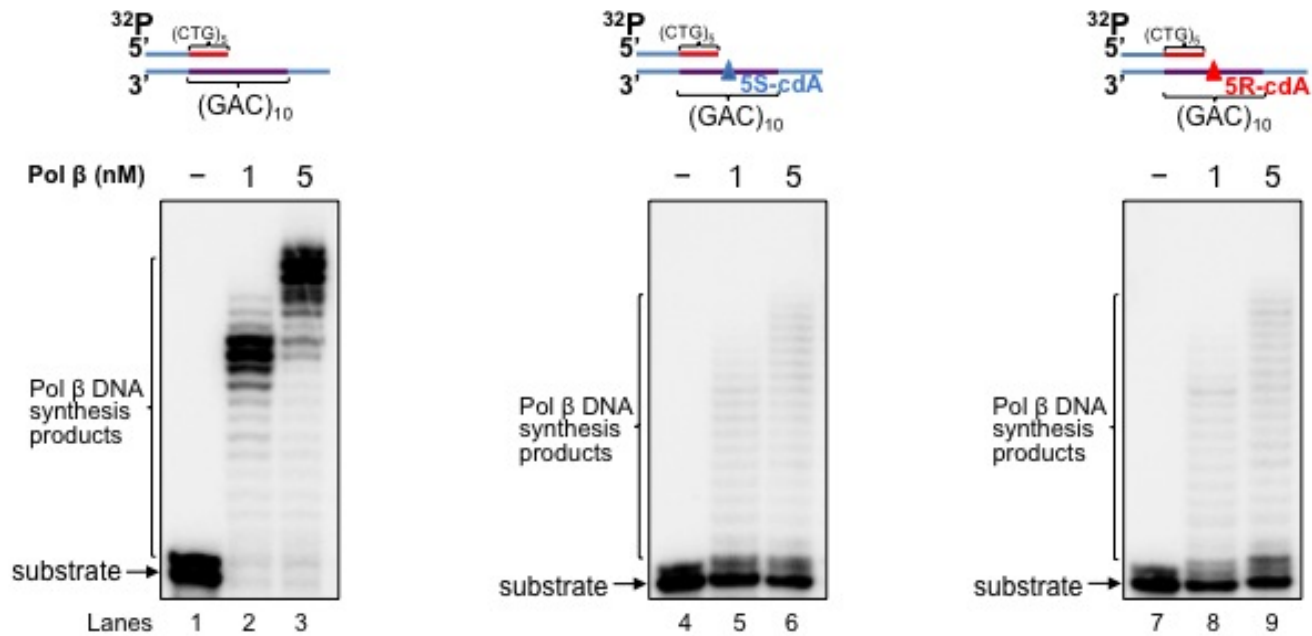
**Figure 3.6 Pol  $\beta$  bypass of 5',8-cdA in a (CAG)<sub>10</sub> tract at upstream of a 1-nt gap during DNA replication.**

Pol  $\beta$  lesion bypass of a 5',8-cdA at the downstream of a 1-nt gap located in the (CAG)<sub>10</sub> repeat template strand and ligation reaction during Okazaki fragment maturation, were examined with the substrates with a 1-nt gap downstream of a template dA, or 5'S-cdA, or 5'R-cdA located at the fifth CAG of the (CAG)<sub>10</sub> repeats and a 5'-phosphate in the downstream primer. Lanes 1, 6 and 11 correspond to substrates only. Lanes 2-3, 7-8 and 12-13 correspond to reaction mixtures with 1 nM pol  $\beta$  in the absence or presence of 10 nM FEN1. Lanes 4, 9 and 14 correspond to reaction mixtures with 1 nM pol  $\beta$ , 10 nM FEN1 and 5 nM LIG I. Lanes 5, 10 and 15 correspond to a series of synthesized size markers (M) for illustrating the size of ligated products. Substrates were <sup>32</sup>P-labeled at the 5'-end of their damaged strands and are illustrated schematically above the gel.



**Figure 3.7 Pol β bypass of a 5',8-cdA located in a (CAG)<sub>10</sub> tract opposite to a 6-nt gap during DNA replication.**

Pol β bypass of 5',8-cdA opposite to a 6-nt gap in a (CAG)<sub>10</sub> repeat tract and ligation reaction during Okazaki fragment maturation were examined with the substrates containing a 6-nt gap along with a 5'-phosphate opposite to a template dA, or 5'S-cdA, or 5'R-cdA located in the fifth CAG of the (CAG)<sub>10</sub>-containing template strand. Lanes 1, 6 and 11 correspond to substrate only. Lanes 2-3, 7-8 and 12-13 correspond to reaction mixtures with 1 nM pol β in the absence or presence of 10 nM FEN1. Lanes 4, 9 and 14 correspond to reaction mixtures with 1 nM pol β, 10 nM FEN1 and 5 nM LIG I. Lanes 5, 10 and 15 correspond to a series of synthesized size markers (M) for illustrating the size of ligated products. Substrates were <sup>32</sup>P-labeled at the 5'-end of their damaged strands and are illustrated schematically above the gel.



**Figure 3.8 Pol  $\beta$  bypass of a 5',8-cdA located in a (CAG)<sub>10</sub> tract in the open template during DNA replication.**

Pol  $\beta$  bypass of a 5',8-cdA on an open template during DNA replication was examined with the substrates with a dA, or 5'S-cdA, or 5'R-cdA in a (CAG)<sub>10</sub>-containing template strand and an upstream primer annealed at the upstream of the lesion. Lanes 1, 4 and 7 correspond to substrate only. Lanes 2-3, 5-6 and 8-9 correspond to reaction mixtures with 1 nM and 5 nM pol  $\beta$ . Substrates were <sup>32</sup>P-labeled at the 5'-end of their damaged strands and are illustrated schematically above the gel.

## **Pol $\beta$ bypass of a 5',8-cdA in the CAG repeat tract results in CTG repeat deletion during BER and DNA replication**

Previous study showed that pol  $\beta$  bypass of a template AP site opposite to a gap can result in DNA slippage and upstream primer realignment, resulting in DNA sequence expansion or deletion (268). We reason that pol  $\beta$  bypass of a template 5',8-cdA may also induce TNR instability. To test this, we reconstituted BER or DNA lagging strand maturation on the substrates with or without a template 5',8-cdA in a (CAG)<sub>10</sub> tract to determine the length change of during BER or Okazaki fragment maturation.. Reconstituted BER was initially constructed with the substrates that mimic a APE1 pre-cut BER intermediate with a 1-nt gap opposite to a template normal A or 5',8-cdA (Figure 3.1, lanes 4, 9 and 14). The results showed that pol  $\beta$  directly bypassed the 5',8-cdA in a (CAG)<sub>10</sub> tract during BER leading to the formation of the full length repaired product as well as a series of CTG repeat deletion products (Figure 3.1, lanes 9 and 14). However, pol  $\beta$  bypass of a dA in the template strand only resulted in the formation of the full length repaired product (Figure 3.1, lane 4). Ligation products resulting from pol  $\beta$  bypass of a 5',8-cdA in the repeat tract during Okazaki fragment maturation was determined with substrates containing a 1-nt gap opposite to the lesion. The results showed that pol  $\beta$  lesion bypass resulted in the full length ligation product that has the same length as the template as well as a series of CTG repeat deletion products. However, pol  $\beta$  DNA synthesis on the template with a dA only led to the production of the full length ligation product (Figure 3.2, lanes 4, 9 and 14).

To further determine whether the position of a 1-nt gap relative to the lesion can also affect repeat length change, reconstituted BER and lagging strand DNA maturation

were also performed with the substrates containing a 1-nt gap located at the upstream or downstream of a template dA or 5',8-cdA (Figure 3.3-3.6, lanes 4, 9 and 14). With the substrate containing a 1-nt gap located at the upstream of a template dA, pol  $\beta$  lesion bypass synthesis only resulted in the formation of the full length repaired or ligated product (Figure 3.3 and 3.4, lane 4). However, with the substrates containing a template 5',8-cdA, pol  $\beta$  lesion bypass led to both the full length product as well as deletion products mainly with one CTG deletion (Figure 3.3 and 3.4, lanes 9 and 14). Pol  $\beta$  DNA synthesis on the substrates with a 1-nt gap located at the downstream of a template dA or 5',8-cdA in a (CAG)<sub>10</sub> repeat tract resulted in the formation of the full length repaired or ligated product (Figure 3.5 and 3.6, lanes 4, 9 and 14). This indicates that CTG repeat deletion can be induced by pol  $\beta$  lesion bypass at the upstream or opposite to a template 5',8-cdA in a (CAG)<sub>10</sub> tract, but not at the downstream of the lesion during BER and Okazaki fragment maturation.

To further determine whether the size of a gap opposite to a template 5',8-cdA can modulate the instability of repeats mediated by pol  $\beta$  lesion bypass synthesis during Okazaki fragment maturation, we reconstituted lagging strand maturation with the substrates containing a 6-nt gap and determined the effects the large gap on repeat length change through pol  $\beta$  bypass of a template 5',8-cdA. The results showed that pol  $\beta$  lesion bypass mainly resulted in a ligated product with 2 repeats deletion and a series of products with 3-4 repeat deletion (Figure 3.7, lanes 9 and 14). Interestingly, pol  $\beta$  synthesis on the substrate with a template dA opposite to the middle of a 6 nt-gap resulted in production of the full length ligation product as well as ligation products with 1 and 2 repeat deletion (Figure 3.7, lane 4). This indicates that a large gap promoted CTG repeat

deletion during pol  $\beta$  bypass of a template 5',8-cdA. Our results indicate that the position of pol  $\beta$  lesion bypass synthesis relative to a template 5',8-cdA in a (CAG)<sub>10</sub> tract governs production of CTG repeat deletion during BER and DNA replication, i.e. pol  $\beta$  lesion bypass synthesis that occurs at the upstream or opposite to a 5',8-cdA damage, results in CTG repeat deletion, whereas its lesion bypass at the downstream of a 5',8-cdA does not alter CTG repeat length. The results further suggest that a template 5',8-cdA can induce looping out of the template CAG repeat to form a bubble during pol  $\beta$  bypass of the damage, thereby promoting pol  $\beta$  to skip over the bubble formed on the template strand and promote CTG repeat deletion. The results also suggest that a large gap allows the formation of a large bubble in the template strand that in turn promotes pol  $\beta$  to skip over more CAG repeats than a small gap, inducing a relatively large CTG repeat deletion. To test this, we then examined whether a 5',8-cdA on the template strand can induce the formation of bubble structures.

### **A template 5',8-cyclo-dA in the CAG repeat tract induces the formation of different sizes of bubble structures**

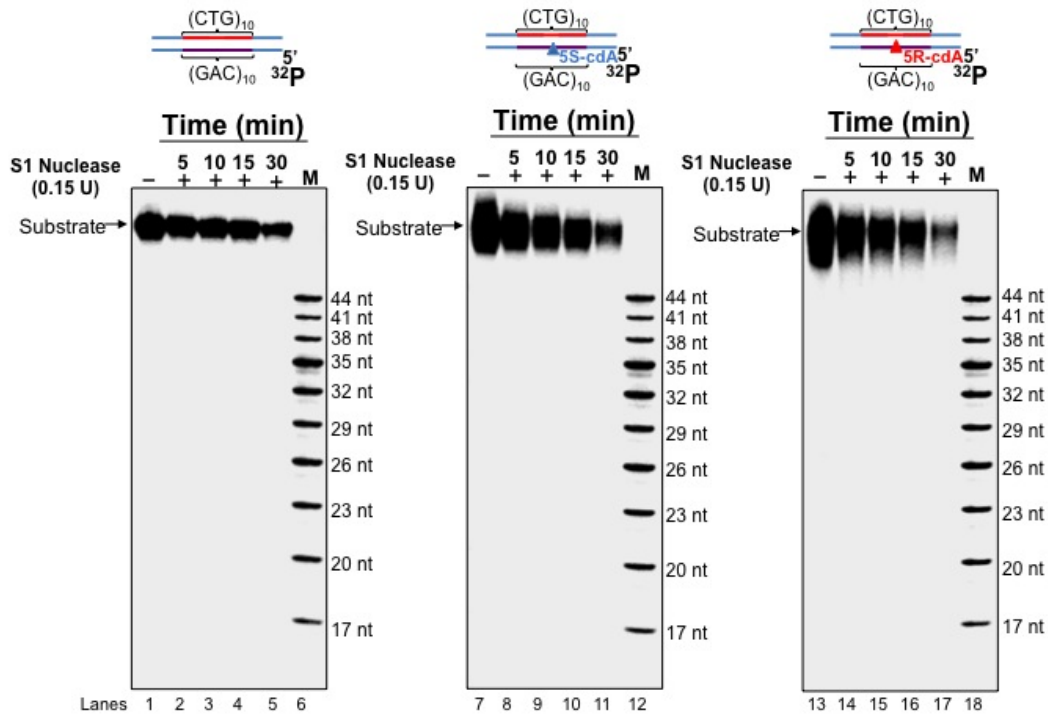
Previous studies showed that during BER of oxidative DNA damage in a TNR tract, both of the damaged and template strand can form different sizes of hairpin structures(200,257). The hairpin formed on the damaged strand leads to TNR expansion and a template hairpin results in TNR deletion (200,257). Interestingly, it has been found that formation of cyclo-ring linking the C5' of the deoxyribose and the C8 of a purine results in a sugar pucker that is energetically unfavorable. This can then further induce DNA backbone distortion causing over twisting of double-strand DNA (239). This

distortion in a created by a cyclo-deoxypurine in a TNR tract may promote the formation of secondary structures such as hairpins. Thus, the presence of a 5',8-cdA in TNR tract could facilitate a hairpin/bubble formation during pol  $\beta$  bypass of the damage and promote pol  $\beta$  skip over the template hairpin/bubble, thereby resulting in TNR deletion. To test this possibility, we initially examined the formation of hairpin/bubble in an intact double-strand DNA using S1 nuclease, a single strand DNA and RNA specific nuclease (269) (Figure 3.9-3.13). The results showed that no S1 nuclease cleavage products were detected with double-strand DNA substrates with a template dA or a 5',8-cdA (Figure 3.9, lanes 2-5, 8-11, 14-17). This indicated that in the context of intact double-strand DNA, 5',8-cdA failed to induce the formation of a hairpin/loop structure. To further determine whether a 5',8-cdA may induce the formation of the hairpin/bubble in the substrates with a 1-nt gap opposite or at the upstream of the lesion, we initially examined S1 Nuclease digestion of the template strand with a dA or a 5',8-cdA in the absence of pol  $\beta$ . For the substrates containing a 1-nt gap opposite to a dA, S1 cleavage resulted in the products with 28 nt and 29 nt (Figure 3.10, lanes 2-6) indicating that the enzyme cleavage occurred opposite to the 1-nt gap and no hairpin/bubble formed in the template strand. However, for the template strand with a 5'S-cdA or 5'R-cdA, the enzyme resulted in the cleavage products with 25 nt to 34 nt (Figure 3.10, lanes 9-13 and 18-22). The results indicated that a (CAG)<sub>3</sub> repeat bubble opposite to the 1-nt gap formed (Figure 3.10, the panels in the middle and on the right). To further examine whether pol  $\beta$  lesion bypass synthesis of a template 5',8-cdA can promote the formation of a CAG repeat hairpin/bubble, we probed bubble structures that may form on the same set of substrates in the presence of 1 nM of pol  $\beta$ . S1 nuclease resulted in the same sizes of products on the



