5-29-2018

Trinucleotide Repeat Instability Modulated by DNA Repair Enzymes and Cofactors

Yaou Ren

Florida International University, yren002@fiu.edu

DOI: 10.25148/etd.FIDC006873

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

Part of the [Biochemistry Commons](#), and the [Molecular Biology Commons](#)

Recommended Citation

Ren, Yaou, "Trinucleotide Repeat Instability Modulated by DNA Repair Enzymes and Cofactors" (2018). *FIU Electronic Theses and Dissertations*. 3762.

<https://digitalcommons.fiu.edu/etd/3762>

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.

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

TRINUCLEOTIDE REPEAT INSTABILITY MODULATED BY DNA REPAIR
ENZYMES AND COFACTORS

A dissertation submitted in partial fulfillment of
the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

BIOCHEMISTRY

by

Yaou Ren

2018

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

This dissertation, written by Yaou Ren, and entitled Trinucleotide Repeat Instability Modulated by DNA Repair Enzymes and Cofactors, 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

Yuk-Ching Tse-Dinh

Xiaotang Wang

Yuan Liu, Major Professor

Date of Defense: May 29, 2018

The dissertation of Yaou Ren is approved.

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

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

Florida International University, 2018

© Copyright 2018 by Yaou Ren

All rights reserved.

DEDICATION

I dedicate this dissertation to my parents, Zhilin Ren and Jihui Li, for their love,
support, and encouragement.

ACKNOWLEDGMENTS

I would like to express my gratitude to the many people who have had a part in helping me to complete this work. First of all, I would like to thank my advisor, Dr. Yuan Liu, for her unending assistance, advice, and support throughout this process. Under her excellent mentorship and guidance, I have gained invaluable experience in scientific research, writing, and presentation. She has encouraged and supported me at every step of this endeavor.

I would like to thank my committee members, Dr. Yuk-Ching Tse-Dinh, Dr. Irina AgoulNIK, and Dr. Xiaotang Wang, for their advice, guidance, and support throughout my Ph.D. training and career development.

I would like to thank Dr. Yanhao Lai, whose unselfish assistance and guidance in the lab supported my research at every step. She has offered patient instruction and support throughout the process of my research training. I would also like to thank Dr. Zhongliang Jiang, Eduardo Laverde, and Pawlos Tsegay as well as all other past and present members of the Liu Research Lab, for their help and support.

Finally, I would like to thank Dr. Samuel H. Wilson, Genome Integrity and Structural Biology Laboratory, National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health (NIH) for generously providing us with expression vectors of BER enzymes.

ABSTRACT OF THE DISSERTATION
TRINUCLEOTIDE REPEAT INSTABILITY MODULATED BY DNA REPAIR
ENZYMES AND COFACTORS

by

Yaou Ren

Florida International University, 2018

Miami, Florida

Professor Yuan Liu, Major Professor

Trinucleotide repeat (TNR) instability including repeat expansions and repeat deletions is the cause of more than 40 inherited incurable neurodegenerative diseases and cancer. TNR instability is associated with DNA damage and base excision repair (BER). In this dissertation research, we explored the mechanisms of BER-mediated TNR instability via biochemical analysis of the BER protein activities, DNA structures, protein-protein interaction, and protein-DNA interaction by reconstructing BER *in vitro* using synthesized oligonucleotide TNR substrates and purified human proteins. First, we evaluated a germline DNA polymerase β (pol β) polymorphic variant, pol β R137Q, in leading TNR instability-mediated cancers or neurodegenerative diseases. We find that the pol β R137Q has slightly weaker DNA synthesis activity compared to that of wild-type (WT) pol β . Because of the similar abilities between pol β R137Q and WT pol β in bypassing a template loop structure, both pol β R137Q and WT pol β induces similar amount of repeat deletion. We conclude that the slightly weaker DNA synthesis activity of pol β R137Q does not alter the TNR instability compared to that of WT pol β , suggesting that the pol β R137Q carriers do not have an altered risk in developing TNR

instability-mediated human diseases. We then investigated the role of DNA synthesis activities of DNA polymerases in modulating TNR instability. We find that pol β Y265C and pol ν with very weak DNA synthesis activities predominantly promote TNR deletions. We identify that the sequences of TNRs may also affect DNA synthesis and alter the outcomes of TNR instability. By inhibiting the DNA synthesis activity of pol β using a pol β inhibitor, we find that the outcome of TNR instability is shifted toward repeat deletions. The results provide the direct evidence that DNA synthesis activity of DNA polymerases can be utilized as a potential therapeutic target for treating TNR expansion diseases. Finally, we explored the role of post-translational modification (PTM) of proliferating cell nuclear antigen (PCNA) on TNR instability. We find that ubiquitinated PCNA (ub-PCNA) stimulates Fanconi associated nuclease 1 (FAN1) 5'-3' exonucleolytic activities directly on hairpin structures, coordinating flap endonuclease 1 (FEN1) in removing difficult secondary structures, thereby suppressing TNR expansions. The results suggest a role of mono-ubiquitination of PCNA in maintaining TNR stability by regulating nucleases switching. Our results suggest enzymatic activities of DNA polymerases and nucleases and the regulation of the activities by PTM play important roles in BER-mediated TNR instability. This research provides the molecular basis for future development of new therapeutic strategies for prevention and treatment of TNR-mediated neurodegenerative diseases.

TABLE OF CONTENTS

CHAPTER	PAGE
INTRODUCTION	1
A. DNA Damage and DNA Repair.....	1
B. Mechanism of Base Excision Repair, Repair Enzymes, and Cofactors.....	4
C. Trinucleotide Repeat Instability and Human Diseases	12
D. Trinucleotide Repeat Instability and DNA Replication	17
E. Trinucleotide Repeat Instability and Base Excision Repair	21
OVERVIEW	37
CHAPTER 1: MODULATION OF TRINUCLEOTIDE REPEAT INSTABILITY	
BY DNA POLYMERASE β POLYMORPHIC VARIANT R137Q	39
ABSTRACT.....	39
INTRODUCTION	40
MATERIALS AND METHODS	43
RESULTS	51
DISCUSSION	72
CHAPTER 2: WEAK DNA SYNTHESIS BY DNA POLYMERASES	
PROMOTES TRINUCLEOTIDE REPEAT DELETIONS DURING BASE	
EXCISION REPAIR.....	76
ABSTRACT.....	76
INTRODUCTION	77
MATERIALS AND METHODS	80
RESULTS	86
DISCUSSION	114
CHAPTER 3: MONO-UBIQUITINATION OF PROLIFERATING CELL	
NUCLEAR ANTIGEN MODULATES TRINUCLEOTIDE REPEAT	
INSTABILITY BY COORDINATING WITH FANCONI ANEMIA-	
ASSOCIATED NUCLEASE 1 AND FLAP ENDONUCLEASE 1	119
ABSTRACT.....	119
INTRODUCTION	120
MATERIALS AND METHODS.....	122
RESULTS	129
DISCUSSION	157
SUMMARY	161
REFERENCES	165
VITA	196

LIST OF FIGURES

FIGURE	PAGE
INTRODUCTION	
I.1 DNA damage and repair	2
I.2 Removal of a base lesion by a DNA glycosylase	5
I.3 DNA polymerase β	6
I.4 Base excision repair pathway	8
I.5 Trinucleotide repeat expansions in encoding and noncoding regions	13
I.6 Protein aggregation and cellular toxicity	15
I.7 GAA repeat expansions suppress frataxin expression	16
I.8 Formation of secondary structures on TNRs	18
I.9 Secondary structures disrupt DNA synthesis of DNA polymerases	19
I.10 Replication restart leads to TNR instability	20
I.11 BER-mediated TNR expansions	22
I.12 BER-mediated toxic oxidation cycle	24
I.13 MSH2-MSH3 coordinates with BER enzymes to promote TNR expansions	26
I.14 Locations of base lesions govern BER-mediated TNR instability	29
I.15 APE1 suppresses TNR expansions by facilitating removal of upstream repeats	31
I.16 Mono-ubiquitinated PCNA	34
I.17 Ubiquitin may prevent interaction between FEN1 and PCNA via PIP box	35

CHAPTER 1

1.1 Pol β R137Q variant exhibits weak DNA synthesis activity during BER of a base lesion in context with (CAG) ₂₀ and (CTG) ₂₀ repeats.....	54
1.2 Pol β R137Q variant exhibits similar gap-filling synthesis activity compared to that of wild-type pol β	56
1.3 PCNA does not affect DNA synthesis activity of pol β WT and pol β R137Q during BER of a base lesion on (CAG) ₂₀ repeats	58
1.4 Pol β R137Q leads to weak FEN1 cleavage during BER in the context of (CAG) ₂₀ repeats	61
1.5 Effects of pol β R137Q on the instability of (CAG) ₂₀ and (CTG) ₂₀ repeats.....	63
1.6 S1 nuclease probing of loops formed in (CAG) ₂₀ repeats with a THF located at the first CAG of the (CAG) ₂₀ repeat substrate in the absence and presence of pol β WT or pol β R137Q DNA synthesis	67
1.7 S1 nuclease probing of loops formed in (CAG) ₂₀ repeats with a THF located at the tenth CAG of the (CAG) ₂₀ repeat substrate in the absence and presence of pol β WT or pol β R137Q DNA synthesis	71
1.8 S1 nuclease digestion of the (CAG) ₂₀ substrates with a nick located after first CAG or after the tenth CAG	72

CHAPTER 2

2.1 Pol β Y265C exhibits weak DNA synthesis activity during BER of a base lesion in the context of (CAG) ₂₀ repeats.....	88
2.2 Pol β Y265C promotes (CAG) ₂₀ repeat deletions.....	91
2.3 Pol β Y265C leads to weak FEN1 cleavage during BER in the context of (CAG) ₂₀ repeats	94
2.4 Weak DNA synthesis of pol ν promotes (CAG) ₂₀ repeat deletions.....	96
2.5 Pol θ does not induce (CAG) ₂₀ repeat deletions	98
2.6 Weak DNA synthesis of pol β Y265C promotes CTG repeat deletion during BER in the context of (CTG) ₂₀ repeats	100

2.7 Weak DNA synthesis of pol ν promotes (CGG) ₂₀ repeat deletions.....	104
2.8 Pol θ induces (GAA) ₂₀ repeat deletions.....	107
2.9 DNA synthesis activity of pol β is inhibited by a pol β inhibitor, NSC666719	109
2.10 Inhibition of DNA synthesis activity of pol β promotes (CAG) ₂₀ repeat deletions.....	110
2.11 Pol β Y265C does not significantly alter (GT) ₃₀ repeat instability.....	113
2.12 Pol β Y265C does not significantly alter (TTAGGG) ₁₀ repeat instability	114

CHAPTER 3

3.1 P11 phosphocellulose cation exchange chromatography of PCNA	131
3.2 DEAE ion-exchange chromatography of PCNA	131
3.3 Phenyl sepharose chromatography of PCNA	133
3.4 Q sepharose anion exchange chromatography of PCNA.....	135
3.5 The first Heparin chromatography of PCNA.....	137
3.6 The second Heparin chromatography of PCNA	137
3.7 Mono Q anion exchange chromatography of PCNA.....	139
3.8 Expression of UBE1 and UbcH5c S22R	141
3.9 The Mono Q anion exchange chromatography of UBE1	141
3.10 The SP sepharose cation exchange chromatography of UbcH5c S22R	142
3.11 Mono-ubiquitination of PCNA	143
3.12 Q sepharose anion exchange chromatography of ub-PCNA	145
3.13 Phenyl sepharose chromatography of ub-PCNA	146
3.14 Mono Q anion exchange chromatography of ub-PCNA.....	147
3.15 The 5'-flap and hairpin structures formed during BER in the context of TNR tracts.....	149

3.16 FAN1 and FEN1 have different substrate cleavage patterns in cleaving TNRs.....	150
3.17 Ub-PCNA stimulates FAN1 5'-3' exonucleolytic activity, but not its 5'-endonucleolytic activity.....	152
3.18 PCNA stimulates FEN1 5'-flap endonucleolytic activity on TNRs	154
3.19 Effects of FAN1 and FEN1 on TNR instability.....	155
3.20 Ub-PCNA facilitates FAN1 and FEN1 in removing a big hairpin	157

ABBREVIATIONS AND ACRONYMS

°C	degree Celsius
A	adenine
APE1	AP Endonuclease 1
AP site	apurinic/apyrimidinic site
AR	androgen receptor
ATP	adenosine triphosphate
BER	base excision repair
Bis-Tris	2,2-Bis(hydroxymethyl)-2,2',2'-nitrilotriethanol
BRCT	BRCA1 C terminus domain
BSA	bovine serum albumin
C (Cys)	cysteine
CDK	cyclin-dependent kinase
cdPu	5',8-cyclo-2'-deoxypurines
Cordycepin	3'-deoxyadenosine
DM	mytonic dystrophy
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
dRP	deoxyribose phosphate
dsDNA	double-stranded DNA
DTT	dithiothreitol
8-oxoG	8-oxoguanine

<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	ethylenediaminetetraacetic acid
EME1	essential meiotic structure-specific endonuclease 1
EXO1	endonuclease 1
5',8-cdA	5',8-cyclo-2'-deoxyadenosine
5'-dRP	5'-deoxyribose phosphate
5-hU	5-hydroxyuracil
FAN1	Fanconi associated nuclease 1
FEN1	flap endonuclease 1
FMR1	fragile X mental retardation 1
FMSF	phenylmethanesulfonyl fluoride
FRDA	Friedreich's ataxia
FXN	frataxin
FXS	fragile X syndrome
G	guanine
Gln	glutamine (Q)
HCl	hydrogen chloride
HD	Huntington's disease
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HMGB1	high-mobility group protein B1
HR	homologous recombination
H ₂ O ₂	hydrogen peroxide

I (Ile)	Isoleucine
ICL	interstrand crosslink
IPTG	isopropyl β -D-1-thiogalactopyranoside
IR	ionizing radiation
K (Lys)	lysine
KCl	potassium chloride
L (Leu)	leucine
LB	lysogeny broth
LIG I	DNA ligase I
LIG III	DNA ligase III
LP-BER	long patch-base excision repair
MBD4	methyl-binding domain glycosylase 4
MD	myotonic dystrophy
MEF	mouse embryonic fibroblast
MgCl ₂	magnesium chloride
MLH	MutL homologue
MMR	mismatch repair
MMS	methyl methanesulfonate
MRN	Mre11-Rad50-Nbs1 complex
MSH	MutS homologue
Mus81	MUS81 structure-specific endonuclease subunit
MutS β	MSH2/MSH3

MW	molecular weight
MYH	MutY homologue
N	nitrogen
NaCl	sodium chloride
NaH ₂ PO ₄	sodium dihydrogen phosphate
NaOH	sodium hydroxide
N ⁷ -meG	N ⁷ -methylguanine
NEIL	endonuclease VIII-like
NER	nucleotide excision repair
NHEJ	nonhomologous end joining
NP-40	nonidet P-40
nM	nanomolar
nt	nucleotide
O	oxygen
O ₂ ^{-•}	superoxide
OH [•]	hydroxyl radical
O ⁶ -meG	O ⁶ -methylguanine
¹ O ₂	singlet oxygen
OGG1	8-oxoguanine DNA glycosylase
OPMD	oculopharyngeal muscular dystrophy
P (Pro)	proline
PAGE	polyacrylamide gel electrophoresis

PAH	polycyclic aromatic hydrocarbon
PARP1	poly(ADP-ribose) polymerase 1
PCNA	proliferating cell nuclear antigen
pH	potential of hydrogen
pI	isoelectric point
PIP	PCNA-interacting protein
PMSF	phenylmethanesulfonyl fluoride
pol β	DNA polymerase β
pol δ	DNA polymerase δ
pol ϵ	DNA polymerase ϵ
pol η	DNA polymerase η
pol κ	DNA polymerase κ
pol ι	DNA polymerase ι
pol ν	DNA polymerase ν
pol θ	DNA polymerase θ
polyA	polyalanine
polyQ	polyglutamine
PTM	post-translational modification
Q	glutamine
R (Arg)	arginine
REV1	REV1, DNA directed polymerase
RNA	ribonucleic acid

ROS	reactive oxygen species
RPA	replication protein A
S (Ser)	serine
SBMA	spinal and bulbar muscular atrophy
SCA10	spinocerebellar ataxia type 10
6-FAM	6-carboxyfluorescein
SMUG 1	single-strand selective monofunctional uracil DNA glycosylase
SP-BER	short patch base excision repair
SPD	synpolydactyly type II
ssDNA	single-stranded DNA
T	thymine
TDG	thymine DNA glycosylase
Tg	thymine glycol
THF	tetrahydrofuran
3meA	3-methyladenine
TL	translesion synthesis
TNR	trinucleotide repeat
Tris	tris(Hydroxymethyl)aminomethane
2-hA	2-hydroxyuracil
XPG	xeroderma pigmentosum complementation group G
XRCC1	X-ray repair cross-complementing protein 1

U	uracil
UbcH5c	ubiquitin-conjugating enzyme H5c
UBE1	ubiquitin activating enzyme E1
UBM	ubiquitin-binding module
UBZ	ubiquitin-binding domain
ub-PCNA	ubiquitinated PCNA
UDG	uracil DNA glycosylase
μM	micromolar
UNG	uracil DNA N-glycosylase
UTR	untranslated region
UV	ultraviolet
WT	wild-type
XPG	xeroderma pigmentosum complementation group G
Zic2	zinc-finger protein of the cerebellum 2
ZnCl ₂	zinc chloride

INTRODUCTION

A. DNA Damage and DNA Repair

Deoxyribonucleic acid (DNA) is a long chain of nucleotides that carry the genetic information of all living organisms [15-17]. The integrity and stability of DNA in human cells are of particular importance because it determines cell survival and determines proper functioning, growth, and reproduction of cells [15-17]. However, the integrity of the DNA molecule is often challenged by many assaults from endogenous and exogenous sources [18-25]. In cells, DNA is not stable as it is prone to self-depurination: spontaneous loss of purines Adenine (A) and Guanine (G) at neutral pH and physiological temperature, resulting in an abasic or apurinic/apyrimidinic (AP) site [20, 26]. The AP site can further form single-strand DNA (ssDNA) breaks. Moreover, cellular metabolic processes and other biochemical reactions can generate reactive oxygen species (ROS) that damage DNA [25]. The ROS include superoxide ($O_2^{\cdot-}$), hydroxyl radicals ($OH^{\cdot}$), singlet oxygen (1O_2), and hydrogen peroxide (H_2O_2) that can oxidize DNA bases, causing ssDNA and double-strand DNA (dsDNA) breaks [25]. The ROS can also indirectly attack DNA through oxidized polyunsaturated fatty acid residues of phospholipids or oxidized amino acid residue. The oxidized lipid or proteins attack DNA molecules, resulting in lipid-DNA crosslinks and peptide-DNA crosslinks, respectively [25]. On the other hand, endogenous agents such as X-ray and UV radiations, plant toxins, environmental pollutants, and other chemicals can cause different types of DNA lesions [18, 19, 21-24]. For example, both ionizing radiation and UV radiation (UV A light) can damage DNA by generating free radicals, which oxidize DNA bases [23, 24].

Meanwhile, ionizing radiation can also directly cause ssDNA and dsDNA breaks [24]; whereas, UV radiation (UV B light) can induce crosslinks between two adjacent thymines or cytosines, generating pyrimidine dimers [23]. Exposure to the algal toxins, such as okadaic acid from dinoflagellates, can also cause oxidated base lesions, DNA strand breaks, and hydroxyl-deoxyguanine adducts [22]. Environmental pollutants, such as bromate in drinking water, can induce 7,8-dihydro-8-oxoguanine (8-oxoG), an oxidized base lesion [21]. Polycyclic aromatic hydrocarbon (PAH) in the photochemical smog (air pollution) and tobacco smoke can cause bulky DNA adducts [18, 19].

Figure I.1 DNA damage and repair [10]

To maintain the integrity of the DNA, the cells have evolved to respond precisely to repair such DNA lesions via several different mechanisms (Figure I.1). Small modifications of DNA bases, including oxidized bases, alkylated bases, and deaminated bases, can be repaired through base excision repair (BER) [27, 28]. Aside from that, BER can also repair AP lesions and ssDNA breaks [29]. However, to repair mismatched DNA bases requires the mismatch repair (MMR), which also repairs small DNA loops (12 nucleotides or smaller), an error caused by misalignment during DNA synthesis [30-32]. Moreover, nucleotide excision repair (NER) repairs larger and more complex lesions, including pyrimidine dimers and other bulky DNA adducts [23]. The dsDNA breaks can be repaired either by homologous recombination (HR), which can faithfully restore the genetic information [33], or non-homologous end joining (NHEJ), which can introduce mutations [34]. However, not all the lesions can be repaired through one single repair mechanism. For example, to repair a large DNA loop (26 nucleotides), it requires coordination between the NER and the MMR [35]. To repair the interstrand crosslinks (ICLs), it requires crosstalk between the Fanconi anemia (FA) DNA repair pathway and the NER [36]. Occasionally, some lesions escape the repair; cells carry out the translesion DNA synthesis (TL) to prevent catastrophic cellular events, such as replication forks collapse and cell death [37].

The most frequently occurring DNA damage is the DNA-base lesion. For example, the 8-oxoGs frequently occur in the cells: the naturally occurring 8-oxoGs *in vivo* have an estimation of 0.53 and 4.01 per 10^6 in single cell because guanine is particularly susceptible to oxidation because of its low-reduction potential [38-41]. Meanwhile, AP lesions are readily formed through self-depurination with an estimation

of 10,000 AP lesions occurring in a 24-hour period [26]. Besides spontaneously occurring 8-oxoG and AP lesions, they can be easily induced by the insults from environmental factors, such as UV radiation (UV A light), ionizing radiation, oxidative stress, industrial pollutants, and plant toxins [23, 24, 38]. Accumulated 8-oxoG can be very mutagenic: it tends to promote G to T and A to C substitution [42]. Meanwhile, accumulated AP lesions can be both mutagenic and lethal [43, 44]. Apurinic/apyrimidinic lesions can induce -1 frame shift mutation, a mutation with a deletion of one nucleotide during DNA synthesis [43]. Apurinic/apyrimidinic lesions can also potentially block DNA replication and transcription [44], leading to cell death. However, all such lesions can be repaired by BER [20, 38]. Base excision repair is one of the most efficient DNA repair pathways in both dividing and non-dividing cells [45]. In addition to repair 8-oxoG and AP lesions, BER also specializes in repairing other small non-bulky base lesions, such as misincorporated U and ssDNA breaks, which are also common lesions in the cells [20, 46]. Therefore, a highly efficient BER ensures the integrity and stability of the genome.

B. Mechanism of Base Excision Repair, Repair Enzymes, and Cofactors

Efficient BER is carried out by a plethora of BER enzymes that process and repair the lesions in a coordinated manner [8, 27, 28]. Initial recognition and removal of base lesions determine the success of BER. There are eleven DNA glycosylases that specialize in recognition and the removal of different kinds of base lesions [3]. The uracil DNA glycosylase 1 (UNG1), the uracil DNA glycosylase 2 (UNG2), and the single-strand-specific monofunctional uracil DNA glycosylase (SMUG1) specialize in removal of the uracil lesions in the ssDNA and dsDNA [3]. The thymine DNA glycosylase (TDG) and

the methyl-binding domain glycosylase 4 (MBD4) remove the thymine from G-T mismatches [3]. The MutY homolog DNA glycosylase (MYH) removes adenosine (A) opposite 8-oxoG or 2-hA opposite G; while the 8-OxoG DNA glycosylase (OGG1) removes the 8-oxoG [3]. The methylpurine glycosylase (MPG) removes the alkylated purines like 3-meA, hypoxanthin, etc [3]. The endonuclease III-like (NTHL) 1, the NTHL2, and the NTHL3 carry out excision on oxidized, ring-fragmented or saturated pyrimidines, such as 5-hydroxyuracil (5-hU), thymine glycol (Tg), etc.[3].

The DNA glycosylases “scan” diligently on the DNA and search for the base lesions [3, 47]. Once they identify a lesion, they flip the damaged base out from the DNA helix and fit the modified base into their active site for further verification (Figure I.2) [3]. After verification, they carry out the nucleophilic attack on the N-glycosidic bond [3]. Removal of the base lesions leads to AP sites [48]. The AP sites from both glycosylase cleavage and DNA damage are further processed by AP endonuclease 1 (APE1): APE1 exercises cleavage on phosphodiester bond at the 5'-side of AP site, and the cleavage results in single strand DNA break with a 3'-hydroxyl group located at the upstream strand and a 5'-deoxyribose phosphate (5'-dRP) group located at the downstream strand [48].

Figure I.2 Removal of a base lesion by a DNA glycosylase [3]

After APE1 cleavage, BER diverges into two sub-pathways: the short-patch (SP) BER and the long-patch (LP) BER [8, 27, 28]. The SP-BER requires a native 5'-dRP group, and it leads to a single-nucleotide replacement [8, 27, 28]. The LP-BER comes into effect when the deoxyribose of the 5'-dRP is modified, and the LP-BER is a more complicated repair pathway with the requirement of coordination of additional enzymes; the repair leads to two or more nucleotides replacements [8, 27, 28]. The ability of polymerase β (pol β), a bifunctional enzyme consists of a 31 kDa polymerase domain and an 8 kDa deoxyribose phosphate (dRP) domain, in removing the deoxyribose of the 5'-dRP group determines whether BER subjects to SP-BER or LP-BER [49].