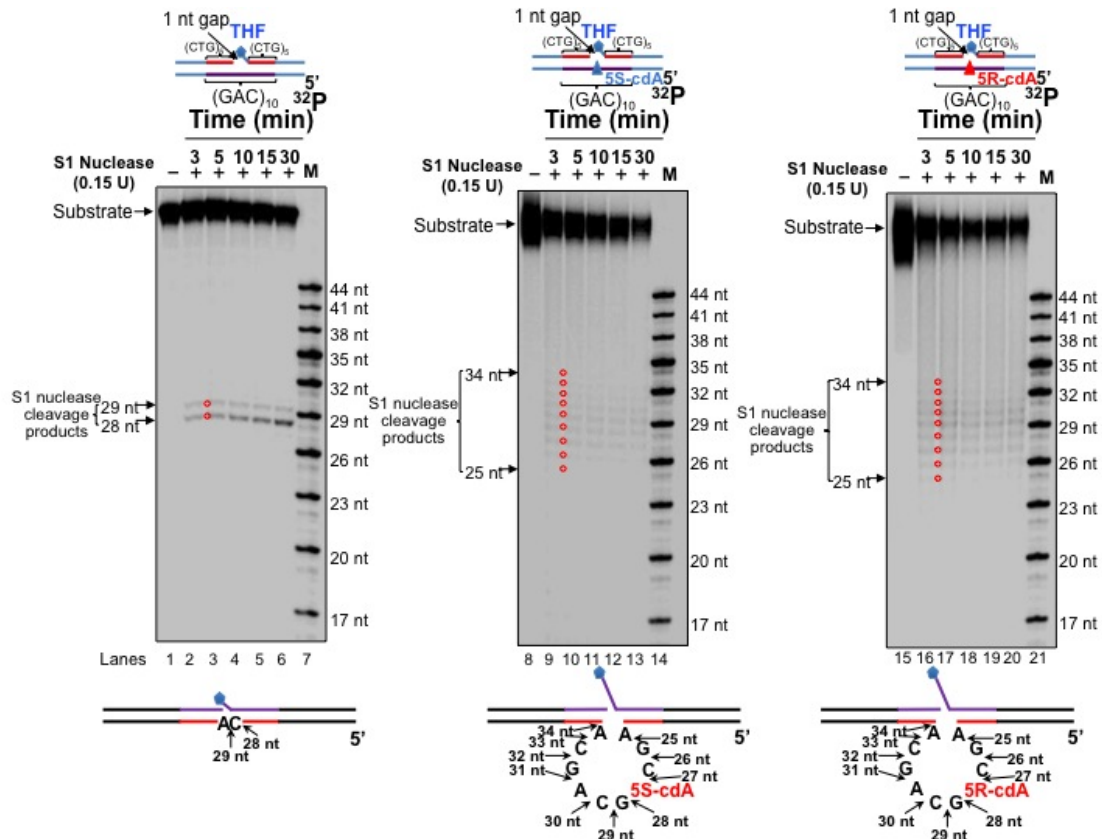
substrates with or without a template base lesion as it did in the absence of pol  $\beta$  (Figure 3.11). This also indicated the formation of a (CAG)<sub>3</sub> repeat bubble on the template containing a 5'S- or 5'R-dcA lesion in presence of pol  $\beta$  during BER (Figure 3.11). The results demonstrated that the bubbles formed in the template strand were solely induced by a 5',8-cdA lesion, and pol  $\beta$  lesion bypass synthesis failed to affect the formation of CAG repeat bubbles in the template strand. The results further indicated that CAG repeat bubbles only formed in the presence of single-stranded DNA region. To examine if the position of a 1nt-gap may affect the formation of a CAG repeat bubble induced by a 5',8-cdA lesion during BER, we determined the bubble formation with the substrates containing a 1-nt gap at 6 nt of the upstream of a template dA or a 5',8-cdA in the absence and presence of pol  $\beta$  (Figure 3.12 and Figure 3.13). We found that S1 Nuclease cleavage on the substrate containing a 1-nt gap located at the upstream of a template dA only generated a series of higher molecular weight cleavage products indicating S1 cleavages on the 3'-side of the template strand that were not fully annealed with the upstream strand in the absence of pol  $\beta$  (Figure 3.12, left panel, lanes 2-6). S1 nuclease cleavage on the substrates with a template 5'S-cdA or 5'R-cdA in the absence of pol  $\beta$  led to products with 21 nt to 27 nt (Figure 3.12, lanes 9-11) and products with 23 nt to 30 nt (Figure 3.12, lanes 16-18), respectively indicating the formation of a loop with (CAG)<sub>3</sub> repeats. S1 nuclease cleavage on the substrates with a template 5'S-cdA or 5'R-cdA in the presence of pol  $\beta$  resulted in products with 22 nt to 29 nt and products with 24 nt to 34 nt (Figure 3.13, lanes 9-11 and lanes 16-18) indicating the formation of a (CAG)<sub>3</sub> and (CAG)<sub>4</sub> loop, respectively. Similar to its cleavage in the absence of pol  $\beta$ , S1 nuclease only made cleavages at the 3'-side of the template strand of the substrate containing a

template dA to generate a series of high molecular weight products in the presence of 1 nM pol  $\beta$  (Figure 3.13, lanes 2-4). These results indicated that a template 5',8-cdA induced the formation of a bubble that was skipped over by pol  $\beta$ . Thus, our results demonstrate that a 5',8-induced bubble formed whenever single-strand DNA breakage occurs, and the bubble can then be skipped over by pol  $\beta$  via DNA synthesis.



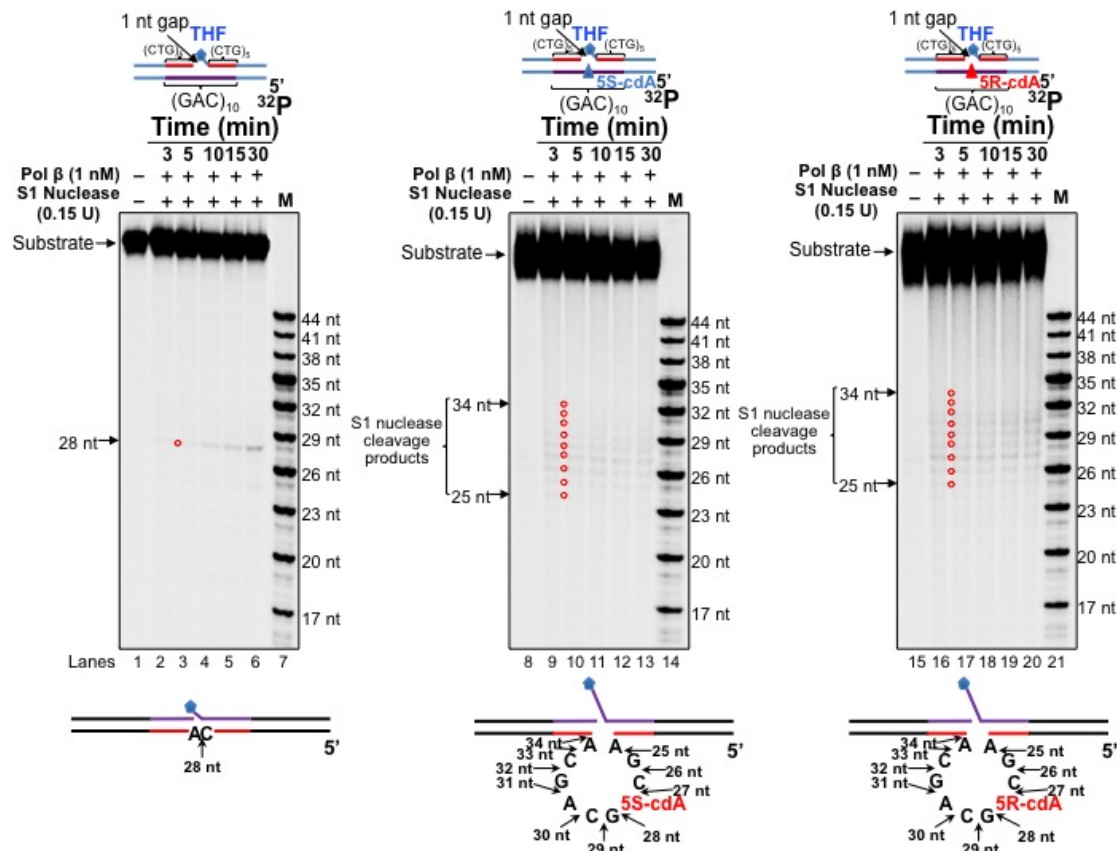
**Figure 3.9 Formation of a loop on the template strand of an intact double-strand DNA.**

Formation of a loop on the template strand of an intact double-strand DNA was probed by S1 Nuclease digestion with substrates containing a dA (left panel) or 5'S-cdA (middle panel) or 5'R-cdA (right panel) in a (CAG)<sub>10</sub> repeat tract of the template strand. The substrates were radiolabeled at the 5'-end of their template strands. Subsequently, substrates were incubated with 0.15 units of S1 Nuclease at 5-, 10-, 15-, and 30-minute time intervals (lanes 2-5, 8-11 and 14-17). Lanes 1, 7 and 13 represent the undigested substrate. Lanes 6, 12 and 18 represent synthesized size markers (M) with 17, 20, 23, 26, 29, 32, 35, 38, 41, and 44 nucleotides, respectively. For all the experiments, 25 nM of substrate was used. The substrates are illustrated schematically above the gel.



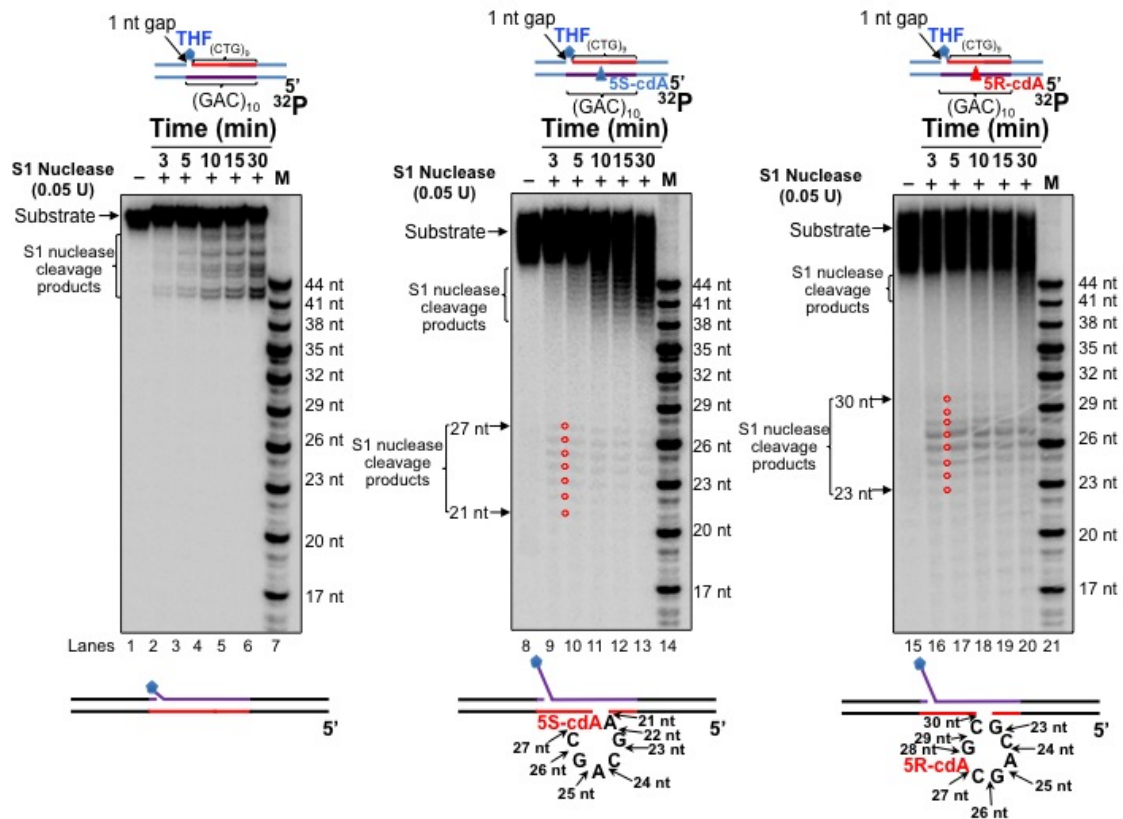
**Figure 3.10 Formation of a loop on the template with a 5',8-cdA lesion opposite to 1-nt gap in the absence of pol  $\beta$ .**

Formation of a loop on the template with a 5',8-cdA lesion opposite to 1-nt gap with a downstream 5'-THF residue was probed by S1 Nuclease with substrates containing a dA (left panel), or 5'S-cdA (middle panel) or 5'R-cdA (right panel) in the (CAG)<sub>10</sub> repeats located in the template strand. Substrates were radiolabeled at the 5'-end of the template strand and incubated with S1 Nuclease in the absence of pol  $\beta$ . For all the experiments, 25 nM of substrate was used. Arrows and circles indicate the major S1 cleavage products. The substrates are illustrated schematically above the gel. A loop deduced by a specific nuclease cleavage pattern and the nuclease digestion sites is illustrated schematically below the gel.



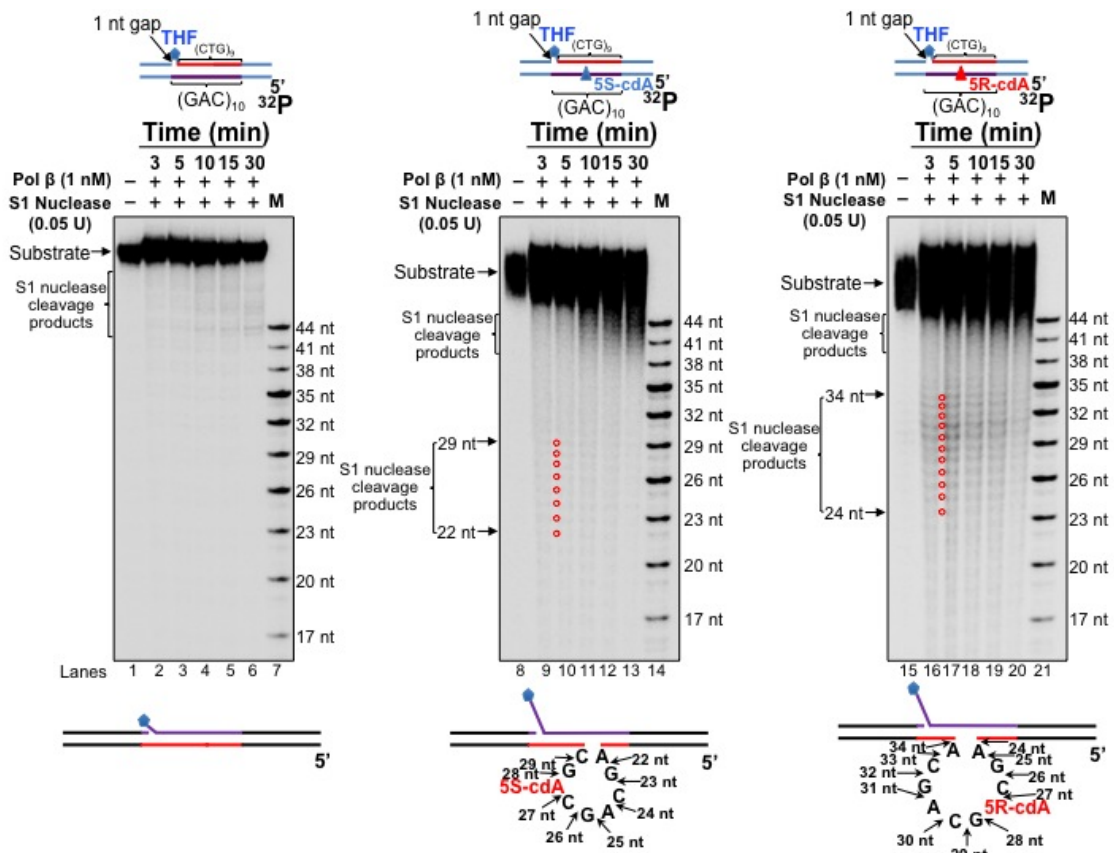
**Figure 3.11** Formation of a loop on the template with a 5',8-cdA lesion opposite to 1-nt gap in the presence of pol  $\beta$ .

Substrates containing a dA, or 5'S-cdA or 5'R-cdA that is opposite to a 1-nt gap were pre-incubated with 1 nM pol  $\beta$  for 30 min at 37 °C and subsequently incubated with S1 Nuclease. Substrates were incubated with 0.15 units of S1 Nuclease at 3-, 5-, 10-, 15-, and 30-minute time intervals (lanes 2-6, 9-13 and 16-20). Lanes 1, 8 and 15 represent the undigested substrate. Lanes 7, 14 and 21 represent synthesized size markers (M). For all the experiments, 25 nM of substrate was used. Arrows and circles indicate the major S1 cleavage products. The substrates are illustrated schematically above the gel. A loop deduced by a specific nuclease cleavage pattern and the nuclease digestion sites is illustrated schematically below the gel.



**Figure 3.12 Formation of a loop in the template strand of the substrates with a 1-nt gap located at the upstream of a 5',8-cdA in the absence of pol  $\beta$ .**

Substrates containing a dA (left panel), 5'S-cdA (middle panel) or 5'R-cdA (right panel) in a (CAG)<sub>10</sub> repeat located at four CAG repeats downstream of a 1-nt gap with a 5'-phosphorylated THF residue were incubated with S1 Nuclease in the absence of pol  $\beta$ . Substrates were radiolabeled at the 5'-end of the template strand. For all the experiments, 25 nM of substrate was used. Arrows and circles indicate the major S1 cleavage products. The substrates are illustrated schematically above the gel. A loop deduced by a specific nuclease cleavage pattern and the nuclease digestion sites is illustrated schematically below the gel.