Figure I.3 DNA polymerase β [12]

When the 5'-dRP is native, pol β removes the deoxyribose through β -elimination with its dRP lyase domain [49]. The cleavage leaves a one-nucleotide gap with a 5'-phosphate group [49]. The pol β further incorporates one nucleotide through its polymerase domain, leaving a nick substrate [8, 27, 28]. The nick is then sealed by X-ray cross-complementing group 1 (XRCC1)-DNA ligase III α (LIG III α) complex [8, 27, 28],

leading to completion of SP-BER. The SP-BER results in one-nucleotide replacement (Figure I.4). The ability of pol β in processing 5'-dRP quickly (at least 20-fold faster than the polymerase gap filling) also contributes to the efficiency of SP-BER [50]. In contrast, when the 5'-dRP group is modified, pol β cannot process the modified deoxyribose, subjecting BER to LP-BER [8, 27, 28, 49]. Long-patch BER can be further divided into two sub-pathways: Pol β /Flap endonuclease 1 (FEN1)-mediated LP-BER ("Hit and Run" BER) and strand-displacement-mediated LP-BER (Figure I.4). During hit and run pathway, pol β and FEN1 coordinate with high efficiency to repair the damage. The pol β first incorporates one nucleotide, and then FEN1 removes one nucleotide that is associated with the modified 5'-dRP at the downstream strand [51]. The cleavage by pol β leaves another one nucleotide gap, and pol β continues to incorporate a second nucleotide, which then leaves a nick substrate [51]. The ligase I (LIG I) comes and seals the nick, completing the repair with two-nucleotide replacement [8]. The hit-and-run LP sub-pathway is the most efficient LP-BER pathway when the downstream 5'-dRP is modified, and the pathway requires good coordination between pol β and FEN1. However, pol β is not essential for the strand-displacement-mediated LP-BER, the polymerases involved in the repair can be DNA polymerase δ (pol δ) or DNA polymerase ϵ (pol ϵ) [8, 27, 28]. Each of the polymerases synthesize continuously, displacing the downstream strand and facilitating the formation of a long flap [8, 27, 28]. The long flap is then captured and removed by FEN1, and the removal results in a nick that subsequently sealed by LIG I, which finishes the repair [8, 27, 28]. The strand-displacement-mediated LP-BER is the least efficient BER sub-pathway, and it usually involves 3-10 nt replacement [8, 52].

Figure I.4 Base excision repair pathway [8]

The figure illustrates the 1-nucleotide-patch BER on the left, the “Hit and Run” BER in the middle, and the long-patch BER subpathways on the right.

Base excision repair is a sophisticated and complex pathway. Proper functioning and coordination of BER enzymes guarantee a precise and efficient repair of DNA base

lesions. The DNA glycosylases and APE1 play vital roles in initiation steps of BER, and the proper functioning of pol β and FEN1 guarantee the efficiency of BER. For example, APE1 is essential for cell survival and its knockout is embryonic lethal [53]. Although OGG1 is not necessary for survival, it plays a vital role in maintaining genomic stability: the OGG1 knockout mice presented an increased risk for cancer development [54]. On the other hand, pol β also plays a critical role in survival and growth. Disruption of the pol β gene causes massive apoptotic events among the postmitotic neuron cells, resulting in death of neonatal mice [55]. This indicates that pol β is essential for cell survival. Moreover, although the embryonic mouse fibroblasts lack of pol β are viable, they are hypersensitive to base damaging DNA agents, which cause accumulative repair intermediates [56]. In addition, the knock-out of FEN1 also causes early embryonic lethality of mice [57], suggesting an essential role of FEN1 for cell survival.

On the one hand, population-based studies showed that germline polymorphisms of some BER enzymes are associated with elevated risks of cancers [58-60]. For instance, the APE1 Asp148Glu genetic carriers has been reported to have an increased risk of colorectal cancer [58]. Meanwhile, the OGG1 Ser326Cys genetic carriers show an increased risk of developing breast cancer [59]. In addition, the heterozygous pol β variant with deletion of exon 4-6 and 11-13 is associated with an increased risk of ovarian carcinoma [60]. In addition, some functional studies of polymorphic variants of BER enzymes also revealed a linkage between variants of BER enzymes and cancers [61, 62]. Studies of pol β Arg137Gln indicate the impaired DNA synthesis activity, which results in accumulative DNA repair intermediates, contributes to the higher risk in cancer development compared to the wild-type pol β carriers [61, 62]. In fact, many non-

germline mutants of BER enzymes have also been identified in a variety of cancers: more than 140 pol β mutants with multiple or single amino acids alternation have been identified in different cancers including gastric, colorectal, prostate, lung, breast, bladder, and esophageal cancers [63]. Some mutations also cause an altered enzymatic activity of pol β . The pol β Leu22Pro mutant with the mutation at the dRP lyase domain, which was identified in a gastric carcinoma, exhibits impaired dRP lyase activity [64]. Another pol β mutant, pol β Glu295Lys, has a reduced DNA synthesis fidelity, which can cause high mutation frequency during DNA synthesis, resulting in cellular transformation [65, 66]. On the other hand, the colorectal cancer-related FEN1 mutant (FEN1 Leu209Pro) was identified to have an impaired endonucleolytic activity, and the reduced endonucleolytic activity induces cellular transformation [67]. Therefore, the integrity of BER enzymes is essential for genomic stability.

The interaction between BER enzymes and cofactors guarantees the efficiency of BER enzymes. The BER cofactors modulate enzymatic activity and regulate the coordination between the BER enzymes. For instance, one BER cofactor (XRCC1) that functions as a scaffold protein, which does not possess any catalytic activities, can enhance the enzymatic activity of APE1 and several DNA glycosylases including OGG1, UNG2, NEIL1, and NEIL2 [68-72]. The functional stimulation of APE1 and DNA glycosylases by XRCC1 indicate an important role of XRCC1 in regulating the enzymatic coordination during the initiation of BER. In addition, the XRCC1 interacts with pol β via its N-terminal domain, while the C-terminal BRCT domain of XRCC1 forms a strong complex with LIG III α [68]. Disruption of the interaction between pol β and XRCC1 leads to a reduced efficiency of ligation by LIG III α [73], which

compromises the efficiency of BER. Meanwhile, an alternation in the BRCT domain of XRCC1 that interact with LIG III α directly affects expression of LIG III α [74], causing reduced ligation, which further allows pol β strand displacement synthesis, thus resulting in the least-efficient strand-displacement LP-BER [75]. The impaired BER in the presence of XRCC1 mutants indicates a critical role of XRCC1 in ensuring the efficiency of SP-BER by regulating the coordination between pol β and LIG III α .

Another important BER cofactor, the proliferating cell nuclear antigen (PCNA), also plays an important role during BER [76]. Proliferating cell nuclear antigen, a central scaffold protein, acts as the central regulator that coordinates not only BER but also other metabolic pathways including DNA replication through protein-protein interactions [77]. At least 200 proteins contain the PCNA-interacting protein (PIP) box, a conserved 8-amino-acid motif that allows the proteins to physically interact with PCNA [78]. During BER, several DNA glycosylases, such as UNG2 and MPG, interact with PCNA [69, 79]. In addition, PCNA can stimulate replicative DNA polymerase pol δ , FEN1, and LIG I via the interaction with their PIP box during LP-BER [80-82]. Although PCNA regulates its binding partners through interaction with PIP boxes, the binding affinities of different PIP boxes and the availability of these PIP boxes to the hydrophobic surface of PCNA also affect PCNA interactions [78]. The differential binding affinity to PIP boxes of PCNA allows PCNA to regulate BER enzymes (pol δ , FEN1, and LIG I) through a sequential switching (from pol δ to FEN1 and then to LIG I), thereby ensuring an efficient LP-BER [83]. Disruption of the interaction between FEN1 and PCNA or between LIG I and PCNA results in a defective LP-BER [84, 85]. Therefore, BER cofactors, such as XRCC1 and PCNA, not only affect the enzymatic activity of the BER core enzymes, but also

regulate the coordination between the enzymes, which ensure a precise and efficient BER.

Efficient BER is critical in safeguarding the integrity and stability of the genome. However, inefficient BER leads to the accumulation of DNA repair intermediates and causes mutation and genomic instability, thereby inducing the development of cancers and lupus [86, 87]. Also, inefficient BER can induce trinucleotide repeat (TNR) instability [14], a typical form of genomic instability that is mainly associated with neurodegenerative diseases and several cancers [88-90].

C. Trinucleotide Repeat Instability and Human Disease

Trinucleotide repeats (TNRs) are a subset of microsatellite DNA that consist of 3 nucleotides in each repeat unit [91]. Trinucleotide repeats are highly polymorphic and are prone to a gain or loss of repeats resulting in repeat expansions and deletions [91, 92]. Both repeat expansions and deletions are associated with human diseases [88-90, 93]. The repeat deletions are associated with human cancers; e.g., the CAG repeat deletion in androgen receptor (AR) gene is associated with prostate cancer [93]. Studies of the association between TNR expansions and human diseases is more established. Currently, more than 40 neurodegenerative diseases, such as Huntington's disease (HD), Friedreich's ataxia (FRDA), fragile X syndrome (FXS), and myotonic dystrophy (DM), are identified as the results of progressive TNR expansions, which can cause cellular toxicity through protein aggregations, induce functional impairment of respective proteins, suppress essential mRNA production, and generate RNA toxicity via production of immature RNA transcripts [88-90]. All such abnormal cellular events can result in

death of neuronal cells, leading to the onset of TNR-expansion-mediated neurodegenerative diseases.

The likelihood of the repeat expansions tremendously increases if the repeat length exceeds a crucial threshold length [92]. Normal individuals carry the same repeats below the threshold length in the respective regions; however, some individuals carry more repeats than the threshold length, the pre-mutation length [92]. The repeats below the threshold are relatively stable. Whereas, the repeats within the pre-mutation range are unstable and tend to result in repeat expansions: typically the longer the repeat tract is, and the earlier the onset of the diseases occurs [92, 94]. The TNR expansions can occur at both the encoding region and the noncoding region (Figure I.5): generally, small repeat expansions as short as 10 repeats occur at the encoding region, and more substantial repeat expansions of 100-10,000 occur at the noncoding region [92]. The difference in the scales of repeat expansions between the encoding and noncoding region indicates that the small repeat expansions occurred in the encoding region can directly cause protein aggregation and cell death, while the large repeat expansions occurred in the noncoding region can result in the deregulation of gene expression and indirect cytotoxicity.

Figure I.5 Trinucleotide repeat expansions in encoding and noncoding regions [7]

The functional loss of certain proteins can be very destructive. For examples, the expansions of the GNC repeats can result in partial functional loss of zinc-finger protein of the cerebellum 2 (Zic2), which causes abnormality during early development and differentiation [95, 96], leading to holoprosencephaly [97, 98]. On the other hand, the protein aggregates resulted from misfolded protein can be equally deleterious [97, 99]. The repeat expansions occurring at the encoding regions are either (CAG)_n or (GNC)_n, which encodes polyglutamine (polyQ) or polyalanine (polyA), respectively, and expansions of such repeats can induce deleterious cellular events [97, 99]. Both the polyQ and polyA have an increasing tendency to aggregate when the number of the CAG or GNC repeats increases [97, 99-101]. The protein aggregates (Figure I.6), which are refractory to protein degradations, are toxic to the cells and can induce cell death [102, 103], thereby leading to development of disease. The exact mechanisms of how the protein aggregates formed by polyQ or polyA tracts and how they induce neuronal degradations remain to be elucidated. However, it has been proposed that proteins with the expanded polyQs tend to form amyloid-like fibrils abundant in β sheets, leading to dysregulation of transcription, impairment of the ubiquitin-proteasome system, mitochondrial dysfunction, and autophagy defects and thus induce cellular toxicity [100]. There are nine polyglutamine disorders, including spinal and bulbar muscular atrophy, Huntington's disease, and six types of spinocerebellar ataxias [99]. On the other hand, polyA has been suggested to form similar protein aggregates with polyQ [97]. However, a more plausible theory has been proposed that polyA tend to form α -helical clusters [101], given the fact that the difference in the hydrophobicity between alanine and glutamine

may induce aggregations with a distinct mechanism [104]. There are at least nine polyalanine disorders caused by the expansion of GNC repeats, including synpolydactyly type II (SPD) and oculopharyngeal muscular dystrophy (OPMD) [105].

Figure I.6 Protein aggregation and cellular toxicity [11]

In contrast, repeat expansions occurring in the noncoding regions (5'-UTR, introns or 3'-UTR) of certain genes are more dramatic and can indirectly cause cellular toxicity by suppressing gene transcription of essential proteins or by generating immature RNA transcripts, resulting in progressive functional loss of essential proteins or accumulation of RNA toxicity [106-112]. For example, CGG repeat expansion in the 5'-UTR usually cause hypermethylation of CpG of the promoter region of the fragile X mental retardation 1 (*FMR1*), leading to a reduction in mRNA of FMR1 that is responsible for normal cognitive development and female reproductive function [111]. The loss or reduction of the protein leads to progressive cognitive declination and

learning disabilities [111]. Similarly, expansion of GAA repeats in the first intron of the frataxin gene (*FXN*) can reach to 1700 repeats [110]. The mega repeat expansions lead to suppression of *FXN* transcription and reduction of frataxin (Figure I.7). The functional loss of frataxin consequently results in the mitochondrial iron overload and extensive oxidative stress [109], causing neuronal death and the development of Friedreich ataxia [108]. Meanwhile, the CTG repeat expansions in 3'-UTR of myotonic dystrophy type 1 (DM1) can reach to 1000 repeats, causing defective mRNA splicing and the accumulative RNA toxicity [106, 107]. The accumulative RNA toxicity leads to neuronal death and neurodegenerative symptoms such as progressive muscle wasting and feebleness [107].

Figure I.7 GAA repeats expansions suppress frataxin expression [6]

The upper panel illustrates the normal unaffected allele. The lower panel illustrates the affected FRDA allele.

The expansions of TNRs in both encoding and noncoding regions interfere with vital cellular metabolism in a variety of ways. The outcomes of TNR expansions usually result in neurodegenerative disorders. Without available cures, such neurological

disorders cause huge financial and health burdens to the patients' family and the society. Therefore, it is essential to understand the mechanisms of the TNR-related diseases. Such knowledge can assist in identification of potential targets for treatment and prevention of TNR-related diseases.

D. Trinucleotide Repeat Instability and DNA Replication

The TNR instability occurring in the cells is a consequence of imbalanced DNA transactions from the DNA metabolism that involved in addition of new nucleotides into the DNA or removal of existing nucleotides from the DNA. Imbalanced addition and removal of TNRs lead to gain or loss of repeats during DNA metabolism, thereby causing TNR instability. One reason that the TNRs are prone to repeat expansions or deletions is because TNRs have high tendency to form non-B DNA structures/secondary structures [7, 113]. Formation of secondary structures modulates DNA metabolic enzymatic activities, facilitating additions or removal of TNRs and promoting repeat instability. In normal condition, DNA adopts right-handed B-form double helix structure via Watson-Crick base-pairing (A-T and G-C base-pairing) [114]. Whereas, during DNA replication, DNA repairs, DNA transcription, and DNA recombination in TNR tracts, the two strands (sense strand and anti-sense strand) that are separated can form secondary structures such as loops and hairpins, G4 DNA structures, H-DNA, and sticky DNA (Figure I.8) by self-basepairing [7, 92]. The non-B DNA structures can disrupt the coordination among DNA metabolic enzymes, causing repeat instability. For example, when the newly synthesized strand loops out forming a hairpin structure, the polymerases continuously synthesize from the 3'-end of the upstream primer, thereby introducing more repeats during the

synthesis and causing TNR expansions (Figure I.9). When the secondary structures formed on the template strand, the polymerase directly bypasses the hairpin, which results in fewer repeats synthesized on the template, leading to repeat deletions (Figure I.9). Moreover, formation of the secondary structures is sequence dependent. Both $(CAG)_n/(CTG)_n$ and $(CGG)_n/(GCC)_n$ can form hairpin structures [7]. In addition, $(CGG)_n/(GCC)_n$ can also form G4 DNA structures with four G's base-pairing with each other via Hoogsteen base-pairing [7, 115]. The $(GAA)_n/(CTT)_n$, on the other hand, can loop out forming a loop, or H-DNA, or sticky DNA [7]. The stability of the secondary structures also correlates with the length of the repeats with a high stability of the secondary structures formed in the long repeat tracts [116, 117].

Figure I.8 Formation of secondary structures on TNRs [7]

a. Hairpin structure, b. G-quadruplex, c. Hairpin structures on both CAG and CTG repeats, d. H-DNA, and e. Sticky DNA.

Figure I.9 Secondary structures disrupt DNA synthesis of DNA polymerases [7]

Although several DNA metabolic pathways including DNA replication, NER, BER, and DNA recombination may induce the formation of secondary structures, many *in vivo* and *in vitro* studies support the DNA replication [118-121] and the BER [1, 2, 14, 122, 123] are two primary mechanisms that cause repeat expansions since TNR expansions occur in diving cells where both the DNA replication and the BER may be involved and TNR expansions occur in non-diving cells where only the BER is involved [92]. Because secondary structure can easily form during DNA replication and the replicative DNA polymerases, pol ϵ and pol δ , exhibit rapid DNA synthesis, the DNA replication has been proposed to be responsible for substantial repeat expansions in diving cells. Several studies of TNR instability from bacteria, yeast, and mammalian cells have shown that TNR expansions can occur during DNA replication, and repeat expansions occur only at the leading strand [118, 120, 121], exhibiting repeat-sequence orientation dependency. Moreover, it has been shown that the replication fork stalling during DNA replication is responsible for TNR expansions [124]. *In vivo* studies show that the replication fork stalls in expanded TNR tracts during the lagging strand synthesis [125, 126], which has been directly visualized by using electron microscopy during the DNA replication of CTG/CAG repeats [127], suggesting that secondary structures formed

in the repeat tracts trigger the replication fork regression during the lagging strand DNA synthesis. Such studies led to a proposed the replication restart model [7, 120]. In the replication restart model (Figure I.10), secondary structures formed on the lagging strand can induce both repeat expansions and repeat deletions. When the pol δ skips the secondary structure, it results in repeat deletions. On the other hand, the secondary structures formed on the lagging strand cause the replication fork stalling, and the subsequent replication restart leads to resolution of the secondary structure on the lagging strand, forcing formation of a secondary structure on the leading strand through the re-alignment of the nascent synthesized leading strand to its template strand. Consequently, pol ϵ synthesized from the 3'-end of the newly formed secondary structure on the leading strand, which results in repeat expansions (Figure I.10).

Figure I.10 Replication restart leads to TNR instability [7]

The upper panel illustrates that the secondary structure induce repeat contraction during lagging strand synthesis. The lower panel illustrates that the secondary structure formed on the lagging strand triggers the replication collapse and restart, resulting in repeat expansion.

The DNA replication fork stall model has successfully explained TNR expansions in dividing cells. However, it cannot explain why substantial repeat expansions occur in the post-mitotic cells [92] where DNA replication is absent. Given the fact that the neuron cells are post-mitotic cells where the DNA replication is absence, and the fact that many disease patients have their first symptoms of the diseases appeared at the age of 50s or 60s [128-130], other mechanisms, such as oxidative DNA damage and its repair by the BER pathway, have been proposed to be responsible for TNR expansions in somatic cells.

E. Trinucleotide Repeat Instability and Base Excision Repair

Base excision repair is one of the most abundant and efficient DNA repair pathways that exist in both dividing and non-dividing cells [45]. It repairs many oxidative DNA damage including the most commonly generated base lesions, 8-oxoG, abasic lesions, and ssDNA breaks, which can also occur in TNR tracts [27, 28]. In fact, oxidative DNA lesions frequently accumulates in TNR tracts since the abundance of guanines with a low-reduction-potential, which increases the frequency of oxidative DNA damage on TNRs [131]. To repair DNA base lesions on the TNR tracts, DNA glycosylases, such as OGG1, initiate BER by removing the damaged bases, and APE1 incises the DNA backbone at the 5'-side of an abasic site, generating a single strand DNA break that in turn result in DNA strand separation [14]. The single strand break then promotes the formation of secondary structures. The secondary structures may further affect the activities of DNA metabolic proteins, leading to TNR expansions or deletions. Thus, oxidative stress has been reported to be closely associated with TNR expansions in

both stem cells and somatic cells [14, 123]. However, oxidative damage alone are not sufficient for TNR instability [14, 122], and the co-existence of OGG1 and other BER enzymes along with the oxidative lesions has been reported to be crucial for TNR instability [14, 122], suggesting BER is a mediator for TNR expansions.

Figure I.11 BER-mediated TNR expansions [1]

Subpathway 1 and 2 illustrate that no repeat instability induced by FEN1 efficient 5'-flap cleavage during "Hit and Run" BER and long-patch BER. Subpathway 3 illustrates that repeat expansion induced by hairpin structure formed on the downstream damage strand, which inhibits FEN1 cleavage.

In the BER-mediated TNR instability, only the strand-displacement LP-BER has been reported to promote repeat instability [1]. In SP-BER and the hit-and-run pathways, the number of TNRs synthesized by pol β equals the one removed, and therefore no repeat instability occurs in such situations [8]. However, in the LP-BER sub-pathway, the repeat instability occurs because the secondary structures disrupt the coordination among BER enzymes, particularly the coordination between pol β and FEN1 [1]. In a model proposed for TNR expansion (Figure I.11) [1]. A hairpin structure is formed on the downstream substrate, which leaves a multi-nucleotide gap for pol β to fill in [1]. The pol β generates multi-nucleotide gap filling synthesis, which results in addition of extra repeats [1]. Meanwhile, the hairpin cannot be removed by FEN1 since FEN1 requires specific 5'-flaps as its substrates [1]. The inability to remove the hairpin by FEN1 consequently leads to alternate cleavage by FEN1 in removing a short 5'-flap, leading to ligation of the hairpin [1]. During the repair, pol β synthesizes more repeats than FEN1 removes, leading to repeat expansion [1]. However, the analysis of TNR expansion profiles in non-dividing cells shows that a single TNR expansion mediated by BER of a single DNA base lesion cannot explain substantial repeat expansion of 3000 GAA repeats in the *FXN* gene and 100-1000 CTG repeats in the *DMI* gene [92]. Given the fact that the onsets of TNR-mediated neurodegenerative disorders require progressive development of the diseases for 10 years or longer in some cases [132, 133], it is possible that multiple rounds of BER-mediated TNR expansions must occur to substantially expand large TNR tracts. The hypothesis of multiple BER leads to large repeat expansion is supported by the fact that large TNR expansions can be induced through BER in a Huntington's transgenic mouse model through aging, which presumably result in accumulation of

oxidative DNA damage in expanded CAG repeats [14]. This has been further supported by the fact that the CAG repeat expansions in mouse neuron cells occur in an age-dependent manner [14]. It is conceivable that during each round of BER, CAG repeats are expanded by a small number, which make the repeats more unstable [14]. With the increasing repeat expansion of the repeat tract and accumulative oxidative damage, multiple rounds of BER in CAG repeats lead to a “toxic oxidation cycle” in a lesion-BER-expansion manner (Figure I.12), which ultimately promote the repeat length to be expanded over the threshold, leading to onset of the diseases.

Figure I.12 BER-mediated toxic oxidation cycle [14]

Moreover, the mismatch repair protein complex, the MSH2-MSH3 complex, also contributes to repeat expansion in a BER-dependent manner (Figure I.13). Under a normal condition, MSH2-MSH3 serves as a structure recognition protein that specifically binds to small insertion/deletion loops; MSH2-MSH3 undergoes a conformational change and carries out its normal function, which coordinates the removal of the structures by adenosine triphosphate (ATP) hydrolysis [134]. However, binding to the relatively large loop and hairpin structures formed on expanded CAG repeats tracts can abolish the ability of MSH2-MSH3 for conformational change and ATP hydrolysis. The binding on the hairpin then traps MSH2-MSH3 on these secondary structures [135]. During BER on TNRs, the entrapment of MSH2-MSH3 on the secondary structures stabilizes the secondary structures, which subsequently prevents FEN1 cleavage [2]. Meanwhile, MSH2-MSH3 physically interacts with pol β at the lesion sites [2]. The interaction stimulates pol β synthesis of TNRs, which further prevents pol β hairpin-bypass and removal of TNRs by FEN1, leading to repeat expansions [2].

Figure 1.13 MSH2-MSH3 coordinates BER enzymes to promote TNR expansions [2]

Subpathway on the left illustrates that repeat deletion occurs during BER in the absence of MSH2-MSH3. Subpathway on the right illustrates that MSH2-MSH3 facilitates repeat expansion by inhibiting 5'-flap cleavage activity of FEN1 and stimulating DNA synthesis activity of pol β.