**Figure 3.13 Formation of a loop in the template strand of the substrates with a 1-nt gap located at the upstream of a 5',8-cdA in the presence of pol β.**

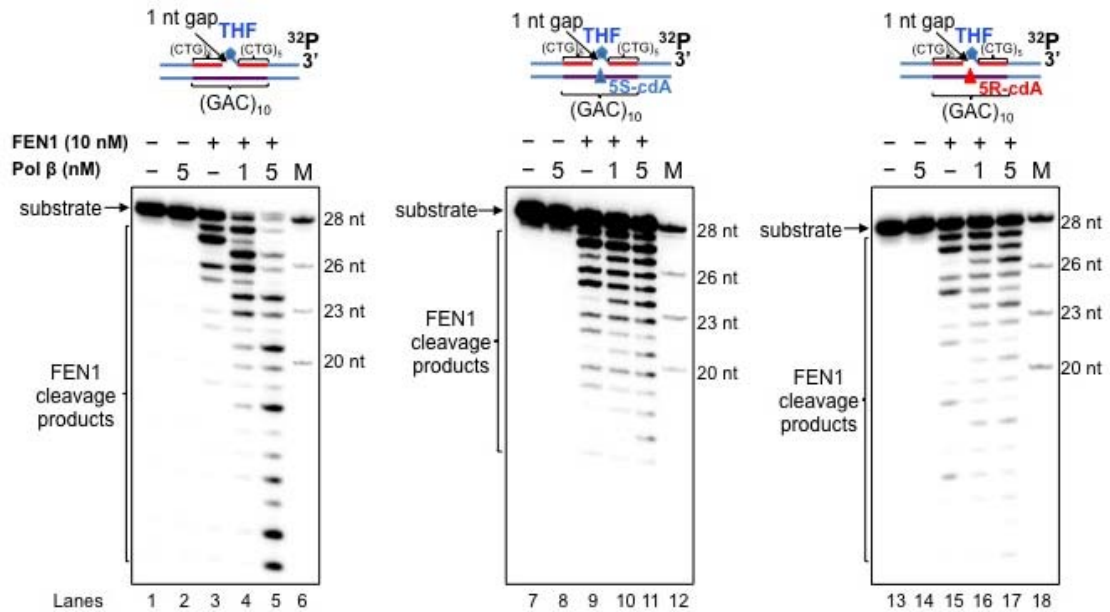
Substrates were pre-incubated with 1 nM pol β for 30 min at 37 °C and subsequently incubated with 0.05 units of S1 Nuclease at 3-, 5-, 10-, 15-, and 30-minute time intervals (lanes 2-6, 9-13 and 16-20). Lanes 1, 8 and 15 represent the undigested substrate. Lanes 7, 14 and 21 represent synthesized size markers (M). For all the experiments, 25 nM of substrate was used. Arrows and circles indicate the major S1 cleavage products. The substrates are illustrated schematically above the gel. A loop deduced by a specific nuclease cleavage pattern and the nuclease digestion sites is illustrated schematically below the gel.

## **A template 5',8-cdA stimulates FEN1 cleavage on CTG repeats promoting repeat deletion**

FEN 1 is a core enzyme in both Okazaki fragment maturation and long-patch BER (270). FEN1 removes a 5'-flap generated by strand displacement synthesis that is mediated by replicative DNA polymerases such as pol  $\delta/\epsilon$  (11). FEN1 is also responsible for removing an oxidized deoxyribose phosphate that is specifically subject to long-patch BER (11). In addition, FEN1 has been shown to be a critical enzyme that prevents trinucleotide repeat expansion (166,192,261) by cleaving a TNR flap during DNA lagging strand synthesis and Okazaki fragment maturation. Moreover, FEN1 can promote the removal of a hairpin to prevent TNR expansion by cleaving a downstream 5'-flap during BER of a base lesion located in the loop region of the hairpin (265). Flap endonuclease 1 flap cleavage activity is readily inhibited by a stable hairpin-containing flap during DNA replication and BER, and this forces the enzyme to perform an alternate flap cleavage that subsequently leads to TNR expansion (264). However, coordination between pol  $\beta$  hairpin bypass synthesis and FEN1 alternate flap cleavage can also result in TNR deletion during BER (257). It is possible that FEN1 coordinates with pol  $\beta$  skip-over of a template 5',8-cdA to remove a downstream flap generated by pol  $\beta$  lesion bypass synthesis, thereby facilitating CTG repeat deletion. To test this possible, we initially determined if FEN1 cleavage on CTG repeats can be affected by a template 5',8-cdA. We examined FEN1 cleavage activity on the substrate with a 1-nt gap opposite or at the upstream of a template dA and 5',8-cdA in the absence of pol  $\beta$  (Figure 3.14 and 3.15, lanes 3, 9 and 15). We found that FEN1 flap cleavage with the substrates containing a template 5',8-cdA generated more cleavage products than its cleavage the substrate

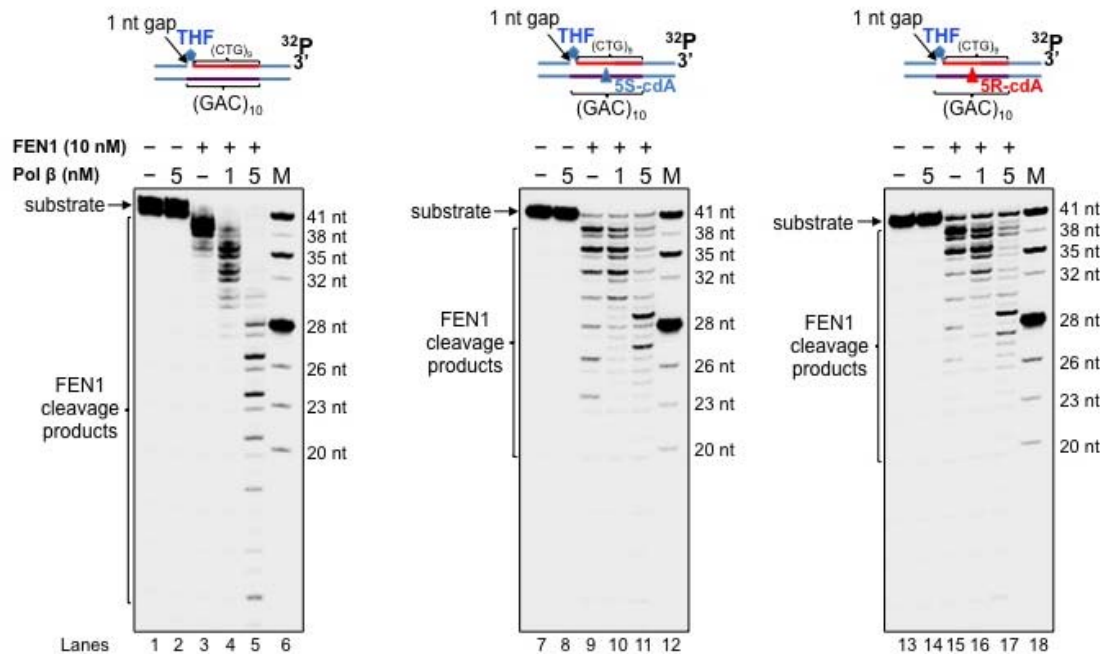


with a template dA (compare lane 9 and 15 with lane 3 of Figure 3.14 and 3.15). This indicates that a template 5',8-cdA stimulated FEN1 flap cleavage. This further suggests that a 5',8-cdA in the CAG repeat template strand can directly induce the formation of a repair or replication intermediate containing a template loop with a long flap, FEN1 then captured the flap and cleaved it efficiently. This further suggests that a template 5',8-cdA induced looping-out of the template strand to produce a CAG repeat bubble, promoting the dissociation of a downstream strand from the template strand that led to a downstream 5'-flap. Surprisingly, a template 5',8-dA failed to exhibit a stimulatory effect on FEN1 flap cleavage with the substrates in the presence of pol  $\beta$  DNA synthesis (compare lanes 10-11 and lanes 16-17 with lanes 4-5 of Figure 3.14 and 3.15). This is because pol  $\beta$  (1 nM or 5 nM) stimulated FEN1 cleavage on the substrates with a template dA (lanes 4-5 of Figure 3.14 and 3.15). The stimulation was increased with increasing concentrations of pol  $\beta$  (compare lane 5 with lane 4 of Figure 3.14 and 3.15) indicating that pol  $\beta$  DNA synthesis created a CTG repeat flap that was efficiently cleaved by FEN1. In contrast, pol  $\beta$  DNA synthesis slightly inhibited FEN1 cleavage on the substrates with a template 5',8-cdA (compare lanes 9-10 and lanes 16-17 with lanes 4-5 of Figure 3.14 and 3.15) indicating a competition between pol  $\beta$  and FEN1 to bind a flap intermediate. This indicated that a template 5',8-cdA base lesion induced the formation of a CTG flap prior to FEN1 cleavage. This further indicated that a downstream CTG repeat flap was directly created by a template 5',8-cdA, and pol  $\beta$  lesion bypass synthesis failed to further facilitate a longer downstream flap formation. Thus, our results suggest that a template 5',8-cdA can induce the formation of an intermediate with a CAG repeat bubble in the template strand and a 5'-flap presumably by distorting the backbone DNA.



**Figure 3.14 FEN1 cleavage during pol  $\beta$  bypass of a 1-nt gap opposite to a template 5',8-cdA.**

FEN1 flap cleavage on the downstream 5'-THF containing substrates with a 1-nt gap opposite to a template dA (left panel), 5'S-cdA (middle panel) or 5'R-cdA (right panel) in a (CAG)<sub>10</sub> repeat. Lanes 1, 7 and 13 represent substrates only. Lanes 2, 8 and 14 correspond to reaction mixtures with 5 nM pol  $\beta$ . Lanes 3-5, 9-11 and 15-17 correspond to reaction mixtures with 10 nM FEN1 and increasing concentrations of pol  $\beta$ . Lanes 6, 12 and 18 correspond to a series of synthesized size markers (M). Substrates were <sup>32</sup>P-labeled at the 3'-end of the 5'-THF containing strands and are illustrated schematically above the gel.



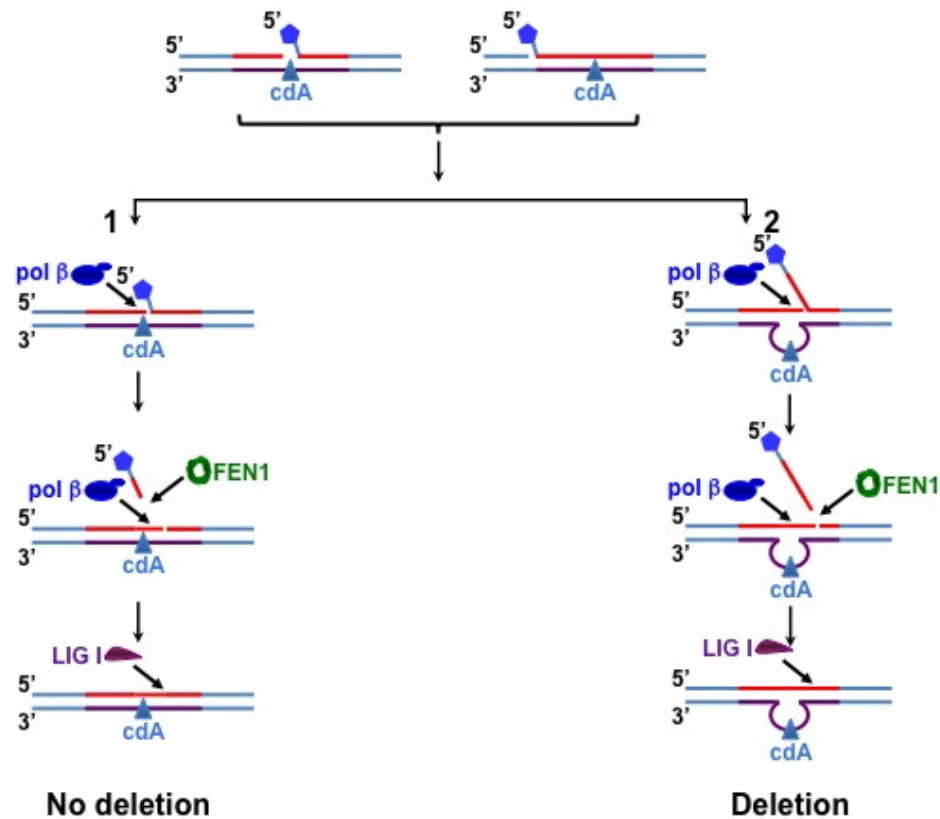
**Figure 3.15 FEN1 cleavage during pol  $\beta$  bypass of a 1-nt gap on the upstream of a template 5',8-cdA.**

FEN1 flap cleavage on the downstream 5'-THF containing substrates with a 1-nt gap at the upstream of a template dA (left panel), 5'S-cdA (middle panel) or 5'R-cdA (right panel) in a (CAG)<sub>10</sub> repeat. Lanes 1, 7 and 13 represent substrates only. Lanes 2, 8 and 14 correspond to reaction mixtures with 5 nM pol  $\beta$ . Lanes 3-5, 9-11 and 15-17 correspond to reaction mixtures with 10 nM FEN1 and increasing concentrations of pol  $\beta$ . Lanes 6, 12 and 18 correspond to a series of synthesized size markers (M). Substrates were <sup>32</sup>P-labeled at the 3'-end of the 5'-THF containing strands and are illustrated schematically above the gel.

## DISCUSSION

In this study, we provide the first evidence that pol  $\beta$  can bypass a template 5',8-cdA located in a CAG repeat tract by skipping over a CAG repeat hairpin loop during BER and Okazaki fragment maturation (Figure 3.1-3.8). This subsequently results in CTG repeat deletion with varying sizes (Figure 3.1-3.7). We demonstrated that pol  $\beta$

failed to bypass a 5'S-cdA located in a random sequence, but managed to bypass a template 5'S-cdA with the same efficiency as it bypassed a 5'R-cdA in a CAG repeat tract (Figures 3.1-3.8, compare the panel in the middle with the one on the right). Further characterization of an alteration in DNA structures indicated that a template 5',8-cdA in a CAG repeat tract induced the formation of a CAG repeat bubble containing the lesion (Figure 3.10 and 3.12). Interestingly, the bubble with the lesion was still sustained after pol  $\beta$  lesion bypass synthesis (Figure 3.13). This indicates that pol  $\beta$  can skip over the bubble to bypass a 5',8-cdA. Furthermore, we found that a template 5',8-cdA also stimulated FEN1 flap cleavage activity (lanes 9 and 15 of Figure 3.14 and 3.15) suggesting the formation of an intermediate with a template hairpin and a 5'-flap that can be efficiently cleaved by FEN1. This further suggests a 5',8-cdA lesion on the template strand also facilitated the dissociation of the downstream strand, thereby subsequently creating an intermediate with a 5'-downstream flap by inducing the formation of a template loop structure. Our data support a model in which different scenarios may occur. In one case, a template 5',8-cdA base lesion in a TNR tract is located opposite or at the downstream of a gap is directly bypassed by pol  $\beta$ , this would lead to a nicked DNA for DNA LIG I to seal, resulting in no repeat deletion product (Figure 3.16, sub-pathway 1). Otherwise, a template 5',8-cyclo-dA can induce the formation of a template loop that is opposite or at the downstream of a 1 nt-gap (Figure 3.16, sub-pathway 2). The template loop promotes dissociation of the downstream strand from the template strand leading to the formation of an intermediate with a template loop structure and a 5'-downstream flap. FEN1 removes the entire downstream 5'-flap leaving a nick for DNA LIG I to seal, resulting in repeat deletion products (Figure 3.16, sub-pathway 2).



**Figure 3.16 CTG repeat deletion during pol  $\beta$  bypass of a template 5',8-cdA.**

A template 5',8-cdA lesion located opposite or at the downstream of a gap in the context of TNR can be directly bypassed by pol  $\beta$  leaving a nick for DNA LIG I to seal. This results in a repaired product that has the same length as the template leading to no repeat deletion (sub-pathway 1). A 5',8-cdA located at the template strand of a TNR tract can result in the formation of TNR loops with varying sizes during BER or DNA lagging strand synthesis. When the lesion is either opposite or at the downstream of a gap (sub-pathway 2), the lesion induces the formation of a loop on the template strand. This subsequently results in the formation of an intermediate containing a template loop and a downstream 5'-flap. This allows pol  $\beta$  to jump over the loop and perform strand-displacement synthesis. Subsequently, FEN1 removes the downstream flap leaving a nick for DNA LIG I to seal resulting in repeat deletion products (sub-pathway 2).

5',8-cyclo-deoxynucleotides can be readily induced by oxidative stress from endogenous, environmental and chemotherapeutic sources (24,234-236). However, they are refractory to DNA repair mechanisms and are inefficiently repaired by nucleotide

excision repair, thereby accumulating in the human genome (239,240). This confers a big challenge for DNA polymerases to bypass the lesion during DNA replication and repair. It has been reported that replicative DNA polymerase, pol  $\delta$  and  $\epsilon$  failed to bypass both 5',8-cdA and 5',8-cdG (240). The lesions can only efficiently be bypassed by translesion DNA polymerases, pol  $\eta$  and pol  $\iota$ . This ultimately leads to replication fork stall and the accumulation of single-strand DNA break intermediates during DNA replication and BER that in turn result in genome instability. However, it remains unknown how this type of base lesions may lead to genome instability. Here we provide the first evidence that a 5',8-cdA lesion can induce TNR deletion by inducing the formation of various sizes of bubbles in the template strand of a TNR repeat tract. This allows pol  $\beta$  to skip over the bubble structures via its lesion bypass synthesis. Furthermore, formation of a loop on the template strand results in the dissociation of the downstream strand from its template strand creating a 5'-flap. This subsequently stimulates FEN1 flap cleavage, thereby promoting TNR deletion. Our study demonstrates a new pathway for TNR deletion mediated by a unique pol  $\beta$  lesion bypass of a 5',8-cdA lesion in a TNR tract during DNA replication and BER.

In this study, we have discovered that 5',8-cdA-induced CTG repeat deletion via pol  $\beta$  lesion bypass during BER and DNA replication can be modulated by the position of a gap relative to the location of a template 5',8-cdA damage in a CAG repeat tract. Our results indicate that a gap that is located at the upstream or opposite to a template 5',8-cdA in a CAG repeat tract can lead to CTG repeat deletion (Figures 3.1-3.4, and 3.7), whereas a gap that is located at the downstream of the lesion leads to sustainment of the length of the repeats (Figure 3.5 and 3.6). Because repeat deletion is mediated by pol  $\beta$

skip-over of a template loop structure and FEN1 cleavage of repeats, this indicates that the formation of a template TNR loop at the downstream of a gap or single-strand break is essential for repeat deletion. This further suggests that a template TNR bubble induced by a 5',8-cdA formed at the downstream of a strand break is critical for pol  $\beta$  to skip over template CAG repeats as well as essential for creating a 5'-TNR flap that can be removed by FEN1, thereby leading to TNR deletion ultimately.

Bypass of a 5',8-cdA lesion in the genome is a big challenge for replicative and translesion polymerases (240,246,248). We have demonstrated that it is also a challenge for pol  $\beta$  lesion bypass. This forces pol  $\beta$  to adopt a unique mechanism to bypass the lesion by its skipping over a loop structure induced by the lesion, causing TNR deletion. Thus, it is conceivable that cofactors that can help pol  $\beta$  to directly bypass a 5',8-cdA lesion may promote the production of undeleted/unexpanded repair or ligation products. It was reported that proliferating cell nuclear antigen (PCNA) and replication protein A (RPA) can stimulate pol  $\beta$  bypass of an AP site during DNA replication that occurs in an open template as well as Okazaki fragment maturation (267). In addition, pol  $\beta$  bypass of an AP site opposite to a 1-nt gap is facilitated by FEN1 and RPA (267). FEN1 and RPA can coordinate with pol  $\beta$  to promote its strand displacement synthesis after pol  $\beta$  bypass a template AP site during Okazaki fragment maturation (4). Thus, it is of interest to determine how possible how PCNA, FEN1 and RPA may help pol  $\beta$  to bypass a 5',8-cdA. This may promote the production of an unexpanded/undeleted products sustaining TNR repeat stability.

Our observation of pol  $\beta$  skip-over various sizes of bubbles induced by a template 5',8-cdA suggests that the stability of the bubble structures is the key to regulate the sizes

of CTG repeat deletion products. Thus, BER and replication cofactors that can modulate the stability of loop structures can modulate the repeat expansion/deletion. For example, mismatch repair proteins MSH2/MSH3 that can bind and stabilize a hairpin/bubbles (113) may facilitate pol  $\beta$  skip over of a TNR bubble on the template strand promoting CTG repeat deletion. In contrast, DNA replication and repair cofactors that can destabilize or disrupt loop structures in the template strand such as DNA helicases may help pol  $\beta$  to directly insert nucleotide to base pair with a template 5',8-cdA, thereby facilitating the formation of a full length repaired or ligation product during DNA replication and BER. For example, Werner syndrome protein and Bloom syndrome protein can unwind a template TNR hairpin/bubble (195,196). Thus it is possible that these helicases can facilitate pol  $\beta$  direct insertion of a dNTP to base pair with the template 5',8-cdA, thereby reducing the TNR deletion products formation.

FEN1 flap cleavage plays an essential role in modulating TNR expansion and deletion. We have shown that a 5',8-cdA base lesion in the context of TNR can directly induce a downstream 5'-flap, and FEN1 removal of the flap creates a nick that is subsequently sealed by DNA ligase I generating deletion products. Thus, proteins that stimulate FEN1 cleavage activity may also promote TNR deletion. For example, PCNA can directly facilitate FEN1 flap cleavage by interacting with FEN1 (232). A 5'-3' exonuclease, Exo I and RPA together with Dna2 can facilitate FEN1 cleavage of a TNR flap by removing a part of a long flap (231,271). These cofactors and enzymes may also play a role in promoting TNR deletion during pol  $\beta$  bypass of a 5',8-cdA in a TNR tract.

Previous studies showed that TNR expansion and deletion occur in a length dependent manner. In this study, we have demonstrated that a 5',8-cdA base lesion



located in a (CAG)<sub>10</sub> repeat can lead to repeat deletion by inducing a small loop in the template strand. It is conceivable that a 5',8-cdA lesion located in a long TNR tract can induce the formation of a more stable TNR loop or hairpin in the template strand. This would facilitate pol  $\beta$  skip-over of the hairpin or loop resulting in larger TNR deletion (257). In addition, since multiple 5',8-cdA may be generated in the same TNR tract and stretches of 5',8-cdA may induce a large distortion on DNA double helix structure inducing the formation of a large loop or hairpin or inducing the formation of multiple loops and hairpins in the template strand. A large TNR loop or hairpin or multiple TNR loops or hairpins in the template strand would result in large TNR deletion by helping pol  $\beta$  to skip over multiple TNR units during its lesion bypass synthesis during DNA replication and BER.

In summary, in this study we have discovered a novel pathway for TNR deletion via pol  $\beta$  bypass of a structurally unusual oxidized DNA base lesion, 5',8-cdA that is located in the template strand of a TNR repeat tract during BER and DNA replication. We have demonstrated that a 5',8-cdA can induce the formation of TNR bubble with varying sizes Pol  $\beta$  can directly bypass a 5',8-cdA lesion by inserting nucleotide to base pair with the lesion leading maintenance of repeat length or by skipping a loop structure formed on the template strand leading to repeat deletion. The formation of a loop on the template strand simultaneously create a 5'-flap that can efficiently cleaved by FEN1. This creates a nick for DNA ligase to seal. Our results reveal that pol  $\beta$  bypass of a 5',8-cdA in a the TNR tract coordinates with FEN1 flap cleavage to mediate TNR deletion. Our work further suggests that a 5',8-cdA-induced TNR deletion during BER and DNA replication is governed by the formation of TNR loops. Pol  $\beta$  bypasses the lesion by skipping over a

loop structure bearing the lesion. In coordinating with FEN1 flap cleavage, this results in repeat deletion.

## SUMMARY

Base excision repair (BER) is a critical DNA repair pathway that maintains genome stability and integrity. The BER pathway is the most efficient repair pathway that removes small DNA base lesions including oxidized base lesions, alkylated base lesions, deaminated base and single-strand breaks. Recent studies have found that oxidative DNA damage and BER are actively involved in modulating trinucleotide repeat (TNR) instability indicating an important role of BER in regulating somatic TNR instability (200,257,262).

TNR expansion has been found as a direct cause of numerous neurodegenerative diseases, and TNR deletion/contraction in androgen receptor is associated with human cancer. Although intensive studies on TNR instability in the past two decades have significantly facilitated our understanding of the molecular mechanisms underlying TNR expansion, however, the mechanisms underlying somatic TNR deletion remain unknown. Because TNR deletion can result in the shortening of expanded TNR tracts, understanding of how somatic TNR deletion is modulated during DNA metabolism is critically important for identifying DNA metabolic proteins as a new target for prevention and treatment of neurodegenerative diseases caused by TNR expansion. To define a mechanism underlying DNA base lesion-induced TNR deletion and the roles of BER to prevent TNR expansion, my Ph.D. dissertation research focuses on exploring TNR deletion induced by oxidative DNA damage and hairpin removal during BER. We characterized TNR instability during BER using a series of oligonucleotide substrates containing oxidative DNA damage in a TNR tract or a TNR hairpin loop and

reconstituted BER reactions using human or mouse cell extracts or purified BER enzymes.

Initially, we demonstrated that TNR deletion can be induced during repair of an oxidized DNA lesion, and this was promoted by an increased level of pol  $\beta$ . Further characterization of secondary structures formed on the damaged and template strand showed that a CAG repeat hairpin formed on both of the damaged and template strand during BER. We showed that pol  $\beta$  bypassed these hairpins in a repeat length dependent manner. With a short repeat that composed an unstable template hairpin, pol  $\beta$  passed through the entire repeat region of the hairpin. This led to maintenance of repeat length. However, with a long repeat tract that composed a stable template hairpin, pol  $\beta$  skipped over several repeat units. In coordination with FEN1 alternate flap cleavage, this resulted in repeat deletion with more repeats removed by FEN1 than those synthesized by pol  $\beta$ . In collaboration with another graduate student in the lab, we have discovered that whether the length of a TNR tract is expanded or deleted/contracted also depends on the position of a DNA base lesion in a TNR tract. This suggests that TNR expansion-induced neurodegenerative diseases can be prevented or treated by modulating the position of an oxidized base lesion in a TNR tract as well as altering the activities of core BER enzymes, pol  $\beta$  and FEN1.

Since oxidative stress attacks the single-stranded hairpin loop region of a hairpin more efficiently than its double-stranded region, and oxidized base lesions are energetically more favorable to be located at a hairpin loop, the efficiency and consequences resulting from BER of an 8-oxoG and oxidized sugar located in a small or large hairpin loop region need to be elucidated. In the second part of my dissertation

research, a mechanism for repairing an oxidized base lesion in a hairpin loop and its effects on TNR expansion was explored. First, we discovered that BER of an 8-oxoG in a CAG repeat hairpin loop was readily initiated by OGG1, and APE1 efficiently incised the 5'-end of an abasic site located in the loop region of both a small and large CAG hairpin resulting in a single strand break in the loop. Subsequently, the incised hairpin is then converted into an intermediate with a 3'- and 5'-flap. This was further confirmed by the result showing that SP-BER of an 8-oxoG in the hairpin loop failed to produce any repair product because of the disruption of the hairpin stem and the formation of a double-flap intermediate with a 3'- and 5'-flap. The results further suggests that the double-flap intermediate can be cleaved by a 3'-5' endonuclease, Mus81/Eme1 and a 5'-3' endonuclease, FEN1, and this could lead to the removal of the hairpin attenuating CAG repeat expansion. The notion was supported by our results showing that with a small hairpin, the presence of either Mus81/Eme1 or FEN1 alone efficiently removed the entire hairpin resulting in the unexpanded repaired product with the same length as the template strand. However, with a large CAG hairpin, coordination among Mus81/Eme1, FEN1 and pol  $\beta$  was required for removal of the entirety of the hairpin. In addition, BER in the loop region of a large hairpin resulted in a series of expanded repair products that are shorter than the expanded strand, but longer than the template strand indicating that BER in a hairpin loop is coupled with removal of a part or the entirety of the hairpin attenuating or preventing TNR expansion. Our discovery further indicates that BER in a hairpin loop may be developed into a novel strategy for treatment for neurodegenerative diseases caused by TNR expansion.

The results from the first two parts of my dissertation research have demonstrated that BER of oxidative DNA damage can directly cause TNR deletion. However, some oxidized DNA base lesions are resistant to BER and other DNA repair pathways and have to be subject to lesion bypass. These base lesions include cyclo-2'-deoxypurines that can accumulate in human genome, especially in a TNR tract. These lesions are usually bypassed by DNA polymerases. A lesion bypass process of a cyclo-deoxypurine located in a TNR tract could also lead to TNR instability. Although pol  $\beta$  is mainly responsible for removing a native 5'-dRP and perform gap-filling synthesis during BER, the enzyme may also perform a lesion bypass synthesis as the Y family translesion synthesis polymerases. During its lesion bypass synthesis in the context of a TNR tract, pol  $\beta$  could modulate TNR instability. To test this hypothesis, we examined pol  $\beta$  lesion bypass of a 5',8-cyclo-deoxyadenosine (5',8-cdA) in the context of (CAG)<sub>10</sub> repeats and its effects on TNR instability during BER and Okazaki fragment maturation. The results showed that pol  $\beta$  efficiently bypassed a 5',8-cdA located in the context of CAG repeats and resulted in CTG repeat deletion during BER and DNA lagging strand maturation. Because recent studies have shown that a 5',8-cdA can induce DNA double helix distortion, it is possible that the distorted double helix favors the formation of secondary structures such as a hairpin or loop in a TNR tract. To test this possibility, we determined the formation of a CAG repeat hairpin/loop during pol  $\beta$  bypass of a 5',8-cdA. We found that a 5',8-cdA opposite or at the downstream of a 1-nt gap, induced the formation of CAG repeat loops with varying sizes prior to pol  $\beta$  lesion bypass synthesis. Interestingly, we found that these CAG repeat loops continued to be present after pol  $\beta$  lesion bypass synthesis indicating that pol  $\beta$  skipped over the CAG repeat loops during its lesion bypass thesis of

a 5',8-cdA. Further characterization of FEN1 cleavage activity indicates that the formation of CAG repeat loops on the template strand also induced realignment of the upstream primer relative to the template strand. The realignment of upstream primer promoted dissociation of the downstream strand from the template strand resulting in the formation of a downstream CAG repeat flap that was efficiently cleaved by FEN1. Thus, our results demonstrate that pol  $\beta$  can bypass a 5',8-cdA by skipping over of a CAG repeat loop, and this further results in CAG repeat deletion via cooperation with FEN1 flap cleavage.

By exploring the roles of BER core enzymes, pol  $\beta$  and FEN1 in modulating TNR instability during repair of oxidative DNA damage or lesion bypass of a cyclo-deoxypurine lesion, we have defined several mechanisms underlying TNR deletion through BER and oxidative DNA damage. Our research has advanced our understanding of how DNA base lesions and repair can regulate TNR stability of. Our discovery can help to develop a novel strategy to treat neurodegenerative diseases via shortening expanded repeats through DNA base lesions and BER.

Thus far, our studies have mainly defined the roles of several core BER enzymes in promoting TNR deletion and preventing TNR expansion in the absence of BER cofactors that can interact with the core BER enzymes to stimulate BER enzymatic activities. Moreover, BER enzymes and cofactors can interact with enzymes and cofactors from other repair pathways through physical interaction or functional cooperation. In addition, some repair proteins from BER or other repair pathways can stabilize or destabilize the secondary structures formed by TNRs such as hairpins during BER that may also modulate TNR instability. Thus, the effects of the interactions among

BER core enzymes, BER cofactors including PCNA, Werner syndrome protein and poly(ADP-ribose) polymerase 1 as well as mismatch repair proteins, MSH2/MSH3, the 3'-5' endonuclease Mus81/Eme1 on TNR instability needs to be elucidated in the future. Our research is limited by using synthesized oligonucleotides that can only represent naked double strand DNA, which are different from the DNA wrapped and compacted in human chromatin in cells. In addition, the posttranslational modifications of BER enzymes and cofactors, histone modifications, DNA methylation pattern could also actively be involved in modulating TNR instability during BER. It appears that the mechanisms underlying TNR instability induced by DNA damage and repair in the context of chromatin, posttranslational modification of repair proteins and histones and their modifications as well as epigenetic features need to be explored in the future. The research in these areas will lead to a comprehensive understanding of mechanisms TNR deletion/contraction mediated by BER *in vivo*.



## REFERENCE

1. Mahabir, A.G., Schaap, M.M., Pennings, J.L., van Benthem, J., Hendriksen, C.F. and van Steeg, H. (2010) Comparison of clastogen-induced gene expression profiles in wild-type and DNA repair-deficient Rad54/Rad54B cells. *BMC genomics*, **11**, 24.
2. [http://utminers.utep.edu/rwebb/html/mutagenesis\\_\\_spontaneous.html](http://utminers.utep.edu/rwebb/html/mutagenesis__spontaneous.html).
3. Brenner, H. In Dnarepair1 (ed.), Wikipedia, Vol. 43K, pp. DNA repair rate is an important determinant of cell pathology.
4. Mirkin, S.M. (2007) Expandable DNA repeats and human disease. *Nature*, **447**, 932-940.
5. Lin, Y., Hubert, L., Jr. and Wilson, J.H. (2009) Transcription destabilizes triplet repeats. *Molecular carcinogenesis*, **48**, 350-361.
6. Budworth, H. and McMurray, C.T. (2013) A brief history of triplet repeat diseases. *Methods in molecular biology*, **1010**, 3-17.
7. David, S.S., O'Shea, V.L. and Kundu, S. (2007) Base-excision repair of oxidative DNA damage. *Nature*, **447**, 941-950.
8. Parniewski, P. and Staczek, P. (2002) Molecular mechanisms of TRS instability. *Advances in experimental medicine and biology*, **516**, 1-25.
9. Hoeijmakers, J.H. (2001) Genome maintenance mechanisms for preventing cancer. *Nature*, **411**, 366-374.
10. McMurray, C.T. (2008) Hijacking of the mismatch repair system to cause CAG expansion and cell death in neurodegenerative disease. *DNA repair*, **7**, 1121-1134.
11. Liu, Y. and Wilson, S.H. (2012) DNA base excision repair: a mechanism of trinucleotide repeat expansion. *Trends in biochemical sciences*, **37**, 162-172.
12. McMurray, C.T. (2010) Mechanisms of trinucleotide repeat instability during human development. *Nature reviews. Genetics*, **11**, 786-799.
13. Pena-Diaz, J. and Jiricny, J. (2010) PCNA and MutLalpha: partners in crime in triplet repeat expansion? *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 16409-16410.
14. Dolle, M.E., Kuiper, R.V., Roodbergen, M., Robinson, J., de Vlugt, S., Wijnhoven, S.W., Beems, R.B., de la Fonteyne, L., de With, P., van der Pluijm, I.