On the other hand, BER can also induce TNR deletions. It has been reported that some chemotherapeutic reagents such as temozolomide, ionizing radiation, UV radiation, hydrogen peroxide, and environmental pollutants such as chromate, bromate can induce TNR deletions in bacteria, mouse stem and kidney cells, and human kidney cells and lymphoblasts [5, 136-141]. The discrepancy of the outcomes of TNR instability (repeat expansions and repeat deletions) from different cell types and different endogenous and exogenous DNA damaging agents indicate that TNR instability induced by DNA base lesion through BER is a complex process that involves many factors. Both *in vivo* and *in vitro* studies have suggested that several factors contribute to the discrepancy TNR instability mediated by BER. These factors include the locations/sites

and the sizes of the secondary structures, the types of lesions on the TNRs, the coordination among BER enzymes, and the involvement of other nucleases and cofactors [1, 5, 9, 14, 122, 142-149].

In vitro studies have demonstrated that the locations of the secondary structure are important in governing TNRs instability [1, 142]. When a secondary structure forms on the downstream damage strand, it inhibits FEN1 cleavage [1, 5]. The inhibition forces FEN1 to perform an alternate flap cleavage activity to remove a shorter flap instead of removing the hairpin, which results in more repeats synthesized by pol β than removed by FEN1, leading to ligation of a hairpin and TNR expansions (Figure I.11 3rd pathway). However, when the secondary formed on the template strand, pol β can skip over the secondary structures on the template strand [5, 142]. The pol β only synthesizes a small portion of repeats within a TNR [142]. The inability of pol β to synthesize within the hairpin results in direct bypass of the big hairpin, leading to FEN1 removal of more repeats and repeat deletions as a consequence [142]. The bulky oxidized DNA base lesions such as the 5',8-cyclo-2'-deoxypurines (cdPus) [150] can distort the DNA backbone because it contains the covalently linkage between the C5 of the 2'-deoxyribose and the C8 of the purine [150]. The distortion on the DNA backbone facilitates formation of a small loop structure with the cdPu lesions [143]. During BER, pol β bypasses the loop, allowing FEN1 to remove more repeats than those synthesized by pol β , leading to TNR deletions [143].

In addition, sites of the base lesions may contribute to the formation of secondary structures with varying sizes located at different DNA strands, which disrupts BER enzymatic coordination [5]. The sites of base lesions are associated with the formation of

different secondary structures on both damage and template strands, contributing to varying outcomes of repeat instability (Figure I.14). When an AP lesion located at the 5'-side of a TNR repeat tract from the random sequence flanking region, it promotes a larger hairpin to form on the downstream damage strand [5]. The secondary structures facilitate pol β multi-nucleotide gap-filling synthesis, FEN1 alternate flap cleavage, and ligation of a hairpin, causing repeat expansion [5]. However, if a base lesion occurs in the middle, multiple small hairpins can occur in the upstream and downstream strand of the damaged strand and template strand [5]. In this scenario, pol β performs weak synthesis of TNRs and bypasses a small hairpin on the template strand while FEN1 removes the unstable downstream hairpins, leading to removal of more repeats than those synthesized and resulting in repeat deletion [5]. In a scenario where multiple abase lesion occurs in the middle of the repeat tract, a large hairpin along with a small hairpin forms on the template strand, which results in pol β skip over the hairpins, leading to large repeat deletion [5]. In contrast, when the lesion occurs at the 3'-side of a TNR tract from the random sequence flanking region, no repeat instability occurs since the sizes of secondary structures on both the damage strand and the template strand are equal [5].

Figure I.14 Locations of base lesions govern BER-mediated TNR instability [5]

Subpathway 1 illustrates that base lesion located on the 5'-end of repeat tract induces small repeat expansion. Subpathway 2 illustrates that base lesion located in the middle of repeat tract induces small repeat deletion. Subpathway 3 illustrates that base lesion located at the 3'-end of repeat tract does not induce repeat instability. Subpathway 4 illustrates that several base lesions located in the middle of repeat tract induce large repeat deletion.

Trinucleotide repeat expansions usually occur in an age- and tissue-dependent manner [14, 122]. It has been shown that CAG repeat expansion preferentially occurs in striatum with a high level of pol β, a low level of FEN1, and a low level of high mobility box 1 (HMGB1), a cofactor in LP-BER that can stimulate FEN1 flap cleavage [122, 144]. In contrast, neither TNR expansions nor high level of pol β expression have been observed in cerebellum [122, 144]. The study suggests that a high level of pol β and a low level of FEN1 is critical in repeat expansions. Therefore, a reduction in DNA synthesis activity of pol β or an enhancement of FEN1 cleavage may contribute to TNR

deletions. In fact, pol β Y265C with impaired DNA synthesis activity has been reported to reduce repeat expansion frequency in both sperm and brain cells of the heterozygous pol β Y265C fragile X-related disorder mice [146, 151], demonstrating that DNA synthesis activity of pol β is critical in modulating TNR instability. On the other hand, other DNA repair nucleases can also contribute to TNR deletions by coordinating with FEN1 during BER. It has been shown that both APE1 and Mus81/Eme1 can be involved in removing a CAG repeat hairpin embedded in the expanded TNRs [9, 145]. The guanines embedding in the single strand regions of secondary structures, such as the loop regions of hairpins, are susceptible to oxidative DNA damage compared to that are in duplex DNA where they base-pairs with cytosines [152, 153]. The DNA base lesions in the loop regions of a CAG repeat hairpin can be removed by DNA glycosylases, OGG1, NEIL1, and NEIL2 [145, 152-154], leaving an abasic site in the loop region of the hairpin. The APE1 cleavage then results in the formation of a double flap intermediate that contains a 3'-flap and a 5'-flap, which can then be removed by the coordinated flap cleavage activity between Mus81/Eme1 [145] or the APE1 3'-5' exonucleolytic activity and FEN1 (Figure I.16).

Figure I.15 APE1 suppresses TNR expansions by facilitating removal of upstream repeats [9]

Subpathway 1 illustrates that APE1 3'-5' exonuclease activity facilitates removal of upstream repeats and prevents repeat expansion. Subpathway 2 illustrates that upstream hairpin structure inhibits APE1 3'-5' exonuclease activity and cause repeat expansion. Subpathway 3 illustrates that coordination between 5'-flap cleavage activity of FEN1 and 3'-5' exonuclease activity of APE1 prevents repeat expansion.

The coordination among BER enzymes and cofactors can also contribute to the outcomes of TNR instability. The X-ray repair cross-complementing protein 1 (XRCC1) and proliferating cell nuclear antigen (PCNA) are two crucial cofactors in SP-BER and LP-BER, respectively. Although there is no direct evidence between XRCC1 and TNR instability, it is likely XRCC1 may promote repeat deletions by interacting with APE1 in resolving a hairpin structure since XRCC1 can stimulate APE1 3'-5' exonucleolytic activity [155]. Moreover, XRCC1 is important for maintaining the efficiency of SP-BER, disruption of the interaction between pol β and XRCC1 can lead to inefficient ligation by

LIG III α [75], which subsequently promotes pol β strand-displacement DNA synthesis leading to LP-BER increasing the probability of TNR instability. The PCNA, on the other hand, has been supported to be crucial in TNR instability by several studies [147-149]. A polymorphic PCNA variant, PCNA Ser228Ile variant with impaired interaction with FEN1 and LIG I, has been reported to be responsible for deficiency in DNA repair, which contributes to a high susceptibility to neurodegeneration [144, 156, 157]. Because PCNA stimulates the activities of both FEN1 and LIG I by interacting with both enzymes via PIP box, the coordination between PCNA and these enzymes are crucial in modulating TNR instability [147-149]. Disruption of the interaction between human PCNA and LIG I leads to an increased TNR instability [147]. Disruption of the interaction between PCNA and LIG I or FEN1 also increases TNR instability frequency in budding yeast [148]. Moreover, PCNA assists the resolution of a CAG repeat hairpin by stimulating FEN1 flap cleavage activity and LIG I activity, facilitating repeat deletion and removal a CAG repeat hairpin and preventing TNR expansion [149]. Furthermore, PCNA interacts with MSH3, and the interaction is important for MSH2-MSH3 mediated lesion recognition [158], suggesting a role of PCNA in modulating TNR instability by interacting with MSH2-MSH3.

Proliferating cell nuclear antigen can also modulate TNR instability through its posttranslational modification. For example, mono-ubiquitination of PCNA can regulate the coordination between BER enzymes and enzymes from other DNA repair pathways, e.g., translesion (TL) DNA synthesis [159-161]. Although there is no direct evidence of mono-ubiquitinated PCNA (ub-PCNA) induces TNR instability, the importance of ub-PCNA in regulating different DNA polymerases and nucleases implicates it may play a

critical role in modulating TNR instability through BER. PCNA can be ubiquitinated at Lys164 by RAD18-RAD6 ubiquitin ligase complex mediated mono-ubiquitination [162, 163]. The RAD18-RAD6 ligase can be recruited to the lesion site by replication protein A (RPA) [164, 165], a single strand DNA binding protein that is involved in the DNA replication and DNA repair, including BER [166]. Mono-ubiquitination of PCNA can also be induced by the DNA base damage agents, methyl methanesulfonate (MMS) and UV irradiation [167]. Since MMS-induced alkylating DNA damage is subject to BER [168], this indicates that PCNA can be mono-ubiquitinated during BER. Mono-ubiquitination of PCNA provides additional protein binding surfaces (Figure I. 19) to interact with other DNA repair proteins with a high binding affinity via either an ubiquitin-binding module (UBM) or an ubiquitin-binding zinc domain (UBZ). All the TL polymerases, including REV1, pol η , pol κ , and pol ι , have UBM or UBZ [161]. It has been shown that interaction between ub-PCNA with pol η , pol κ , and pol ι enhances their DNA synthesis activity. It has also been suggested that TL polymerases are involved in DNA replication-mediated repeat instability where pol δ was mutated in budding yeast [169]. Unlike the replication polymerases (pol ϵ and pol δ) and other DNA repair polymerases, such as pol β , TL polymerases have an enlarged active site. The enlarged active site may allow the polymerases to interact with the secondary structures formed on TNR tracts with more flexibility that further increases the likelihood of repeat instability. Additionally, since ub-PCNA does not physically interact with pol β , the binding between the TL polymerases and ub-PCNA may enhance the DNA synthesis activity of the TL polymerases, leading to TNR instability with a mechanism that is different from the one employed by pol β .

Figure I.16 Mono-ubiquitinated PCNA (the ubiquitin depicted as cyan and the PIP binding box depicted as magenta) [4]

On the other hands, ub-PCNA can also regulate the activity of nucleases during the repair. The ub-PCNA can compromise the interaction between PCNA and FEN1. Structural alignment between ub-PCNA and PCNA-FEN1 complex reveals an overlap between ubiquitin and the active site of FEN1 (Figure I.20), suggesting ub-PCNA may not be able to interact with FEN1 and stimulate its activity. The ub-PCNA can also physically interact with the Fanconi anemia-associated nuclease 1 (FAN1) via UBZ and the non-canonical PIP box located at N-terminal domain of FAN1 [159, 160]. The FAN1 was initially identified as nuclease with the ability to unlock the interstrand crosslink (ICL) [170, 171] in Fanconi anemia (FA) DNA repair pathway, and its deficiency results in a rare genetic kidney disorder [172]. It has been shown that single nucleotide

polymorphisms of FAN1 are associated with the modulation of onset age of Huntington's disease [173, 174], suggesting a critical role of FAN1 in modulating CAG repeat instability. The FAN1 may facilitate in suppressing CAG repeat expansion through its 5'-3' exonucleolytic and 5'-flapendonucleolytic activity. The 5'-flap exonucleolytic activity of FAN1 requires dimerization of two FAN1 monomers in a head-to-tail fashion [171, 175]. However, the 5'-3' exonucleolytic activity only requires its monomer [171, 175] and may be particularly important in resolving the hairpin structures formed on the downstream strand of a TNR tract. Thus, although ub-PCNA may suppress the interaction with FEN1, it may stimulate FAN1 activity to resolve a downstream TNR hairpin modulating TNR instability during BER.