- et al.* (2011) Broad segmental progeroid changes in short-lived *Ercc1(-/Delta7)* mice. *Pathobiology of aging & age related diseases*, **1**.
15. Lindahl, T. (1993) Instability and decay of the primary structure of DNA. *Nature*, **362**, 709-715.
  16. Doetsch, P.W. and Cunningham, R.P. (1990) The enzymology of apurinic/apyrimidinic endonucleases. *Mutation research*, **236**, 173-201.
  17. Lindahl, T. and Andersson, A. (1972) Rate of chain breakage at apurinic sites in double-stranded deoxyribonucleic acid. *Biochemistry*, **11**, 3618-3623.
  18. Shapiro, R. (1981) *Chromosome Damage and Repair*. Plenum, New York.
  19. Lindahl, T. and Nyberg, B. (1974) Heat-induced deamination of cytosine residues in deoxyribonucleic acid. *Biochemistry*, **13**, 3405-3410.
  20. Lindahl, T. (1979) DNA glycosylases, endonucleases for apurinic/apyrimidinic sites, and base excision-repair. *Progress in nucleic acid research and molecular biology*, **22**, 135-192.
  21. Ehrlich, M., Zhang, X.Y. and Inamdar, N.M. (1990) Spontaneous deamination of cytosine and 5-methylcytosine residues in DNA and replacement of 5-methylcytosine residues with cytosine residues. *Mutation research*, **238**, 277-286.
  22. Friedberg, E.C. (2003) DNA damage and repair. *Nature*, **421**, 436-440.
  23. Pfeifer, G.P., Denissenko, M.F., Olivier, M., Tretyakova, N., Hecht, S.S. and Hainaut, P. (2002) Tobacco smoke carcinogens, DNA damage and p53 mutations in smoking-associated cancers. *Oncogene*, **21**, 7435-7451.
  24. Cadet, J., Delatour, T., Douki, T., Gasparutto, D., Pouget, J.P., Ravanat, J.L. and Sauvaigo, S. (1999) Hydroxyl radicals and DNA base damage. *Mutation research*, **424**, 9-21.
  25. Fuss, J.O. and Cooper, P.K. (2006) DNA repair: dynamic defenders against cancer and aging. *PLoS biology*, **4**, e203.
  26. Hashem, V.I. and Sinden, R.R. (2002) Chemotherapeutically induced deletion of expanded triplet repeats. *Mutation research*, **508**, 107-119.
  27. Su, Y., Meador, J.A., Geard, C.R. and Balajee, A.S. (2010) Analysis of ionizing radiation-induced DNA damage and repair in three-dimensional human skin model system. *Experimental dermatology*, **19**, e16-22.
  28. Acharya, P. (1975), *10th International Congress of Gerontology*, Jerusalem.

29. Acharya, P. (1976), *10th International Congress of Biochemistry*, Hamburg, Germany.
30. Acharya, P.N. (1977), *1st International Meeting of Heads of Clinical Biochemistry Laboratories*, Jerusalem, Israel.
31. Breitling, L.P., Yang, R., Korn, B., Burwinkel, B. and Brenner, H. (2011) Tobacco-smoking-related differential DNA methylation: 27K discovery and replication. *American journal of human genetics*, **88**, 450-457.
32. Hashem, V.I., Pytlos, M.J., Klysik, E.A., Tsuji, K., Khajavi, M., Ashizawa, T. and Sinden, R.R. (2004) Chemotherapeutic deletion of CTG repeats in lymphoblast cells from DM1 patients. *Nucleic acids research*, **32**, 6334-6346.
33. Caldecott, K.W. (2008) Single-strand break repair and genetic disease. *Nature reviews. Genetics*, **9**, 619-631.
34. J., L.H.a.D. (2004) *Molecular Biology of the Cell*. Macmillan.
35. Brooks, P.J., Wise, D.S., Berry, D.A., Kosmoski, J.V., Smerdon, M.J., Somers, R.L., Mackie, H., Spoonde, A.Y., Ackerman, E.J., Coleman, K. *et al.* (2000) The oxidative DNA lesion 8,5'-(S)-cyclo-2'-deoxyadenosine is repaired by the nucleotide excision repair pathway and blocks gene expression in mammalian cells. *The Journal of biological chemistry*, **275**, 22355-22362.
36. Best, B.P. (2009) Nuclear DNA damage as a direct cause of aging. *Rejuvenation research*, **12**, 199-208.
37. Seviour, E.G. and Lin, S.Y. (2010) The DNA damage response: Balancing the scale between cancer and ageing. *Aging*, **2**, 900-907.
38. Brasnjevic, I., Hof, P.R., Steinbusch, H.W. and Schmitz, C. (2008) Accumulation of nuclear DNA damage or neuron loss: molecular basis for a new approach to understanding selective neuronal vulnerability in neurodegenerative diseases. *DNA repair*, **7**, 1087-1097.
39. Kubota, Y., Nash, R.A., Klungland, A., Schar, P., Barnes, D.E. and Lindahl, T. (1996) Reconstitution of DNA base excision-repair with purified human proteins: interaction between DNA polymerase beta and the XRCC1 protein. *The EMBO journal*, **15**, 6662-6670.
40. Wilson, S.H. (1998) Mammalian base excision repair and DNA polymerase beta. *Mutation research*, **407**, 203-215.

41. Wilson, D.M., 3rd and Thompson, L.H. (1997) Life without DNA repair. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 12754-12757.
42. Lindahl, T. and Wood, R.D. (1999) Quality control by DNA repair. *Science*, **286**, 1897-1905.
43. Hoeijmakers, J.H. (2009) DNA damage, aging, and cancer. *The New England journal of medicine*, **361**, 1475-1485.
44. Espejel, S., Martin, M., Klatt, P., Martin-Caballero, J., Flores, J.M. and Blasco, M.A. (2004) Shorter telomeres, accelerated ageing and increased lymphoma in DNA-PKcs-deficient mice. *EMBO reports*, **5**, 503-509.
45. de Boer, J., Andressoo, J.O., de Wit, J., Huijmans, J., Beems, R.B., van Steeg, H., Weeda, G., van der Horst, G.T., van Leeuwen, W., Themmen, A.P. *et al.* (2002) Premature aging in mice deficient in DNA repair and transcription. *Science*, **296**, 1276-1279.
46. Dolle, M.E., Busuttil, R.A., Garcia, A.M., Wijnhoven, S., van Drunen, E., Niedernhofer, L.J., van der Horst, G., Hoeijmakers, J.H., van Steeg, H. and Vijg, J. (2006) Increased genomic instability is not a prerequisite for shortened lifespan in DNA repair deficient mice. *Mutation research*, **596**, 22-35.
47. Bjorksten, J., Acharya, P.V., Ashman, S. and Wetlaufer, D.B. (1971) Gerogenic fractions in the tritiated rat. *Journal of the American Geriatrics Society*, **19**, 561-574.
48. Waters, L.S., Minesinger, B.K., Wiltout, M.E., D'Souza, S., Woodruff, R.V. and Walker, G.C. (2009) Eukaryotic translesion polymerases and their roles and regulation in DNA damage tolerance. *Microbiology and molecular biology reviews : MMBR*, **73**, 134-154.
49. Colis, L.C., Raychaudhury, P. and Basu, A.K. (2008) Mutational specificity of gamma-radiation-induced guanine-thymine and thymine-guanine intrastrand cross-links in mammalian cells and translesion synthesis past the guanine-thymine lesion by human DNA polymerase  $\eta$ . *Biochemistry*, **47**, 8070-8079.
50. Raychaudhury, P. and Basu, A.K. (2011) Genetic requirement for mutagenesis of the G[8,5-Me]T cross-link in *Escherichia coli*: DNA polymerases IV and V compete for error-prone bypass. *Biochemistry*, **50**, 2330-2338.
51. Wang, Z. (2001) Translesion synthesis by the UmuC family of DNA polymerases. *Mutation research*, **486**, 59-70.

52. Neeley, W.L. and Essigmann, J.M. (2006) Mechanisms of formation, genotoxicity, and mutation of guanine oxidation products. *Chemical research in toxicology*, **19**, 491-505.
53. Waris, G. and Ahsan, H. (2006) Reactive oxygen species: role in the development of cancer and various chronic conditions. *Journal of carcinogenesis*, **5**, 14.
54. Shibutani, S., Takeshita, M. and Grollman, A.P. (1991) Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG. *Nature*, **349**, 431-434.
55. Paz-Elizur, T., Ben-Yosef, R., Elinger, D., Vexler, A., Krupsky, M., Berrebi, A., Shani, A., Schechtman, E., Freedman, L. and Livneh, Z. (2006) Reduced repair of the oxidative 8-oxoguanine DNA damage and risk of head and neck cancer. *Cancer research*, **66**, 11683-11689.
56. Karihtala, P., Kauppila, S., Puistola, U. and Jukkola-Vuorinen, A. (2011) Divergent behaviour of oxidative stress markers 8-hydroxydeoxyguanosine (8-OHdG) and 4-hydroxy-2-nonenal (HNE) in breast carcinogenesis. *Histopathology*, **58**, 854-862.
57. Pylvas, M., Puistola, U., Laatio, L., Kauppila, S. and Karihtala, P. (2011) Elevated serum 8-OHdG is associated with poor prognosis in epithelial ovarian cancer. *Anticancer research*, **31**, 1411-1415.
58. Lan, L., Nakajima, S., Oohata, Y., Takao, M., Okano, S., Masutani, M., Wilson, S.H. and Yasui, A. (2004) In situ analysis of repair processes for oxidative DNA damage in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 13738-13743.
59. Okano, S., Lan, L., Caldecott, K.W., Mori, T. and Yasui, A. (2003) Spatial and temporal cellular responses to single-strand breaks in human cells. *Molecular and cellular biology*, **23**, 3974-3981.
60. Frosina, G., Fortini, P., Rossi, O., Carrozzino, F., Raspaglio, G., Cox, L.S., Lane, D.P., Abbondandolo, A. and Dogliotti, E. (1996) Two pathways for base excision repair in mammalian cells. *The Journal of biological chemistry*, **271**, 9573-9578.
61. Klungland, A. and Lindahl, T. (1997) Second pathway for completion of human DNA base excision-repair: reconstitution with purified proteins and requirement for DNase IV (FEN1). *The EMBO journal*, **16**, 3341-3348.
62. Fortini, P., Pascucci, B., Parlanti, E., Sobol, R.W., Wilson, S.H. and Dogliotti, E. (1998) Different DNA polymerases are involved in the short- and long-patch base excision repair in mammalian cells. *Biochemistry*, **37**, 3575-3580.

63. Biade, S., Sobol, R.W., Wilson, S.H. and Matsumoto, Y. (1998) Impairment of proliferating cell nuclear antigen-dependent apurinic/apyrimidinic site repair on linear DNA. *The Journal of biological chemistry*, **273**, 898-902.
64. Singhal, R.K., Prasad, R. and Wilson, S.H. (1995) DNA polymerase beta conducts the gap-filling step in uracil-initiated base excision repair in a bovine testis nuclear extract. *The Journal of biological chemistry*, **270**, 949-957.
65. Dianov, G., Price, A. and Lindahl, T. (1992) Generation of single-nucleotide repair patches following excision of uracil residues from DNA. *Molecular and cellular biology*, **12**, 1605-1612.
66. Wilson, S.H. and Kunkel, T.A. (2000) Passing the baton in base excision repair. *Nature structural biology*, **7**, 176-178.
67. Prasad, R., Shock, D.D., Beard, W.A. and Wilson, S.H. (2010) Substrate channeling in mammalian base excision repair pathways: passing the baton. *The Journal of biological chemistry*, **285**, 40479-40488.
68. Mol, C.D., Izumi, T., Mitra, S. and Tainer, J.A. (2000) DNA-bound structures and mutants reveal abasic DNA binding by APE1 and DNA repair coordination [corrected]. *Nature*, **403**, 451-456.
69. Parikh, S.S., Mol, C.D., Hosfield, D.J. and Tainer, J.A. (1999) Envisioning the molecular choreography of DNA base excision repair. *Current opinion in structural biology*, **9**, 37-47.
70. Toth, G., Gaspari, Z. and Jurka, J. (2000) Microsatellites in different eukaryotic genomes: survey and analysis. *Genome research*, **10**, 967-981.
71. Freudenreich, C.H., Kantrow, S.M. and Zakian, V.A. (1998) Expansion and length-dependent fragility of CTG repeats in yeast. *Science*, **279**, 853-856.
72. Paulson, H.L. and Fischbeck, K.H. (1996) Trinucleotide repeats in neurogenetic disorders. *Annual review of neuroscience*, **19**, 79-107.
73. Richards, R.I. and Sutherland, G.R. (1997) Dynamic mutation: possible mechanisms and significance in human disease. *Trends in biochemical sciences*, **22**, 432-436.
74. Pearson, C.E., Nichol Edamura, K. and Cleary, J.D. (2005) Repeat instability: mechanisms of dynamic mutations. *Nature reviews. Genetics*, **6**, 729-742.
75. Wang, J., Yuan, B., Guerrero, C., Bahde, R., Gupta, S. and Wang, Y. (2011) Quantification of oxidative DNA lesions in tissues of Long-Evans Cinnamon rats

- by capillary high-performance liquid chromatography-tandem mass spectrometry coupled with stable isotope-dilution method. *Analytical chemistry*, **83**, 2201-2209.
76. Cummings, C.J. and Zoghbi, H.Y. (2000) Fourteen and counting: unraveling trinucleotide repeat diseases. *Human molecular genetics*, **9**, 909-916.
  77. Freudenreich, C.H. and Lahiri, M. (2004) Structure-forming CAG/CTG repeat sequences are sensitive to breakage in the absence of Mrc1 checkpoint function and S-phase checkpoint signaling: implications for trinucleotide repeat expansion diseases. *Cell cycle*, **3**, 1370-1374.
  78. Everett, C.M. and Wood, N.W. (2004) Trinucleotide repeats and neurodegenerative disease. *Brain : a journal of neurology*, **127**, 2385-2405.
  79. Tapscott, S.J. and Thornton, C.A. (2001) Biomedicine. Reconstructing myotonic dystrophy. *Science*, **293**, 816-817.
  80. Helmlinger, D., Tora, L. and Devys, D. (2006) Transcriptional alterations and chromatin remodeling in polyglutamine diseases. *Trends in genetics : TIG*, **22**, 562-570.
  81. Ashley, C.T., Jr. and Warren, S.T. (1995) Trinucleotide repeat expansion and human disease. *Annual review of genetics*, **29**, 703-728.
  82. Schildkraut, J.M., Murphy, S.K., Palmieri, R.T., Iversen, E., Moorman, P.G., Huang, Z., Halabi, S., Calingaert, B., Gusberg, A., Marks, J.R. *et al.* (2007) Trinucleotide repeat polymorphisms in the androgen receptor gene and risk of ovarian cancer. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, **16**, 473-480.
  83. Andersson, P., Varenhorst, E. and Soderkvist, P. (2006) Androgen receptor and vitamin D receptor gene polymorphisms and prostate cancer risk. *European journal of cancer*, **42**, 2833-2837.
  84. Haberman, Y., Amariglio, N., Rechavi, G. and Eisenberg, E. (2008) Trinucleotide repeats are prevalent among cancer-related genes. *Trends in genetics : TIG*, **24**, 14-18.
  85. Wells, R.D. (1996) Molecular basis of genetic instability of triplet repeats. *The Journal of biological chemistry*, **271**, 2875-2878.
  86. Wells, R.D., Dere, R., Hebert, M.L., Napierala, M. and Son, L.S. (2005) Advances in mechanisms of genetic instability related to hereditary neurological diseases. *Nucleic acids research*, **33**, 3785-3798.