Figure I.17 Ubiquitin may prevent interaction between FEN1 and PCNA via PIP box [4]

In summary, BER plays an active role in modulating somatic TNR instability induced by DNA base lesions and ssDNA breaks. The BER-mediated TNR instability is complex and involves many different factors. Some factors have been identified to promote TNR expansions, the others have been shown to induce TNR deletions. Exploration of the molecular mechanisms underlying BER-mediated TNR instability is particularly important as it can help to identify potential therapeutic targets for attenuation of the expanded TNRs, thereby facilitating the development of the treatment and prevention for TNR expansion diseases. In my Ph.D. dissertation research, I explored the underlying molecular mechanisms of BER-mediated somatic TNR instability from three different areas. The results have been summarized in the three chapters in the dissertation. Chapter 1 evaluates the effects of the polymorphic variants of DNA polymerase β (pol β) on TNR instability during BER. Chapter 2 explores the mechanisms how the DNA synthesis activity of pol β modulates the outcomes of TNR instability, and how the inhibition of pol β activity can preferentially contribute to CAG repeat deletion, and the implications of inhibition of pol β DNA synthesis in the treatment of neurological disorders mediated by TNR expansions. Chapter 3 of my dissertation investigates how the coordination between ub-PCNA and nucleases, FAN1, and FEN1, can modulate TNR instability through BER, and the implications of the coordination as a new target for the treatment TNR expansion diseases.

OVERVIEW

Trinucleotide repeat (TNR) expansions are associated with more than 40 human neurological disorders including Huntington's disease, Friedreich's ataxia, and fragile X syndrome. The repeat expansions can occur through maternal transmission and continue to progress after post-mitotic differentiation. The DNA replication, DNA repairs, DNA recombination, and gene transcription have been proposed to be associated with inherited and somatic TNR expansions and deletions. Recent studies show that BER is one of the major pathways that is responsible for TNR expansion in post-mitotic neuronal cells. The BER-mediated TNR instability is the result of inability of the BER enzymes in resolving the secondary structures, and consequently resulting in an imbalance between addition of TNRs by DNA polymerases or removal by nucleases during BER. Other factors including the sites of DNA base lesions, locations and sizes of the secondary structures, the lesion types, the coordination among BER enzymes and cofactors, and involvement of nucleases and cofactors from other DNA repair pathways all contribute to modulating of TNR instability mediated by BER. Yet the mechanisms as to how the enzymes and cofactor and their posttranslational modifications can modulate TNR instability through regulating the balance between the addition and removal of TNRs during BER remain to be elucidated. Since there are no effective treatment for TNR-mediated neurological disorders, exploration of the molecular mechanisms is essentially important as it provides new insights into development and prevention of TNR expansion diseases, which will further aid in identification of new therapeutic targets and biomarkers for early diagnosis of TNR-related diseases.

In Chapter 1, we evaluate the effects of a germline polymorphic pol β R137Q variant on CAG repeat instability during BER. The results indicate that pol β R137Q does not significantly affect CAG repeat instability during BER suggesting that the polymorphic variant R137Q carriers may not have a higher risk in developing TNR expansion mediated neurological disorders than the normal population (published). In Chapter 2, we explore the mechanisms as to how the DNA synthesis activity of pol β can modulate the outcomes of TNR instability, how the inhibition of pol β activity can preferentially lead to CAG repeat deletion, and the implications of inhibition of pol β DNA synthesis in the treatment of TNR-expansion neurological disorders. We find that weak DNA synthesis activity promotes BER-mediated TNR deletions. Inhibition of pol β DNA synthesis activity may be utilized as a potential therapeutic strategy to shorten the expanded TNR. In Chapter 3, we explore the effect of the ub-PCNA on TNR instability by modulating activities of the nucleases, FEN1 and FAN1. We find that ub-PCNA stimulates FAN1 5'-3' exonucleolytic activity in resolving the downstream hairpins, thereby suppressing TNR expansions.

CHAPTER 1: MODULATION OF TRINUCLEOTIDE REPEAT INSTABILITY BY DNA POLYMERASE β POLYMORPHIC VARIANT R137Q

ABSTRACT

Trinucleotide repeat (TNR) instability is associated with human neurodegenerative diseases and cancer. Recent studies have pointed out that DNA base excision repair (BER) mediated by DNA polymerase β (pol β) plays a crucial role in governing somatic TNR instability in a damage-location dependent manner. It has been shown that the activities and function of BER enzymes and cofactors can be modulated by their polymorphic variations. This could alter the function of BER in regulating TNR instability. However, the roles of BER polymorphism in modulating TNR instability remain to be elucidated. A previous study has shown that a pol β polymorphic variant, pol β R137Q is associated with cancer due to its impaired polymerase activity and its deficiency in interacting with a BER cofactor, proliferating cell nuclear antigen (PCNA). In this study, we have studied the effect of the pol β R137Q variant on TNR instability. We showed that pol β R137Q exhibited weak DNA synthesis activity to cause TNR deletion during BER. We demonstrated that similar to wild-type pol β , the weak DNA synthesis activity of pol β R137Q allowed it to skip over a small loop formed on the template strand, thereby facilitating TNR deletion during BER. Our results further suggests that carriers with pol β R137Q polymorphic variant may not exhibit an elevated risk of developing human diseases that are associated with TNR instability.

INTRODUCTION

Human genome is susceptible to a variety of types of DNA damage that can modify DNA bases, deoxyribose sugar phosphate (dRP) groups as well as directly break DNA backbone [176]. It has been estimated that more than 10,000 base lesions are generated per cell per day [177], and these lesions are efficiently repaired by DNA base excision repair (BER) [49, 177] through the single-nucleotide or long-patch BER sub-pathway [8, 27, 73, 178, 179].

Genome instability, typically microsatellite instability is responsible for many human diseases [180-183] including GT repeat instability that is associated with colon cancer [184] as well as trinucleotide repeat (TNR) expansion diseases [89, 92, 185, 186]. TNR expansion has been identified as the cause of more than 40 neurodegenerative diseases [7, 92] including Huntington's disease (HD), spinocerebellar ataxia (SCA) type 1, 2, 3, 6, 17 and spinal bulbar muscular atrophy (SBMA) (Kennedy's disease) (CAG repeat expansion) [89, 187, 188], myotonic dystrophy type 1 (DM1) (CTG repeat expansion), Friedreich's ataxia (GAA repeat expansion) and fragile X syndrome (CGG repeat expansion) [189-191]. TNR expansions can occur in both the coding or non-coding regions of the genes associated with disease development, leading to aberrant protein aggregation or deficiency of gene expression [7, 192, 193]. On the other hand, CAG repeat deletion is associated with cancer [194]. CAG repeat deletion in the androgen receptor (*AR*) gene can result in a high transcriptional activity of the AR protein [194-196], which may potentially lead to progression of prostate cancer [197, 198].

TNR instability is mediated by the formation of secondary structures, including hairpins, loops and G4-quadruplex [7, 92], during DNA replication [199], repair [14,

200] and recombination [201] as well as gene transcription [202]. Furthermore, it has been shown that DNA repair [92] and gene transcription [202] play crucial roles in modulating somatic TNR instability [203] especially in postmitotic neurons. Since TNR tracts contain a long stretch of guanines that are susceptible to DNA base damage [177, 204], they form hotspots of base lesions and are constantly subject to multiple cycles of BER, which leads to a “toxic oxidation cycle” resulting in TNR expansion [14, 92]. This is supported by the fact that TNR expansion is promoted by increased amount of 8-oxoguanine (8-oxoG) in the neurons of HD transgenic mice [14] and germ cells of fragile X syndrome mice treated with an oxidative DNA damaging agent, potassium bromate [205]. Moreover, TNR deletion can also be induced by an alkylating DNA damaging agent, temozolomide [206] in the lymphoblasts of Friedreich’s ataxia patients through BER [207]. As a core enzyme of BER, pol β plays a critical role in maintaining genome stability [49, 208] as well as modulating TNR instability [1, 14, 209-211]. It has been found that pol β promotes TNR expansion by performing multi-nucleotide gap-filling synthesis on a TNR repeat tract [14] and facilitating FEN1 alternate flap cleavage of a short repeat flap. This subsequently leads to ligation of a hairpin during long-patch BER [1]. It has been suggested that during DNA replication, pol β can also promote repeat expansion by extending the 3'-terminus of a hairpin to produce extra repeats [209]. A recent study has also shown that pol β can interact with mismatch repair proteins MSH2-MSH3 to promote TNR expansion [2]. On the other hand, pol β facilitates TNR deletion by skipping over a TNR hairpin on the template strand [210, 211] or bypass a 5', 8-cyclo-2'-deoxyadenosine (cdA), a bulky base lesion located in a loop on the template of a CTG repeat tract [212]. Our previous studies have shown that pol β coordinates with FEN1 to

govern the balance between the addition and removal of nucleotides during the repair on TNRs, thereby leading to TNR expansion or deletion in a damage-location dependent manner [210]. All these indicate that in coordination with FEN1, pol β modulates TNR expansion or deletion during long-patch BER via its multi-nucleotide gap-filling synthesis, strand displacement DNA synthesis, hairpin-bypass synthesis. Since the long-patch BER in the context of a TNR tract is much less efficient than the single-nucleotide BER, this allows the formation of secondary structures, such as hairpins and loops. Thus, efficient BER prevents TNR instability by inhibiting DNA slippage and the formation of hairpin and loop structures in a TNR tract, whereas inefficient BER can promote the processes and TNR instability [1, 8].

Genetic variations, i.e. polymorphism of DNA repair enzymes and cofactors, 8-oxoguanine DNA glycosylase (OGG1) [59, 213], APE1 [214], X-ray repair cross-complementing protein 1 (XRCC1) [215-217], XPC [218, 219], MSH3 [220-223], RPA-CDK7 [224] among others have been reported to be associated with cancer and neurodegenerative diseases. However, some studies do not support the notion [225-227]. The controversy is due to lack of knowledge of the effects of these BER polymorphic variants on genome stability and integrity. Thus far, three germline polymorphic variants of pol β with single amino acid substitution have been identified in human population. They are pol β R137Q, pol β P242R, and pol β Q8R variant that are associated with cancer [61, 228-232]. Among them, pol β R137Q variant is particularly of interest. It contains the substitution of arginine 137 with glutamine, which occurs in the polymerase domain of pol β and is involved in mediating the interaction between pol β and proliferating cell nuclear antigen (PCNA) [233]. The pol β variant exhibits impaired DNA synthesis

activity and deficiency in interacting with PCNA. This results in cellular hypersensitivity to an alkylating DNA damaging agent, methyl methanesulfonate (MMS) [61]. Thus, it is conceivable that the impaired DNA synthesis activity of pol β R137Q may disrupt the coordination between pol β and other BER proteins, promoting genome instability such as TNR instability during BER. To test this possibility, we initially characterized DNA synthesis activity of the polymorphic pol β R137Q variant and its effects on CAG and CTG repeat instability during BER. We found that pol β R137Q variant showed weaker DNA synthesis activity than wild-type pol β during BER in the context of CAG and CTG repeats. Yet it exhibited similar ability as wild-type pol β to cause deletion of CAG and CTG repeats during BER of an abasic lesion at various locations. We provide the first evidence that pol β R137Q variant modulates TNR instability in a similar manner as wild-type pol β . Our results further suggest that the individuals who carry pol β R137Q polymorphic variant, do not exhibit a higher risk of development of TNR instability and its associated diseases than individuals who have wild-type pol β .

MATERIALS AND METHODS

Materials

DNA oligonucleotide substrates were synthesized by Integrated DNA Technologies (Coralville, IA, USA). The radionucleotides γ - ^{32}P ATP (6000 mCi/mmol) and Cordycepin 5'-triphosphate 3'- α - ^{32}P (5000 mCi/mmol) were purchased from PerkinElmer Inc. (Boston, MA). T4 polynucleotide kinase, terminal deoxynucleotidyl transferase and deoxynucleoside 5'-triphosphates (dNTPs) were purchased from Thermo Fisher Scientific (Waltham, MA). Micro Bio-Spin 6 Columns were purchased from Bio-

Rad Laboratories (Hercules, CA). Pierce Avidin Agarose resin was from Thermo Fisher Scientific (Waltham, MA). QuikChange II XL Site-Directed Mutagenesis kit was purchased from Agilent Technologies (Santa Clara, CA). All other standard chemical reagents were from Thermo Fisher Scientific (Waltham, MA) and Sigma-Aldrich (St. Louis, MO). Purified enzymes including APE1, pol β , FEN1 and LIG I were made according to the procedures described previously [51, 145].

Oligonucleotide substrates

The 100 nt oligonucleotide substrates contain (CAG)₂₀ repeats or (CTG)₂₀ repeats with a tetrahydrofuran (THF), an analog of a modified abasic site. The THF residue substituted the first or tenth repeat unit of the (CAG)₂₀ or (CTG)₂₀ containing substrates. This mimics the scenario that the damage occurred at the 5'- end or in the middle of the repeat tract. Substrates were constructed by annealing the damaged strand with the template strand at a molecular ratio of 1:2. A strand of DNA fragment containing (CAG)₂₀ or (CTG)₂₀ repeats without damage was used as a size marker. The sequences of substrates are listed in Table 1.