87. Mitas, M. (1997) Trinucleotide repeats associated with human disease. *Nucleic acids research*, **25**, 2245-2254.
88. Usdin, K. and Woodford, K.J. (1995) CGG repeats associated with DNA instability and chromosome fragility form structures that block DNA synthesis in vitro. *Nucleic acids research*, **23**, 4202-4209.
89. Gacy, A.M., Goellner, G.M., Spiro, C., Chen, X., Gupta, G., Bradbury, E.M., Dyer, R.B., Mikesell, M.J., Yao, J.Z., Johnson, A.J. *et al.* (1998) GAA instability in Friedreich's Ataxia shares a common, DNA-directed and intraallelic mechanism with other trinucleotide diseases. *Molecular cell*, **1**, 583-593.
90. Kim, J.C. and Mirkin, S.M. (2013) The balancing act of DNA repeat expansions. *Current opinion in genetics & development*, **23**, 280-288.
91. Kang, S., Jaworski, A., Ohshima, K. and Wells, R.D. (1995) Expansion and deletion of CTG repeats from human disease genes are determined by the direction of replication in *E. coli*. *Nature genetics*, **10**, 213-218.
92. Ohshima, K., Kang, S. and Wells, R.D. (1996) CTG triplet repeats from human hereditary diseases are dominant genetic expansion products in *Escherichia coli*. *The Journal of biological chemistry*, **271**, 1853-1856.
93. Iyer, R.R. and Wells, R.D. (1999) Expansion and deletion of triplet repeat sequences in *Escherichia coli* occur on the leading strand of DNA replication. *The Journal of biological chemistry*, **274**, 3865-3877.
94. Sarkar, P.S., Chang, H.C., Boudi, F.B. and Reddy, S. (1998) CTG repeats show bimodal amplification in *E. coli*. *Cell*, **95**, 531-540.
95. Freudenreich, C.H., Stavenhagen, J.B. and Zakian, V.A. (1997) Stability of a CTG/CAG trinucleotide repeat in yeast is dependent on its orientation in the genome. *Molecular and cellular biology*, **17**, 2090-2098.
96. Schweitzer, J.K. and Livingston, D.M. (1998) Expansions of CAG repeat tracts are frequent in a yeast mutant defective in Okazaki fragment maturation. *Human molecular genetics*, **7**, 69-74.
97. Krasilnikova, M.M. and Mirkin, S.M. (2004) Replication stalling at Friedreich's ataxia (GAA)<sub>n</sub> repeats in vivo. *Molecular and cellular biology*, **24**, 2286-2295.
98. Mirkin, E.V. and Mirkin, S.M. (2007) Replication fork stalling at natural impediments. *Microbiology and molecular biology reviews : MMBR*, **71**, 13-35.



99. Liu, L.F. and Wang, J.C. (1987) Supercoiling of the DNA template during transcription. *Proceedings of the National Academy of Sciences of the United States of America*, **84**, 7024-7027.
100. Bowater, R.P., Jaworski, A., Larson, J.E., Parniewski, P. and Wells, R.D. (1997) Transcription increases the deletion frequency of long CTG.CAG triplet repeats from plasmids in *Escherichia coli*. *Nucleic acids research*, **25**, 2861-2868.
101. Parniewski, P., Bacolla, A., Jaworski, A. and Wells, R.D. (1999) Nucleotide excision repair affects the stability of long transcribed (CTG\*CAG) tracts in an orientation-dependent manner in *Escherichia coli*. *Nucleic acids research*, **27**, 616-623.
102. Gonitel, R., Moffitt, H., Sathasivam, K., Woodman, B., Detloff, P.J., Faull, R.L. and Bates, G.P. (2008) DNA instability in postmitotic neurons. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 3467-3472.
103. Kovtun, I.V., Liu, Y., Bjoras, M., Klungland, A., Wilson, S.H. and McMurray, C.T. (2007) OGG1 initiates age-dependent CAG trinucleotide expansion in somatic cells. *Nature*, **447**, 447-452.
104. Mirkin, S.M. (2006) DNA structures, repeat expansions and human hereditary disorders. *Current opinion in structural biology*, **16**, 351-358.
105. Cleary, J.D. and Pearson, C.E. (2005) Replication fork dynamics and dynamic mutations: the fork-shift model of repeat instability. *Trends in genetics : TIG*, **21**, 272-280.
106. Dion, V. and Wilson, J.H. (2009) Instability and chromatin structure of expanded trinucleotide repeats. *Trends in genetics : TIG*, **25**, 288-297.
107. Wells, R.D., Dere, R., Hebert, M.L., Napierala, M. and Son, L.S. (2005) Advances in mechanisms of genetic instability related to hereditary neurological diseases. *Nucleic acids research*, **33**, 3785-3798.
108. Kang, S., Ohshima, K., Shimizu, M., Amirhaeri, S. and Wells, R.D. (1995) Pausing of DNA synthesis in vitro at specific loci in CTG and CGG triplet repeats from human hereditary disease genes. *The Journal of biological chemistry*, **270**, 27014-27021.
109. Kamath-Loeb, A.S., Loeb, L.A., Johansson, E., Burgers, P.M. and Fry, M. (2001) Interactions between the Werner syndrome helicase and DNA polymerase delta specifically facilitate copying of tetraplex and hairpin structures of the d(CGG)<sub>n</sub> trinucleotide repeat sequence. *The Journal of biological chemistry*, **276**, 16439-16446.

110. Spiro, C., Pelletier, R., Rolfsmeier, M.L., Dixon, M.J., Lahue, R.S., Gupta, G., Park, M.S., Chen, X., Mariappan, S.V. and McMurray, C.T. (1999) Inhibition of FEN-1 processing by DNA secondary structure at trinucleotide repeats. *Molecular cell*, **4**, 1079-1085.
111. Henricksen, L.A., Tom, S., Liu, Y. and Bambara, R.A. (2000) Inhibition of flap endonuclease 1 by flap secondary structure and relevance to repeat sequence expansion. *The Journal of biological chemistry*, **275**, 16420-16427.
112. Liu, Y. and Bambara, R.A. (2003) Analysis of human flap endonuclease 1 mutants reveals a mechanism to prevent triplet repeat expansion. *The Journal of biological chemistry*, **278**, 13728-13739.
113. Owen, B.A., Yang, Z., Lai, M., Gajec, M., Badger, J.D., 2nd, Hayes, J.J., Edelman, W., Kucherlapati, R., Wilson, T.M. and McMurray, C.T. (2005) (CAG)(n)-hairpin DNA binds to Msh2-Msh3 and changes properties of mismatch recognition. *Nature structural & molecular biology*, **12**, 663-670.
114. Lang, W.H., Coats, J.E., Majka, J., Hura, G.L., Lin, Y., Rasnik, I. and McMurray, C.T. (2011) Conformational trapping of mismatch recognition complex MSH2/MSH3 on repair-resistant DNA loops. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, E837-844.
115. Liu, Y., Prasad, R., Beard, W.A., Hou, E.W., Horton, J.K., McMurray, C.T. and Wilson, S.H. (2009) Coordination between polymerase beta and FEN1 can modulate CAG repeat expansion. *The Journal of biological chemistry*, **284**, 28352-28366.
116. Richard, G.F., Goellner, G.M., McMurray, C.T. and Haber, J.E. (2000) Recombination-induced CAG trinucleotide repeat expansions in yeast involve the MRE11-RAD50-XRS2 complex. *The EMBO journal*, **19**, 2381-2390.
117. Kovtun, I.V. and McMurray, C.T. (2001) Trinucleotide expansion in haploid germ cells by gap repair. *Nature genetics*, **27**, 407-411.
118. Manley, K., Shirley, T.L., Flaherty, L. and Messer, A. (1999) Msh2 deficiency prevents in vivo somatic instability of the CAG repeat in Huntington disease transgenic mice. *Nature genetics*, **23**, 471-473.
119. Panigrahi, G.B., Slean, M.M., Simard, J.P., Gileadi, O. and Pearson, C.E. (2010) Isolated short CTG/CAG DNA slip-outs are repaired efficiently by hMutSbeta, but clustered slip-outs are poorly repaired. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 12593-12598.

120. Acharya, S., Wilson, T., Gradia, S., Kane, M.F., Guerrette, S., Marsischky, G.T., Kolodner, R. and Fishel, R. (1996) hMSH2 forms specific mispair-binding complexes with hMSH3 and hMSH6. *Proceedings of the National Academy of Sciences of the United States of America*, **93**, 13629-13634.
121. Alani, E., Lee, S., Kane, M.F., Griffith, J. and Kolodner, R.D. (1997) *Saccharomyces cerevisiae* MSH2, a mispaired base recognition protein, also recognizes Holliday junctions in DNA. *Journal of molecular biology*, **265**, 289-301.
122. Shell, S.S., Putnam, C.D. and Kolodner, R.D. (2007) Chimeric *Saccharomyces cerevisiae* Msh6 protein with an Msh3 mispair-binding domain combines properties of both proteins. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 10956-10961.
123. Harrington, J.M. and Kolodner, R.D. (2007) *Saccharomyces cerevisiae* Msh2-Msh3 acts in repair of base-base mismatches. *Molecular and cellular biology*, **27**, 6546-6554.
124. Obmolova, G., Ban, C., Hsieh, P. and Yang, W. (2000) Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA. *Nature*, **407**, 703-710.
125. Gradia, S., Subramanian, D., Wilson, T., Acharya, S., Makhov, A., Griffith, J. and Fishel, R. (1999) hMSH2-hMSH6 forms a hydrolysis-independent sliding clamp on mismatched DNA. *Molecular cell*, **3**, 255-261.
126. Gacy, A.M., Goellner, G., Juranic, N., Macura, S. and McMurray, C.T. (1995) Trinucleotide repeats that expand in human disease form hairpin structures in vitro. *Cell*, **81**, 533-540.
127. van den Broek, W.J., Nelen, M.R., Wansink, D.G., Coerwinkel, M.M., te Riele, H., Groenen, P.J. and Wieringa, B. (2002) Somatic expansion behaviour of the (CTG)<sub>n</sub> repeat in myotonic dystrophy knock-in mice is differentially affected by Msh3 and Msh6 mismatch-repair proteins. *Human molecular genetics*, **11**, 191-198.
128. Moore, H., Greenwell, P.W., Liu, C.P., Arnheim, N. and Petes, T.D. (1999) Triplet repeats form secondary structures that escape DNA repair in yeast. *Proceedings of the National Academy of Sciences of the United States of America*, **96**, 1504-1509.
129. Ahn, B. and Grossman, L. (1996) The binding of UvrAB proteins to bubble and loop regions in duplex DNA. *The Journal of biological chemistry*, **271**, 21462-21470.

130. Kirkpatrick, D.T. and Petes, T.D. (1997) Repair of DNA loops involves DNA-mismatch and nucleotide-excision repair proteins. *Nature*, **387**, 929-931.
131. Jung, J. and Bonini, N. (2007) CREB-binding protein modulates repeat instability in a Drosophila model for polyQ disease. *Science*, **315**, 1857-1859.
132. Lin, Y., Dion, V. and Wilson, J.H. (2006) Transcription promotes contraction of CAG repeat tracts in human cells. *Nature structural & molecular biology*, **13**, 179-180.
133. Lin, Y. and Wilson, J.H. (2007) Transcription-induced CAG repeat contraction in human cells is mediated in part by transcription-coupled nucleotide excision repair. *Molecular and cellular biology*, **27**, 6209-6217.
134. Cleaver, J.E., Lam, E.T. and Revet, I. (2009) Disorders of nucleotide excision repair: the genetic and molecular basis of heterogeneity. *Nature reviews. Genetics*, **10**, 756-768.
135. Hanawalt, P.C. and Spivak, G. (2008) Transcription-coupled DNA repair: two decades of progress and surprises. *Nature reviews. Molecular cell biology*, **9**, 958-970.
136. Le May, N., Egly, J.M. and Coin, F. (2010) True lies: the double life of the nucleotide excision repair factors in transcription and DNA repair. *Journal of nucleic acids*, **2010**.
137. Staresincic, L., Fagbemi, A.F., Enzlin, J.H., Gourdin, A.M., Wijgers, N., Dunand-Sauthier, I., Giglia-Mari, G., Clarkson, S.G., Vermeulen, W. and Scharer, O.D. (2009) Coordination of dual incision and repair synthesis in human nucleotide excision repair. *The EMBO journal*, **28**, 1111-1120.
138. Dragileva, E., Hendricks, A., Teed, A., Gillis, T., Lopez, E.T., Friedberg, E.C., Kucherlapati, R., Edelman, W., Lunetta, K.L., MacDonald, M.E. *et al.* (2009) Intergenerational and striatal CAG repeat instability in Huntington's disease knock-in mice involve different DNA repair genes. *Neurobiology of disease*, **33**, 37-47.
139. Hubert, L., Jr., Lin, Y., Dion, V. and Wilson, J.H. (2011) Xpa deficiency reduces CAG trinucleotide repeat instability in neuronal tissues in a mouse model of SCA1. *Human molecular genetics*, **20**, 4822-4830.
140. Tang, W., Dominska, M., Greenwell, P.W., Harvanek, Z., Lobachev, K.S., Kim, H.M., Narayanan, V., Mirkin, S.M. and Petes, T.D. (2011) Friedreich's ataxia (GAA)<sub>n</sub>\*(TTC)<sub>n</sub> repeats strongly stimulate mitotic crossovers in *Saccharomyces cerevisiae*. *PLoS genetics*, **7**, e1001270.

141. Napierala, M., Parniewski, P., Pluciennik, A. and Wells, R.D. (2002) Long CTG.CAG repeat sequences markedly stimulate intramolecular recombination. *The Journal of biological chemistry*, **277**, 34087-34100.
142. Pluciennik, A., Iyer, R.R., Napierala, M., Larson, J.E., Filutowicz, M. and Wells, R.D. (2002) Long CTG.CAG repeats from myotonic dystrophy are preferred sites for intermolecular recombination. *The Journal of biological chemistry*, **277**, 34074-34086.
143. Jankowski, C. and Nag, D.K. (2002) Most meiotic CAG repeat tract-length alterations in yeast are SPO11 dependent. *Molecular genetics and genomics : MGG*, **267**, 64-70.
144. Bishop, A.J. and Schiestl, R.H. (2000) Homologous recombination as a mechanism for genome rearrangements: environmental and genetic effects. *Human molecular genetics*, **9**, 2427-2334.
145. Kanaar, R., Hoeijmakers, J.H. and van Gent, D.C. (1998) Molecular mechanisms of DNA double strand break repair. *Trends in cell biology*, **8**, 483-489.
146. Sundararajan, R., Gellon, L., Zunder, R.M. and Freudenreich, C.H. (2010) Double-strand break repair pathways protect against CAG/CTG repeat expansions, contractions and repeat-mediated chromosomal fragility in *Saccharomyces cerevisiae*. *Genetics*, **184**, 65-77.
147. Chandok, G.S., Patel, M.P., Mirkin, S.M. and Krasilnikova, M.M. (2012) Effects of Friedreich's ataxia GAA repeats on DNA replication in mammalian cells. *Nucleic acids research*, **40**, 3964-3974.
148. Klungland, A. and Bjelland, S. (2007) Oxidative damage to purines in DNA: role of mammalian Ogg1. *DNA repair*, **6**, 481-488.
149. Schuster, G.B. (2009) One-electron oxidation of DNA: mechanism and consequences. *Nucleic acids symposium series*, 85-86.
150. Reagan, M.S., Pittenger, C., Siede, W. and Friedberg, E.C. (1995) Characterization of a mutant strain of *Saccharomyces cerevisiae* with a deletion of the RAD27 gene, a structural homolog of the RAD2 nucleotide excision repair gene. *Journal of bacteriology*, **177**, 364-371.
151. Lindahl, T. (1982) DNA repair enzymes. *Annual review of biochemistry*, **51**, 61-87.
152. Kennedy, L., Evans, E., Chen, C.M., Craven, L., Detloff, P.J., Ennis, M. and Shelbourne, P.F. (2003) Dramatic tissue-specific mutation length increases are an

- early molecular event in Huntington disease pathogenesis. *Human molecular genetics*, **12**, 3359-3367.
153. Bogdanov, M.B., Andreassen, O.A., Dedeoglu, A., Ferrante, R.J. and Beal, M.F. (2001) Increased oxidative damage to DNA in a transgenic mouse model of Huntington's disease. *Journal of neurochemistry*, **79**, 1246-1249.
  154. Entezam, A., Lokanga, A.R., Le, W., Hoffman, G. and Usdin, K. (2010) Potassium bromate, a potent DNA oxidizing agent, exacerbates germline repeat expansion in a fragile X premutation mouse model. *Human mutation*, **31**, 611-616.
  155. Goula, A.V., Berquist, B.R., Wilson, D.M., 3rd, Wheeler, V.C., Trottier, Y. and Merienne, K. (2009) Stoichiometry of base excision repair proteins correlates with increased somatic CAG instability in striatum over cerebellum in Huntington's disease transgenic mice. *PLoS genetics*, **5**, e1000749.
  156. Stadtman, E.R. (1992) Protein oxidation and aging. *Science*, **257**, 1220-1224.
  157. Cardozo-Pelaez, F., Song, S., Parthasarathy, A., Hazzi, C., Naidu, K. and Sanchez-Ramos, J. (1999) Oxidative DNA damage in the aging mouse brain. *Movement disorders : official journal of the Movement Disorder Society*, **14**, 972-980.
  158. Shibata, A., Kamada, N., Masumura, K., Nohmi, T., Kobayashi, S., Teraoka, H., Nakagama, H., Sugimura, T., Suzuki, H. and Masutani, M. (2005) Parp-1 deficiency causes an increase of deletion mutations and insertions/rearrangements in vivo after treatment with an alkylating agent. *Oncogene*, **24**, 1328-1337.
  159. Heidenfelder, B.L., Makhov, A.M. and Topal, M.D. (2003) Hairpin formation in Friedreich's ataxia triplet repeat expansion. *The Journal of biological chemistry*, **278**, 2425-2431.
  160. Liu, Y., Prasad, R. and Wilson, S.H. (Wells R.D. and Warren S.T.), . (2006) In Ashizawa, R. D. W. a. T. (ed.), *Genetic Instabilities and Neurological Disease*. 2nd ed, pp. 667-678.
  161. Morel, P., Reverdy, C., Michel, B., Ehrlich, S.D. and Cassuto, E. (1998) The role of SOS and flap processing in microsatellite instability in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 10003-10008.
  162. Nagata, Y., Mashimo, K., Kawata, M. and Yamamoto, K. (2002) The roles of Klenow processing and flap processing activities of DNA polymerase I in chromosome instability in Escherichia coli K12 strains. *Genetics*, **160**, 13-23.

163. Johnson, R.E., Kovvali, G.K., Prakash, L. and Prakash, S. (1995) Requirement of the yeast RTH1 5' to 3' exonuclease for the stability of simple repetitive DNA. *Science*, **269**, 238-240.
164. Kokoska, R.J., Stefanovic, L., Tran, H.T., Resnick, M.A., Gordenin, D.A. and Petes, T.D. (1998) Destabilization of yeast micro- and minisatellite DNA sequences by mutations affecting a nuclease involved in Okazaki fragment processing (rad27) and DNA polymerase delta (pol3-t). *Molecular and cellular biology*, **18**, 2779-2788.
165. White, P.J., Borts, R.H. and Hirst, M.C. (1999) Stability of the human fragile X (CGG)(n) triplet repeat array in *Saccharomyces cerevisiae* deficient in aspects of DNA metabolism. *Molecular and cellular biology*, **19**, 5675-5684.
166. Spiro, C. and McMurray, C.T. (2003) Nuclease-deficient FEN-1 blocks Rad51/BRCA1-mediated repair and causes trinucleotide repeat instability. *Molecular and cellular biology*, **23**, 6063-6074.
167. Spiro, C., Pelletier, R., Rolfsmeier, M.L., Dixon, M.J., Lahue, R.S., Gupta, G., Park, M.S., Chen, X., Mariappan, S.V. and McMurray, C.T. (1999) Inhibition of FEN-1 processing by DNA secondary structure at trinucleotide repeats. *Molecular cell*, **4**, 1079-1085.
168. Henricksen, L.A., Veeraraghavan, J., Chafin, D.R. and Bambara, R.A. (2002) DNA ligase I competes with FEN1 to expand repetitive DNA sequences in vitro. *The Journal of biological chemistry*, **277**, 22361-22369.
169. Refsland, E.W. and Livingston, D.M. (2005) Interactions among DNA ligase I, the flap endonuclease and proliferating cell nuclear antigen in the expansion and contraction of CAG repeat tracts in yeast. *Genetics*, **171**, 923-934.
170. Prasad, R., Lavrik, O.I., Kim, S.J., Kedar, P., Yang, X.P., Vande Berg, B.J. and Wilson, S.H. (2001) DNA polymerase beta -mediated long patch base excision repair. Poly(ADP-ribose)polymerase-1 stimulates strand displacement DNA synthesis. *The Journal of biological chemistry*, **276**, 32411-32414.
171. Asagoshi, K., Liu, Y., Masaoka, A., Lan, L., Prasad, R., Horton, J.K., Brown, A.R., Wang, X.H., Bdour, H.M., Sobol, R.W. *et al.* (2010) DNA polymerase beta-dependent long patch base excision repair in living cells. *DNA repair*, **9**, 109-119.
172. Beard, W.A., Prasad, R. and Wilson, S.H. (2006) Activities and mechanism of DNA polymerase beta. *Methods in enzymology*, **408**, 91-107.
173. Bjelland, S. and Seeberg, E. (2003) Mutagenicity, toxicity and repair of DNA base damage induced by oxidation. *Mutation research*, **531**, 37-80.

174. Kaplan, S., Itzkovitz, S. and Shapiro, E. (2007) A universal mechanism ties genotype to phenotype in trinucleotide diseases. *PLoS computational biology*, **3**, e235.
175. Zhang, Y., Monckton, D.G., Siciliano, M.J., Connor, T.H. and Meistrich, M.L. (2002) Detection of radiation and cyclophosphamide-induced mutations in individual mouse sperm at a human expanded trinucleotide repeat locus transgene. *Mutation research*, **516**, 121-138.
176. Pineiro, E., Fernandez-Lopez, L., Gamez, J., Marcos, R., Surralles, J. and Velazquez, A. (2003) Mutagenic stress modulates the dynamics of CTG repeat instability associated with myotonic dystrophy type 1. *Nucleic acids research*, **31**, 6733-6740.
177. Gomes-Pereira, M. and Monckton, D.G. (2004) Chemically induced increases and decreases in the rate of expansion of a CAG\*CTG triplet repeat. *Nucleic acids research*, **32**, 2865-2872.
178. Nelson, K.A. and Witte, J.S. (2002) Androgen receptor CAG repeats and prostate cancer. *American journal of epidemiology*, **155**, 883-890.
179. Lieberman, A.P. and Robins, D.M. (2008) The androgen receptor's CAG/glutamine tract in mouse models of neurological disease and cancer. *Journal of Alzheimer's disease : JAD*, **14**, 247-255.
180. Savouret, C., Brisson, E., Essers, J., Kanaar, R., Pastink, A., te Riele, H., Junien, C. and Gourdon, G. (2003) CTG repeat instability and size variation timing in DNA repair-deficient mice. *The EMBO journal*, **22**, 2264-2273.
181. Kremer, B., Almqvist, E., Theilmann, J., Spence, N., Telenius, H., Goldberg, Y.P. and Hayden, M.R. (1995) Sex-dependent mechanisms for expansions and contractions of the CAG repeat on affected Huntington disease chromosomes. *American journal of human genetics*, **57**, 343-350.
182. Panigrahi, G.B., Cleary, J.D. and Pearson, C.E. (2002) In vitro (CTG)\*(CAG) expansions and deletions by human cell extracts. *The Journal of biological chemistry*, **277**, 13926-13934.
183. McMurray, C.T. (1999) DNA secondary structure: a common and causative factor for expansion in human disease. *Proceedings of the National Academy of Sciences of the United States of America*, **96**, 1823-1825.
184. Jaworski, A., Rosche, W.A., Gellibolian, R., Kang, S., Shimizu, M., Bowater, R.P., Sinden, R.R. and Wells, R.D. (1995) Mismatch repair in *Escherichia coli* enhances instability of (CTG)<sub>n</sub> triplet repeats from human hereditary diseases.



*Proceedings of the National Academy of Sciences of the United States of America*, **92**, 11019-11023.

185. Pearson, C.E., Ewel, A., Acharya, S., Fishel, R.A. and Sinden, R.R. (1997) Human MSH2 binds to trinucleotide repeat DNA structures associated with neurodegenerative diseases. *Human molecular genetics*, **6**, 1117-1123.
186. Wheeler, V.C., Lebel, L.A., Vrbanac, V., Teed, A., te Riele, H. and MacDonald, M.E. (2003) Mismatch repair gene Msh2 modifies the timing of early disease in Hdh(Q111) striatum. *Human molecular genetics*, **12**, 273-281.
187. Entezam, A., Lokanga, A.R., Le, W., Hoffman, G. and Usdin, K. (2010) Potassium bromate, a potent DNA oxidizing agent, exacerbates germline repeat expansion in a fragile X premutation mouse model. *Human mutation*, **31**, 611-616.
188. Kovtun, I.V., Johnson, K.O. and McMurray, C.T. (2011) Cockayne syndrome B protein antagonizes OGG1 in modulating CAG repeat length in vivo. *Aging*, **3**, 509-514.
189. Lopez Castel, A., Cleary, J.D. and Pearson, C.E. (2010) Repeat instability as the basis for human diseases and as a potential target for therapy. *Nature reviews. Molecular cell biology*, **11**, 165-170.
190. Kroutil, L.C. and Kunkel, T.A. (1999) Deletion errors generated during replication of CAG repeats. *Nucleic acids research*, **27**, 3481-3486.
191. Marsin, S., Vidal, A.E., Sossou, M., Menissier-de Murcia, J., Le Page, F., Boiteux, S., de Murcia, G. and Radicella, J.P. (2003) Role of XRCC1 in the coordination and stimulation of oxidative DNA damage repair initiated by the DNA glycosylase hOGG1. *The Journal of biological chemistry*, **278**, 44068-44074.
192. Yang, J. and Freudenreich, C.H. (2007) Haploinsufficiency of yeast FEN1 causes instability of expanded CAG/CTG tracts in a length-dependent manner. *Gene*, **393**, 110-115.
193. Prasad, R., Liu, Y., Deterding, L.J., Poltoratsky, V.P., Kedar, P.S., Horton, J.K., Kanno, S., Asagoshi, K., Hou, E.W., Khodyreva, S.N. *et al.* (2007) HMGB1 is a cofactor in mammalian base excision repair. *Molecular cell*, **27**, 829-841.
194. Bachrati, C.Z. and Hickson, I.D. (2008) RecQ helicases: guardian angels of the DNA replication fork. *Chromosoma*, **117**, 219-233.

195. Brosh, R.M., Jr., Majumdar, A., Desai, S., Hickson, I.D., Bohr, V.A. and Seidman, M.M. (2001) Unwinding of a DNA triple helix by the Werner and Bloom syndrome helicases. *The Journal of biological chemistry*, **276**, 3024-3030.
196. Mohaghegh, P., Karow, J.K., Brosh, R.M., Jr., Bohr, V.A. and Hickson, I.D. (2001) The Bloom's and Werner's syndrome proteins are DNA structure-specific helicases. *Nucleic acids research*, **29**, 2843-2849.
197. Goula, A.V., Pearson, C.E., Della Maria, J., Trottier, Y., Tomkinson, A.E., Wilson, D.M., 3rd and Merienne, K. (2012) The nucleotide sequence, DNA damage location, and protein stoichiometry influence the base excision repair outcome at CAG/CTG repeats. *Biochemistry*, **51**, 3919-3932.
198. Lahue, R.S. and Slater, D.L. (2003) DNA repair and trinucleotide repeat instability. *Frontiers in bioscience : a journal and virtual library*, **8**, s653-665.
199. Xu, M., Gabison, J. and Liu, Y. (2013) Trinucleotide repeat deletion via a unique hairpin bypass by DNA polymerase beta and alternate flap cleavage by flap endonuclease 1. *Nucleic acids research*, **41**, 1684-1697.
200. Lai, Y., Xu, M., Zhang, Z. and Liu, Y. (2013) Instability of CTG repeats is governed by the position of a DNA base lesion through base excision repair. *PLoS one*, **8**, e56960.
201. Ross, S.A. and Burrows, C.J. (1996) Cytosine-specific chemical probing of DNA using bromide and monoperoxysulfate. *Nucleic acids research*, **24**, 5062-5063.
202. Jarem, D.A., Wilson, N.R. and Delaney, S. (2009) Structure-dependent DNA damage and repair in a trinucleotide repeat sequence. *Biochemistry*, **48**, 6655-6663.
203. Volle, C.B., Jarem, D.A. and Delaney, S. (2012) Trinucleotide repeat DNA alters structure to minimize the thermodynamic impact of 8-oxo-7,8-dihydroguanine. *Biochemistry*, **51**, 52-62.
204. Jarem, D.A., Wilson, N.R., Schermerhorn, K.M. and Delaney, S. (2011) Incidence and persistence of 8-oxo-7,8-dihydroguanine within a hairpin intermediate exacerbates a toxic oxidation cycle associated with trinucleotide repeat expansion. *DNA repair*, **10**, 887-896.
205. Liu, Y., Beard, W.A., Shock, D.D., Prasad, R., Hou, E.W. and Wilson, S.H. (2005) DNA polymerase beta and flap endonuclease 1 enzymatic specificities sustain DNA synthesis for long patch base excision repair. *The Journal of biological chemistry*, **280**, 3665-3674.

206. Ciccia, A., Constantinou, A. and West, S.C. (2003) Identification and characterization of the human mus81-eme1 endonuclease. *The Journal of biological chemistry*, **278**, 25172-25178.
207. Ciccia, A., McDonald, N. and West, S.C. (2008) Structural and functional relationships of the XPF/MUS81 family of proteins. *Annual review of biochemistry*, **77**, 259-287.
208. Bastin-Shanower, S.A., Fricke, W.M., Mullen, J.R. and Brill, S.J. (2003) The mechanism of Mus81-Mms4 cleavage site selection distinguishes it from the homologous endonuclease Rad1-Rad10. *Molecular and cellular biology*, **23**, 3487-3496.
209. Gaillard, P.H., Noguchi, E., Shanahan, P. and Russell, P. (2003) The endogenous Mus81-Eme1 complex resolves Holliday junctions by a nick and counternick mechanism. *Molecular cell*, **12**, 747-759.
210. Osman, F., Dixon, J., Doe, C.L. and Whitby, M.C. (2003) Generating crossovers by resolution of nicked Holliday junctions: a role for Mus81-Eme1 in meiosis. *Molecular cell*, **12**, 761-774.
211. Whitby, M.C., Osman, F. and Dixon, J. (2003) Cleavage of model replication forks by fission yeast Mus81-Eme1 and budding yeast Mus81-Mms4. *The Journal of biological chemistry*, **278**, 6928-6935.
212. Shin, Y.K., Amangyeld, T., Nguyen, T.A., Munashingha, P.R. and Seo, Y.S. (2012) Human MUS81 complexes stimulate flap endonuclease 1. *The FEBS journal*, **279**, 2412-2430.
213. Connelly, J.C., Kirkham, L.A. and Leach, D.R. (1998) The SbcCD nuclease of Escherichia coli is a structural maintenance of chromosomes (SMC) family protein that cleaves hairpin DNA. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 7969-7974.
214. Leach, D.R., Okely, E.A. and Pinder, D.J. (1997) Repair by recombination of DNA containing a palindromic sequence. *Molecular microbiology*, **26**, 597-606.
215. Pearson, C.E., Tam, M., Wang, Y.H., Montgomery, S.E., Dar, A.C., Cleary, J.D. and Nichol, K. (2002) Slipped-strand DNAs formed by long (CAG)<sup>n</sup>(CTG)<sup>m</sup> repeats: slipped-out repeats and slip-out junctions. *Nucleic acids research*, **30**, 4534-4547.
216. Panigrahi, G.B., Lau, R., Montgomery, S.E., Leonard, M.R. and Pearson, C.E. (2005) Slipped (CTG)<sup>n</sup>(CAG)<sup>m</sup> repeats can be correctly repaired, escape repair or undergo error-prone repair. *Nature structural & molecular biology*, **12**, 654-662.

217. Hou, C., Chan, N.L., Gu, L. and Li, G.M. (2009) Incision-dependent and error-free repair of (CAG)(n)/(CTG)(n) hairpins in human cell extracts. *Nature structural & molecular biology*, **16**, 869-875.
218. Chan, N.L., Hou, C., Zhang, T., Yuan, F., Machwe, A., Huang, J., Orren, D.K., Gu, L. and Li, G.M. (2012) The Werner syndrome protein promotes CAG/CTG repeat stability by resolving large (CAG)(n)/(CTG)(n) hairpins. *The Journal of biological chemistry*, **287**, 30151-30156.
219. Panigrahi, G.B., Slean, M.M., Simard, J.P. and Pearson, C.E. (2012) Human mismatch repair protein hMutLalpha is required to repair short slipped-DNAs of trinucleotide repeats. *The Journal of biological chemistry*, **287**, 41844-41850.
220. Zhang, T., Huang, J., Gu, L. and Li, G.M. (2012) In vitro repair of DNA hairpins containing various numbers of CAG/CTG trinucleotide repeats. *DNA repair*, **11**, 201-209.
221. Cloud, K.G., Shen, B., Strniste, G.F. and Park, M.S. (1995) XPG protein has a structure-specific endonuclease activity. *Mutation research*, **347**, 55-60.
222. Hou, C., Zhang, T., Tian, L., Huang, J., Gu, L. and Li, G.M. (2011) The Role of XPG in Processing (CAG)n/(CTG)n DNA Hairpins. *Cell & bioscience*, **1**, 11.
223. Stracker, T.H. and Petrini, J.H. (2011) The MRE11 complex: starting from the ends. *Nature reviews. Molecular cell biology*, **12**, 90-103.
224. Trujillo, K.M., Yuan, S.S., Lee, E.Y. and Sung, P. (1998) Nuclease activities in a complex of human recombination and DNA repair factors Rad50, Mre11, and p95. *The Journal of biological chemistry*, **273**, 21447-21450.
225. Hill, J.W., Hazra, T.K., Izumi, T. and Mitra, S. (2001) Stimulation of human 8-oxoguanine-DNA glycosylase by AP-endonuclease: potential coordination of the initial steps in base excision repair. *Nucleic acids research*, **29**, 430-438.
226. Vidal, A.E., Hickson, I.D., Boiteux, S. and Radicella, J.P. (2001) Mechanism of stimulation of the DNA glycosylase activity of hOGG1 by the major human AP endonuclease: bypass of the AP lyase activity step. *Nucleic acids research*, **29**, 1285-1292.
227. Vidal, A.E., Boiteux, S., Hickson, I.D. and Radicella, J.P. (2001) XRCC1 coordinates the initial and late stages of DNA abasic site repair through protein-protein interactions. *The EMBO journal*, **20**, 6530-6539.