Table 1. Oligonucleotide Sequences [13]

Oligonucleotide	nt	Sequence (5'-3')
Undamaged strand		
CAG-0 (undamaged)	100	CGA GTC ATC TAG CAT CCG TA CAG TA CGT AGA CTT ACT CAT TGC
CTG-0 (undamaged)	100	GCA ATG AGT AAG TCT ACG TA CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG

		CTG CTG CTG TA CGG ATG CTA GAT GAC TCG
<hr/>		
THF containing substrate		
CAG-1 (THF 5'-end)	99	CGA GTC ATC TAG CAT CCG TA CATHF CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG TA CGT AGA CTT ACT CAT TGC
CAG-10 (THF middle)	99	CGA GTC ATC TAG CAT CCG TA CAG CAG CAG CAG CAG CAG CAG CAG CAG CATHF CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG TA CGT AGA CTT ACT CAT TGC
CTG-1 (THF 5'-end)	99	CGA GTC ATC TAG CAT CCG TA CTTHF CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG TA CGT AGA CTT ACT CAT TGC
CTG-10 (THF middle)	99	GCA ATG AGT AAG TCT ACG TA CTG CTG CTG CTG CTG CTG CTG CTG CTG CTTHFCTG CTG CTG CTG CTG CTG CTG CTG CTG CTG TA CGG ATG CTA GAT GAC TCG
<hr/>		
Template strand		
CAG-T (CAG-1 and CAG-10 template)	100	GCA ATG AGT AAG TCT ACG TA CTG TA CGG ATG CTA GAT GAC TCG
CTG-1-T (CTG-1 template)	100	GCA ATG AGT AAG TCT ACG TAC AG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG TA CGG ATG CTA GAT GAC TCG
CTG-10-T (CTG-10 template)	100	CGA GTC ATC TAG CAT CCG TA CAG TA CGT AGA CTT ACT CAT TGC
<hr/>		
1-nt gap substrate		
CAG-gap-upstream	50	CGA GTC ATC TAG CAT CCG TA CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG
CAG-gap-downstream	49	AG CAG CAG CAG CAG CAG CAG CAG CAG CAG TA CGT AGA CTT ACT

CAG-gap-template	100	CAT TGC GCA ATG AGT AAG TCT ACG TA CTG TA CGG ATG CTA GAT GAC TCG
Random-gap-upstream	50	CGA GTC ATC TAG CAT CCG TA TCG CAC TGT TAT CAT TTC GTG TAC TTC ATG
Random-gap-downstream	49	GT ATG TGT CAT ATA TTC ATT TGC GCT AAC TA CGT AGA CTT ACT CAT TGC
Random-gap-template	100	5'- GCA ATG AGT AAG TCT ACG TAG TTA GCG CAA ATG AAT ATA TGA CAC ATA CAC ATG AAG TAC ACG AAA TGA TAA CAG TGC GAT ACG GAT GCT AGA TGA CTC G
<hr/>		
Nick substrate		
Nick-CAG-1-up	23	CGA GTC ATC TAG CAT CCG TA CAG
Nick-CAG-1-down	77	CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG TA CGT AGA CTT ACT CAT TGC
Nick-CAG-1-Template	100	GCA ATG AGT AAG TCT ACG TA CTG TA CGG ATG CTA GAT GAC TCG
Nick-CAG-10-up	50	CGA GTC ATC TAG CAT CCG TA CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG
Nick-CAG-10-down	50	CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG TA CGT AGA CTT ACT CAT TGC
Nick-CAG-10-Template	100	GCA ATG AGT AAG TCT ACG TA CTG TA CGG ATG CTA GAT GAC TCG
<hr/>		
Primers for DNA fragment analysis		
CAG reverse primer	23	6-FAM -CAA TGA GTA AGT CTA CGT ACT GC
CAG forward primer	21	CGA GTC ATC TAG CAT CCG TAC
CTG-1 reverse primer	23	6-FAM -CAA TGA GTA AGT CTA CGT

		ACA GC
CTG-1 forward primer	20	CGA GTC ATC TAG CAT CCG TAC
CTG-10 reverser primer	19	6-FAM- CGA GTC ATC TAG CAT CCG TAC
CTG-10 forward primer	21	GCA ATG AGT AAG TCT ACG TA
Primers for site directed mutagenesis		
Pol beta R137Q fwd	44	AAT TGA ACC ATC ATC AGC AGA TTG GG CTG A AAT ATT TTG GGG AC
Pol beta R137Q rev	44	GTC CCC AAA ATA TTT CAG CCC AAT CTG CTG ATG ATG GTT CAA TT

Construction of pol β R137Q variant expression vector and expression and purification of pol β R137Q variant protein

The expression vector of pol β R137Q variant was constructed by site-directed mutagenesis using the encoding region of wild-type pol β -(His)₆ cloned in pET15b vector as the template. Site-directed mutagenesis was conducted with the QuickChange II XL Site-Directed Mutagenesis Kit. A forward PCR primer and a reverse primer (Table S1) were used for PCR reactions under the conditions as follows: 1 cycle of 95°C for 30 s; then 16 cycles of 95°C for 30 s, 52°C for 1 min, 68°C for 7 min. The expression vector with pol β R137Q variant with a (His)₆ tag was then transformed into *E. coli* BL21DE3 (Aligent Technologies, Santa Clara, CA) for their expression according to the procedures described previously [51, 145]. Briefly, cell pellets were resuspended in the lysis buffer that contains 50 mM NaH₂PO₄, 30 mM NaCl, 10 mM imidazole, 10 mM ethylenediaminetetraacetic acid (EDTA), 10 mM dithiothreitol (DTT), 1 mM phenylmethylsulfonyl fluoride (PMSF), 1 mM benzamidin, 1 μ g/ml leupeptin and 1 μ g/ml pepstatin A. Soluble proteins and cell debris were then separated by centrifugation at 12,000 rpm, 4°C for 30 min. The supernatant was subjected to Ni-NTA agarose column

from Qiagen (Hilden, Germany) for purification. Proteins were eluted by elution buffer containing 30 mM 4-(2-hydroxyethyl)-1-piperazineethane-sulfonic acid (HEPES), pH7.5, 300 mM NaCl, 500 mM imidazole, 10 mM EDTA and 10 mM DTT. Peak fractions were collected and dialyzed into buffer that contains 30 mM HEPES, pH 7.5, 0.5% inositol, 1.7M (NH₄)₂SO₄ and 1 mM PMSF. Proteins were separated by phenyl sepharose 6 fast flow column (GE Healthcare Bio-Science, Uppsala, Sweden). Eluted peak fractions were combined and dialyzed into 30 mM HEPES, pH 7.5, 0.5% inositol, 30 mM KCl, 1 mM EDTA and 1 mM PMSF. Proteins were then subjected to purification with the Q sepharose (GE Healthcare Bio-Science, Uppsala, Sweden). Peak fractions were combined and dialyzed into storage buffer that contains 30 mM HEPES, 100 mM KCl, 20% glycerol and 1mM PMSF, aliquoted and frozen at -80°C for storage.

***In vitro* reconstituted BER**

Reconstituted BER was performed by incubating purified APE1, wild-type pol β or pol β R137Q variant, along with FEN1, LIG I and (CAG)₂₀ or (CTG)₂₀ substrates (25 nM) containing a THF residue in a reaction mixture (20 μ l) that contained 5 mM MgCl₂, 50 μ M dNTPs, 2 mM ATP, 50 mM Tris-HCl (pH 7.5), 50 mM KCl, 0.1 mg/ml BSA, 0.1 mM EDTA and 0.01% NP-40. Reaction mixtures were assembled on ice and incubated at 37°C for 15 min. Reactions were terminated with 2X stopping buffer (95% formamide and 10 mM EDTA) and incubation at 95°C for 10 min. Substrates and products were then separated in 15% urea-denaturing polyacrylamide gel (PAGE) and detected by a Pharos FX Plus PhosphorImager (Bio-Rad Laboratory, CA).

Probing of secondary structures formed in a TNR tract by S1 nuclease

The formation of hairpin structures on the damaged and template strands of the (CAG)₂₀ or (CTG)₂₀ substrate was probed with the S1 nuclease that makes cleavage specifically on a single-strand DNA (Promega Life Science, Madison, WI). 100 nM substrates containing a THF residue that substitutes the G at the first or tenth repeat were initially incubated with 10 nM APE1 in the absence or presence of 5 nM wild-type pol β or pol β R137Q for 30 min in the BER buffer as described previously. Subsequently, the 10 μ l reaction mixtures were subjected to S1 nuclease digestion at 37°C for 1, 3, 5, 10, 15 min, respectively with 5 μ l S1 nuclease reaction mixtures that contained optimized concentrations of S1 nuclease in the buffer containing 30 mM sodium acetate (pH 4.6), 50 mM NaCl, 1 mM ZnCl₂, 0.5 mg/ml denatured calf thymus DNA and 5% glycerol. The S1 nuclease digestion was optimized by employing 25 U and 2 U of S1 nuclease for the digestion of the the damaged strand and template strand of the (CAG)₂₀ containing substrate that contains a THF located at the first repeat, respectively; and by employing 3 U, 25 U and 5 U of S1 nuclease for the digestion of the upstream damaged strand, downstream damaged strand and the template strand of the (CAG)₂₀ containing substrate that contains a THF located at the tenth repeat, respectively. Reactions were then quenched with addition of 2 μ g proteinase K and incubation at 55°C for 30 min. Reaction mixtures were then mixed with the same volume of 2X stopping buffer and denatured at 95°C for 10 min. Substrates and products were separated in an 18% urea-denaturing PAGE and detected by a PhosphorImager.

Enzymatic activity assay

Pol β DNA synthesis activity were determined by incubating wild-type pol β or pol β R137Q variant in the absence or presence of PCNA with 25 nM (CAG)₂₀ or (CTG)₂₀ substrate with a THF residue in 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 along with 50 μ M dNTPs and 5 mM Mg²⁺ at 37°C for 15 min. FEN1 cleavage activity on the substrates was measured in the absence and presence of wild-type pol β or pol β R137Q variant at 37°C for 15 min in reaction buffer under the conditions described previously. The 20 μ l of reaction mixture was assembled on ice, and reaction was quenched by addition of 20 μ l stopping buffer and incubation at 95°C for 10 min. Substrates and products were separated by a 15% urea denaturing gel and detected by a PhosphoImager.

Sizing analysis of BER products by DNA fragment analysis

To isolate a repaired strand specifically, the template strand of all substrates was tagged by a biotin residue at the 5'-end. BER reactions were terminated with 1 μ l of 100 mM EDTA, and reaction mixtures were incubated with 50 μ l avidin agarose beads (Pierce-Thermo Scientific, Rockford, IL) for 2 hrs, allowing the binding of avidin beads to the biotin on the template strand. Reaction mixtures were then subjected to incubation with 0.15 M NaOH at room temperature for 15 min with rotation allowing separation of repaired strands from the template strands. This was followed by 2 min centrifugation at 5000 rpm pelleting the template strands bound by avidin beads. The repaired strands in the supernatant were precipitated with ethanol and subsequently dissolved in TE buffer

for PCR amplification and size analysis. Repaired products were amplified through PCR with the AmpliTaq Gold 360 DNA polymerase Kit (Applied Biosystems, Foster City, CA) at the conditions: denaturation at 95°C for 30 s, annealing at 50°C for 30 s, extension at 72°C for 90 s for 35 cycles with a final extension at 72°C for 1 hr. CAG forward primer and 6-carboxyfluorescein (6-FAM) tagged CAG reverse primer, were used for PCR amplification of the repaired products from the (CAG)₂₀ substrate with a THF at 5'-end or in the middle of the repeat tract. CTG-1 forward primer and 6-FAM tagged reverse primer were used for PCR amplification of the repaired products from the (CTG)₂₀ substrate with a THF at 5'-end of the repeat tract; while CTG-10 forward primer and 6-FAM tagged reverse primer were used to amplify the repaired products of (CTG)₂₀ substrate with a THF in the middle of the repeat tract. The sequences of the primers are indicated in Table1. PCR products along with the size marker MapMarker 1000 (Bioventures, Murfreesboro, TN) were then subject to capillary electrophoresis via an ABI 3130XL Genetic Analyzer (Applied Biosystems, Foster City, CA) at Florida International University DNA sequencing core facility. The sizes of PCR products were determined by DNA fragment analysis with the Peak Scanner version 1.0 software (Applied Biosystems, Foster City, CA).