228. Liu, Y., Zhang, H., Veeraraghavan, J., Bambara, R.A. and Freudenreich, C.H. (2004) Saccharomyces cerevisiae flap endonuclease 1 uses flap equilibration to maintain triplet repeat stability. *Molecular and cellular biology*, **24**, 4049-4064.
229. Kao, H.I., Veeraraghavan, J., Polaczek, P., Campbell, J.L. and Bambara, R.A. (2004) On the roles of Saccharomyces cerevisiae Dna2p and Flap endonuclease 1 in Okazaki fragment processing. *The Journal of biological chemistry*, **279**, 15014-15024.
230. Rossi, M.L., Pike, J.E., Wang, W., Burgers, P.M., Campbell, J.L. and Bambara, R.A. (2008) Pif1 helicase directs eukaryotic Okazaki fragments toward the two-nuclease cleavage pathway for primer removal. *The Journal of biological chemistry*, **283**, 27483-27493.
231. Bae, S.H., Bae, K.H., Kim, J.A. and Seo, Y.S. (2001) RPA governs endonuclease switching during processing of Okazaki fragments in eukaryotes. *Nature*, **412**, 456-461.
232. Li, X., Li, J., Harrington, J., Lieber, M.R. and Burgers, P.M. (1995) Lagging strand DNA synthesis at the eukaryotic replication fork involves binding and stimulation of FEN-1 by proliferating cell nuclear antigen. *The Journal of biological chemistry*, **270**, 22109-22112.
233. Vallur, A.C. and Maizels, N. (2010) Complementary roles for exonuclease 1 and Flap endonuclease 1 in maintenance of triplet repeats. *The Journal of biological chemistry*, **285**, 28514-28519.
234. Dizdaroglu, M. (1992) Measurement of radiation-induced damage to DNA at the molecular level. *International Journal of Radiation Biology*, **61**, 175-183.
235. Chatgililoglu, C., Ferreri, C. and Terzidis, M.A. (2011) Purine 5',8-cyclonucleoside lesions: chemistry and biology. *Chemical Society Reviews*, **40**, 1368-1382.
236. Cadet, J., Douki, T., Gasparutto, D. and Ravanat, J.L. (2003) Oxidative damage to DNA: formation, measurement and biochemical features. *Mutation research*, **531**, 5-23.
237. Chatgililoglu, C., D'Angelantonio, M., Kciuk, G. and Bobrowski, K. (2011) New insights into the reaction paths of hydroxyl radicals with 2'-deoxyguanosine. *Chemical research in toxicology*, **24**, 2200-2206.
238. Gates, K.S. (2009) An overview of chemical processes that damage cellular DNA: spontaneous hydrolysis, alkylation, and reactions with radicals. *Chemical research in toxicology*, **22**, 1747-1760.

239. Kropachev, K., Ding, S., Terzidis, M.A., Masi, A., Liu, Z., Cai, Y., Kolbanovskiy, M., Chatgililoglu, C., Broyde, S., Geacintov, N.E. *et al.* (2014) Structural basis for the recognition of diastereomeric 5',8-cyclo-2'-deoxypurine lesions by the human nucleotide excision repair system. *Nucleic acids research*, **42**, 5020-5032.
240. Kuraoka, I., Bender, C., Romieu, A., Cadet, J., Wood, R.D. and Lindahl, T. (2000) Removal of oxygen free-radical-induced 5',8-purine cyclodeoxynucleosides from DNA by the nucleotide excision-repair pathway in human cells. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 3832-3837.
241. Mitra, D., Luo, X., Morgan, A., Wang, J., Hoang, M.P., Lo, J., Guerrero, C.R., Lennerz, J.K., Mihm, M.C., Wargo, J.A. *et al.* (2012) An ultraviolet-radiation-independent pathway to melanoma carcinogenesis in the red hair/fair skin background. *Nature*, **491**, 449-453.
242. Das, R.S., Samaraweera, M., Morton, M., Gascon, J.A. and Basu, A.K. (2012) Stability of N-glycosidic bond of (5'S)-8,5'-cyclo-2'-deoxyguanosine. *Chemical research in toxicology*, **25**, 2451-2461.
243. Kamakura, N., Yamamoto, J., Brooks, P.J., Iwai, S. and Kuraoka, I. (2012) Effects of 5',8-cyclodeoxyadenosine triphosphates on DNA synthesis. *Chemical research in toxicology*, **25**, 2718-2724.
244. Marietta, C., Gulam, H. and Brooks, P.J. (2002) A single 8,5'-cyclo-2'-deoxyadenosine lesion in a TATA box prevents binding of the TATA binding protein and strongly reduces transcription in vivo. *DNA repair*, **1**, 967-975.
245. Andrews, A.D., Barrett, S.F. and Robbins, J.H. (1978) Xeroderma pigmentosum neurological abnormalities correlate with colony-forming ability after ultraviolet radiation. *Proceedings of the National Academy of Sciences of the United States of America*, **75**, 1984-1988.
246. Kuraoka, I., Robins, P., Masutani, C., Hanaoka, F., Gasparutto, D., Cadet, J., Wood, R.D. and Lindahl, T. (2001) Oxygen free radical damage to DNA. Translesion synthesis by human DNA polymerase eta and resistance to exonuclease action at cyclopurine deoxynucleoside residues. *The Journal of biological chemistry*, **276**, 49283-49288.
247. Pednekar, V., Weerasooriya, S., Jasti, V.P. and Basu, A.K. (2014) Mutagenicity and genotoxicity of (5'S)-8,5'-cyclo-2'-deoxyadenosine in Escherichia coli and replication of (5'S)-8,5'-cyclopurine-2'-deoxynucleosides in vitro by DNA polymerase IV, exo-free Klenow fragment, and Dpo4. *Chemical research in toxicology*, **27**, 200-210.

248. You, C., Swanson, A.L., Dai, X., Yuan, B., Wang, J. and Wang, Y. (2013) Translesion synthesis of 8,5'-cyclopurine-2'-deoxynucleosides by DNA polymerases eta, iota, and zeta. *The Journal of biological chemistry*, **288**, 28548-28556.
249. Brooks, P.J. (2008) The 8,5'-cyclopurine-2'-deoxynucleosides: candidate neurodegenerative DNA lesions in xeroderma pigmentosum, and unique probes of transcription and nucleotide excision repair. *DNA repair*, **7**, 1168-1179.
250. Jaruga, P. and Dizdaroglu, M. (2008) 8,5'-Cyclopurine-2'-deoxynucleosides in DNA: mechanisms of formation, measurement, repair and biological effects. *DNA repair*, **7**, 1413-1425.
251. Kirkali, G., de Souza-Pinto, N.C., Jaruga, P., Bohr, V.A. and Dizdaroglu, M. (2009) Accumulation of (5'S)-8,5'-cyclo-2'-deoxyadenosine in organs of Cockayne syndrome complementation group B gene knockout mice. *DNA repair*, **8**, 274-278.
252. Wang, J., Clauson, C.L., Robbins, P.D., Niedernhofer, L.J. and Wang, Y. (2012) The oxidative DNA lesions 8,5'-cyclopurines accumulate with aging in a tissue-specific manner. *Aging Cell*, **11**, 714-716.
253. Paulson, H.L. and Fischbeck, K.H. (1996) Trinucleotide repeats in neurogenetic disorders. *Annual review of neuroscience*, **19**, 79-107.
254. Sinden, R.R. (2001) Neurodegenerative diseases. Origins of instability. *Nature*, **411**, 757-758.
255. Sinden, R.R., Potaman, V.N., Oussatcheva, E.A., Pearson, C.E., Lyubchenko, Y.L. and Shlyakhtenko, L.S. (2002) Triplet repeat DNA structures and human genetic disease: dynamic mutations from dynamic DNA. *Journal of biosciences*, **27**, 53-65.
256. Lai, Y., Zhang, Z., Li, J., Sun, D., Zhou, Y., Jiang, T., Han, Y., Huang, L., Zhu, Y., Li, X. *et al.* (2013) EGFR mutations in surgically resected fresh specimens from 697 consecutive Chinese patients with non-small cell lung cancer and their relationships with clinical features. *International Journal of Molecular Sciences*, **14**, 24549-24559.
257. Xu, M., Gabison, J. and Liu, Y. (2013) Trinucleotide repeat deletion via a unique hairpin bypass by DNA polymerase beta and alternate flap cleavage by flap endonuclease 1. *Nucleic acids research*, **41**, 1684-1697.
258. McMurray, C.T. (2010) Mechanisms of trinucleotide repeat instability during human development. *Nature reviews. Genetics*, **11**, 786-799.

259. Nelson, K.A. and Witte, J.S. (2002) Androgen receptor CAG repeats and prostate cancer. *American journal of epidemiology*, **155**, 883-890.
260. McMurray, C.T. (1999) DNA secondary structure: a common and causative factor for expansion in human disease. *Proceedings of the National Academy of Sciences of the United States of America*, **96**, 1823-1825.
261. Liu, Y. and Bambara, R.A. (2003) Analysis of human flap endonuclease 1 mutants reveals a mechanism to prevent triplet repeat expansion. *The Journal of biological chemistry*, **278**, 13728-13739.
262. Kovtun, I.V. and McMurray, C.T. (2007) Crosstalk of DNA glycosylases with pathways other than base excision repair. *DNA repair*, **6**, 517-529.
263. Bogdanov, M.B., Andreassen, O.A., Dedeoglu, A., Ferrante, R.J. and Beal, M.F. (2001) Increased oxidative damage to DNA in a transgenic mouse model of Huntington's disease. *Journal of neurochemistry*, **79**, 1246-1249.
264. Liu, Y., Prasad, R., Beard, W.A., Hou, E.W., Horton, J.K., McMurray, C.T. and Wilson, S.H. (2009) Coordination between polymerase beta and FEN1 can modulate CAG repeat expansion. *The Journal of biological chemistry*, **284**, 28352-28366.
265. Xu, M., Lai, Y., Torner, J., Zhang, Y., Zhang, Z. and Liu, Y. (2014) Base excision repair of oxidative DNA damage coupled with removal of a CAG repeat hairpin attenuates trinucleotide repeat expansion. *Nucleic acids research*, **42**, 3675-3691.
266. Beard, W.A., Shock, D.D., Batra, V.K., Pedersen, L.C. and Wilson, S.H. (2009) DNA polymerase beta substrate specificity: side chain modulation of the "A-rule". *The Journal of biological chemistry*, **284**, 31680-31689.
267. Maga, G., van Loon, B., Crespan, E., Villani, G. and Hubscher, U. (2009) The block of DNA polymerase delta strand displacement activity by an abasic site can be rescued by the concerted action of DNA polymerase beta and Flap endonuclease 1. *The Journal of biological chemistry*, **284**, 14267-14275.
268. Villani, G., Hubscher, U., Gironis, N., Parkkinen, S., Pospiech, H., Shevelev, I., di Cicco, G., Markkanen, E., Syvaaja, J.E. and Tanguy Le Gac, N. (2011) In vitro gap-directed translesion DNA synthesis of an abasic site involving human DNA polymerases epsilon, lambda, and beta. *The Journal of biological chemistry*, **286**, 32094-32104.
269. Wiegand, R.C., Godson, G.N. and Radding, C.M. (1975) Specificity of the S1 nuclease from *Aspergillus oryzae*. *The Journal of biological chemistry*, **250**, 8848-8855.



270. Liu, Y., Kao, H.I. and Bambara, R.A. (2004) Flap endonuclease 1: a central component of DNA metabolism. *Annual review of biochemistry*, **73**, 589-615.
271. Vallur, A.C. and Maizels, N. (2010) Complementary roles for exonuclease 1 and Flap endonuclease 1 in maintenance of triplet repeats. *The Journal of biological chemistry*, **285**, 28514-28519.

## VITA

### MENG XU

2005-2009	B.E. of Biotechnology Jiangnan University Wuxi, Jiangsu, China
2010-2011	Teaching Assistant Florida International University Miami, Florida
2011-2013	Research Assistant Florida International University Miami, Florida
2013-2014	FIU Dissertation Year Fellow and Doctoral Candidate Florida International University Miami, Florida

### PUBLICATIONS AND PRESENTATIONS

Zhang, Y., Chen, S., Xu, M., Cavoco-Paulo, A., Wu, J., and Chen, J. (2010) Characterization of *Thermobifida fusca* Cutinase-CBM Fusion Proteins and Their Potential Application in Bioscouring. *Appl Environ Microbiol.* 76(23):7896.

Xu, M., Gabison, J. and Liu, Y. (2013) Trinucleotide Repeat Deletion via a Unique Hairpin Bypass by DNA polymerase  $\beta$  and Alternate Flap Cleavage by Flap endonuclease 1. *Nucl. Acids Res.* 41 (3): 1684-1697.

Lai, Y., Xu, M. \*, Zhang, Z. and Liu, Y. (2013) Instability of CTG Repeats Is Governed by the Position of a DNA Base Lesion through Base Excision Repair. *PLoS One.* 8 (2): e56960.

Xu, M., Lai, Y., Torner, J., Osemota, E., Zhang, Y. and Liu, Y. (2014) Removal of a CAG Repeat Hairpin by Base Excision Repair Prevents Trinucleotide Repeat Expansion. *Nucl. Acids Res.* 42 (6): 3675-91

Xu, M., Lai, Y., Torner, J., Zhang, Y., Zhang, Z. and Liu, Y. (2013) Base Excision Repair of Oxidative DNA Damage is Coupled with Removal of a CAG Repeat Hairpin to Attenuate Trinucleotide Repeat Expansion, EMGS 44th Annual Meeting Monterey, CA, September 22-25, 2013. (Oral presentation)

Xu, M., Jiang, Z., Terzidis, M., Masi, A., Chatgialloglu, C. and Liu, Y. (2014) Bypass of a 5',8-cyclo-2'-deoxynucleoside by DNA polymerase  $\beta$  leads to trinucleotide repeat

deletion, FIU's annual Scholarly Forum during Graduate Student Appreciation Week (GSAW), March 31, 2014, Florida International University, Miami, FL (Oral presentation)

Xu, M., Jiang, Z., Terzidis, M., Masi, A., Chatgialiloglu, C. and Liu, Y. (2014). A 5',8-cyclo-deoxypurine lesion induces trinucleotide repeat deletion via a unique lesion bypass by DNA polymerase  $\beta$ . EMGS 45th Annual Meeting, Orlando, FL, September 13-17, 2014.

Xu, M., Jiang, Z., Terzidis, M., Masi, A., Chatgialiloglu, C. and Liu, Y. (2014) Bypass of a 5',8-cyclo-2'-deoxynucleoside by DNA polymerase  $\beta$  leads to trinucleotide repeat deletion, Gordon Research Conferences, DNA Damage, Mutation & Cancer Ventura, CA, March 16-21, 2014.

Xu, M., Jiang, Z., Terzidis, M., Masi, A., Chatgialiloglu, C. and Liu, Y. (2014) Bypass of a 5',8-cyclo-2'-deoxynucleoside by DNA polymerase  $\beta$  leads to trinucleotide repeat deletion, Poster session of Graduate Visitation Day, March 7, 2014, Department of Chemistry and Biochemistry, College of Arts and Sciences, Florida International University, Miami, FL

Xu, M., Lai, Y., Torner, J., Zhang, Y., Zhang, Z. and Liu, Y. (2013) Base Excision Repair of Oxidative DNA Damage is Coupled with Removal of a CAG Repeat Hairpin to Attenuate Trinucleotide Repeat Expansion, EMGS 44th Annual Meeting Monterey, CA, September 22-25, 2013.

Xu, M. and Liu, Y. (2012) Trinucleotide Repeats are Deleted during base excision repair via A Unique Hairpin Bypass by DNA Polymerase  $\beta$  and FEN1 Cleavage, Gordon Research Conferences, DNA Damage, Mutation & Cancer Ventura, CA, March 25-30, 2012.

Xu, M. and Liu, Y. (2012) Trinucleotide Repeats are Deleted during base excision repair via A Unique Hairpin Bypass by DNA Polymerase  $\beta$  and FEN1 Cleavage, Poster session of Graduate Visitation Day, March 23, 2012, Department of Chemistry and Biochemistry, College of Arts and Sciences, Florida International University, Miami, FL

Xu, M. and Liu, Y. (2011) Mechanisms Underlying CAG Repeat Instability during Oxidative DNA Base Damage and Repair, Gordon Research Conferences-Genetic Toxicology, July 10-15, 2011, Lucca (Barga), Italy

Xu, M. and Liu, Y. (2011) Mechanisms Underlying CAG Repeat Instability during Oxidative DNA Base Damage and Repair, Poster session of Graduate Visitation Day, March 4, 2011, Department of Chemistry and Biochemistry, College of Arts and Sciences, Florida International University, Miami, FL