RESULTS

Pol β R137Q polymorphic variant exhibits weaker DNA synthesis activity than wild-type pol β in the context of CAG and CTG repeats

Since pol β DNA synthesis plays a crucial role in mediating TNR expansion [1] and deletion during BER [211] in a damage location dependent manner [210], we initially

examined the DNA synthesis activity of pol β R137Q with the (CAG)₂₀ or (CTG)₂₀ repeat substrate containing an abasic site (THF) located at the 5'-end or in the middle of the repeat tract. We found that with the damage located at the 5'-end of the (CAG)₂₀ repeat tract, pol β R137Q and wild-type pol β at 1 nM, mainly inserted 1 and 3 nucleotides (Figure 1.1A, compare lane 5 with 3). However, 5 nM pol β R137Q variant inserted up to 4 repeats (Figure 1.1A, lane 6), whereas the same concentration of wild-type pol β inserted up to 6 repeats (Figure 1.1A, lane 4). With the damage located in the middle of the repeat tract, pol β R137Q inserted up to 3 nucleotides and wild-type pol β inserted up to 5 nucleotides (Figure 1.1B, compare lane 11 with lane 9). However, 5 nM pol β R137Q variant inserted up to 4 repeats, whereas the same concentration of wild-type pol β inserted up to 5 repeats (Figure 1.1B, compare lane 12 with lane 10). Similarly, for the (CTG)₂₀ substrate with the damage at the 5'-end, 1 nM pol β R137Q mainly inserted 1 nucleotide, while the same concentration of wild-type pol β mainly inserted 3 nucleotide on the substrate (Figure 1.1C, compare lane 6 and lane 3). 2.5 nM and 5 nM pol β R137Q inserted up to 2 and 3 repeats with the same substrate while wild-type pol β inserted up to 3 or 5 repeats (Figure 1.1C, compare lanes 7 and 8 with lanes 4 and 5). With the damage located in the middle of the (CTG)₂₀ substrate, 1 nM and 2.5 nM pol β R137Q variant inserted 2 nucleotides and 1 repeat, respectively (Figure 1.1D, lanes 14-15), whereas the wild-type pol β at the same concentrations inserted 1 nucleotide and up to 2 repeats, respectively (Figure 1.1D, lanes 11-12). At the concentration of 5 nM, pol β R137Q variant inserted up to 3 repeats, while wild-type pol β inserted up to 5 repeats (Figure 1.1D, compare lane 16 with lane 13). The results indicated that pol β R137Q exhibited weaker DNA synthesis activity than wild-type pol β in the context of TNRs.

Figure 1.1 Pol β R137Q variant exhibits weak DNA synthesis activity during BER of a base lesion in context with (CAG)₂₀ and (CTG)₂₀ repeats [13].

DNA synthesis activity of wild-type pol β (pol β WT) or pol β R137Q variant at various concentrations, was measured with 25 nM substrates containing (CAG)₂₀ and (CTG)₂₀ with a THF residue. DNA synthesis activity of pol β on the (CAG)₂₀ and (CTG)₂₀ repeat substrates was measured at the concentrations of 1 nM, 2.5 nM and 5 nM. Substrates were ³²P-labeled at the 5'-end of the damaged strand. **(A)** and **(B)** Measurement of DNA synthesis activity with (CAG)₂₀ substrate containing a THF at the first and tenth repeat, respectively. Lanes 1 and 11 represent substrate only. Lane 2 and 12 represent APE1 cleavage products. Lanes 3-4 and 13-14 represent pol β WT synthesized products. Lanes 5-6 and 15-16 represent pol β R137Q variant synthesized products. **(C)** and **(D)** Measurement of DNA synthesis activity assay in the context with (CTG)₂₀ substrate with a THF at the first and tenth repeat, respectively. Lanes 1 and 15 represent substrate only. Lane 2 and 16 represent the reaction with APE1 along. Lanes 3-5 and 17-19 illustrate the reactions with pol β WT. Lanes 6-8 and 20-22 represent reactions with pol β R137Q. The red circles superimposed in the gels indicate the locations of the synthesized products of pol β WT or pol β R137Q.

To compare the gap-filling synthesis activity of pol β R137Q variant and wild-type pol β at low concentrations, the gap-filling synthesis of pol β R137Q variant and wild-type pol β at the concentrations of 0.1 nM, 0.2 nM and 0.5 nM was measured with the substrates containing the (CAG)₂₀ and (CTG)₂₀ with a THF at the 5'-end or in the middle of the repeat tract. For the (CAG)₂₀ and (CTG)₂₀ substrates with a THF at the 5'-end, both wild-type pol β and pol β R137Q incorporated 1 nucleotide at the concentrations of 0.1 nM and 0.2 nM (Figures 1.2A and 1.2C, lanes 3-4 and 6-7). Moreover, wild-type pol β and pol β R137Q variant at 0.5 nM incorporated up to 3 nucleotides (Figures 1.2A and 1.2C, lanes 5 and 8). Similarly, for the (CAG)₂₀ and (CTG)₂₀ substrates with a THF in the middle, wild-type pol β and pol β R137Q variant at the concentrations of 0.1 nM, 0.2 nM and 0.5 nM mainly inserted one nucleotide (Figures 1.2B and 1.2D, lanes 11-16). The

results indicated that pol β R137Q variant exhibited similar gap-filling synthesis activity as wild-type pol β at low concentrations. We also examined the gap-filling synthesis activity of wild-type pol β and pol β R137Q variant at low concentrations of 0.1 nM, 0.2 nM and 0.5 nM on the 1 nt-gap substrates with (CAG)₂₀ repeats or a random DNA sequence. For the 1-nt gap substrate with (CAG)₂₀ repeats, both wild-type pol β and pol β R137Q variant mainly inserted one nucleotide to fill in the gap at all tested concentrations (Figure 1.2E, lanes 2-7). For the substrate containing a random DNA sequence, both wild-type pol β and pol β R137Q inserted one nucleotide to fill in the gap at concentrations of 0.1 nM and 0.2 nM (Figure 1.2F, lanes 9-10 and 12-13). Furthermore, both wild-type pol β and pol β R137Q at 0.5 nM inserted additional 2 nucleotides to displace the downstream flap (Figure 1.2F, lanes 11 and 14). The results indicated that pol β R137Q variant at low concentrations exhibited slightly weaker one-nucleotide gap-filling synthesis compared to that of wild-type pol β on the 1-nt gap substrate with (CAG)₂₀ repeats (Figure 1.2E) and similar gap filling synthesis and strand displacement synthesis activity with the wild-type polymerase on the 1-nt gap substrates with a random DNA sequence (Figure 1.2F).

Figure 1.2. Pol β R137Q variant exhibits similar gap-filling synthesis activity compared to that of wild-type pol β [13].

Pol β gap-filling synthesis was conducted by incubating 0.1 nM, 0.2 nM and 0.5 nM of wild-type pol β or R137Q variant with 25nM 32 P-labeled substrates (5'-end labeled). The experimental conditions are described in the Materials and Methods. **(A)** and **(B)** The gap-filling synthesis of pol β WT or pol β R137Q variant on the (CAG)₂₀ substrates that contains a THF at the 5'-end or in the middle of the repeat tract. **(C)** and **(D)** The gap-filling synthesis of pol β WT or pol β R137Q variant on the (CTG)₂₀ substrates that contains a THF at the 5'-end or in the middle of the repeat tract. Lanes 1 and 9 represent substrate only. Lanes 2 and 10 represent APE1 cleavage products. Lanes 3-5 and 11-13 represent pol β WT synthesized products. Lanes 6-8 and 14-16 represent pol β R137Q variant synthesized products. **(E)** The gap-filling synthesis of pol β WT or pol β R137Q variant on 1-nt gap substrate containing (CAG)₂₀. **(F)** The gap-filling synthesis of pol β WT or pol β R137Q variant on the 1-nt gap substrate containing a random sequence. Lanes 1 and 8 represent substrate only. Lanes 2-4 and 9-11 represent pol β WT synthesized products. Lanes 5-7 and 12-14 represent pol β R137Q variant synthesized products. The red circles superimposed in the gels indicate the synthesized products of pol β WT or pol β R137Q.

PCNA does not affect DNA synthesis activity of pol β R137Q variant

As a BER cofactor, PCNA can stimulate the activities of both FEN1 and LIG I during long patch BER [234, 235]. However, it remains unknown whether PCNA may affect pol β DNA synthesis activity although it has been found that PCNA can physically interact with pol β [236]. Since pol β R137Q variant has weak interaction with PCNA [61], it is possible that this can affect pol β R137Q DNA synthesis activity, promoting TNR instability. To test this, high concentrations of PCNA (50 nM and 100 nM) were employed to determine if PCNA could stimulate pol β DNA synthesis activity in the context of TNRs at a low concentration (0.2 nM) of wild-type pol β or pol β R137Q variant. We found that in the absence of PCNA or presence of 50 nM and 100 nM PCNA, at 0.2 nM, both pol β R137Q and wild-type pol β generated a similar amount of one nucleotide insertion product during BER of an abasic lesion (THF) located at either the 5'-end (Figure 1.3A, compare lanes 4-5 with lane 3 and lanes 7-8 with lane 6) or in the

middle of the (CAG)₂₀ repeat substrates (Figure 1.3B, compare lanes 12-13 with lane 11 and lanes 15-16 with lane 14). The results showed that at the high concentration of PCNA, no stimulatory effect was detected for the gap-filling synthesis activity of wild-type pol β and pol β R137Q variant (Figure 1.3), indicating that PCNA did not affect the processivity of pol β WT and pol β R137Q variant either. This further demonstrates that the impaired interaction between PCNA and pol β R137Q variant did not affect the polymerase gap-filling synthesis activity of the pol β variant.

Figure 1.3. PCNA does not affect DNA synthesis activity of pol β WT and pol β R137Q during BER of a base lesion on (CAG)₂₀ repeats [13].

The effects of PCNA on pol β was examined by determining pol β DNA synthesis in the presence of 0.2 nM of pol β WT or pol β R137Q along with 50 nM and 100 nM PCNA. 25 nM substrates were ³²P-labeled at the 5'-end of the damaged strand of the substrates. **(A)** and **(B)** Pol β DNA synthesis in the presence of PCNA with the (CAG)₂₀ substrate containing a THF located at the first repeat and tenth repeat, respectively. Lanes 1 and 9 represent substrate only. Lanes 2 and 10 represent reaction with APE1. Lanes 3 and 11 illustrate reactions with pol β WT. Lanes 6 and 14 represent reactions with pol β R137Q. Lanes 4-5 and 12-13 represent reactions with pol β WT in the presence of increasing concentrations of PCNA. Lanes 7-8 and 15-16 represent reaction mixtures with pol β R137Q in the presence of increasing concentrations of PCNA.

The weak DNA synthesis of pol β R137Q variant leads to weak FEN1 cleavage during BER in a TNR tract

Since pol β DNA synthesis in the context of TNR tracts facilitates FEN1 cleavage by creating a TNR flap during BER [207, 210, 211], we then examined the effect of pol β R137Q on FEN1 cleavage activity by measuring FEN1 activity in the presence of 1 nM or 5 nM pol β R137Q variant during repair of an abasic site located at the 5'-end and in the middle of the (CAG)₂₀ repeat substrates (Figure 1.4). We found that in the absence of pol β R137Q variant or wild-type pol β , FEN1 cleaved up to 3 nucleotides at 5 nM and up to 5 nucleotides at 10 nM on the substrate with the lesion at the 5'-end (Figure 1.4A, lanes 3-4). However, FEN1 still removed up to 3 nucleotides at 5 nM and up to 5 nucleotides at 10 nM in the presence of 1 nM pol β R137Q variant or wild-type pol β (Figure 1.4A, compare lanes 5-8 with lanes 3-4). In the presence of 5 nM pol β R137Q, the same concentrations of FEN1 (5 nM and 10 nM) removed up to 2 and 3 CAG repeats from the CAG repeat substrate (Figure 1.4A, lanes 11-12), whereas in the presence of 5 nM wild-

type pol β , FEN1 at 5 nM and 10 nM removed up to 3 and 4 repeats, respectively (Figure 1.4A, lanes 9-10). Similarly, FEN1 cleavage activity on the substrate with the lesion in the middle of the repeat tract, was not altered significantly in the presence of 1 nM pol β R137Q and wild-type pol β by comparing to in the absence of pol β R137Q and wild-type pol β , FEN1 cleaved up to 3 and 5 nucleotides at 5 nM and 10 nM in the absence of pol β , respectively (Figure 1.4B, lanes 14-15); with additional 1 nucleotide cleaved by FEN1 in the presence of 1 nM pol β R137Q variant and wild-type pol β at the concentration of 10 nM of FEN1 (Figure 1.4B, compare lane 18 and lane 20 with lane 16). However, in the presence of 5 nM pol β R137Q, FEN1 (5 nM and 10 nM) mainly cleaved 2 and 3 CAG repeats, respectively, compared to 3 and 4 CAG repeats, respectively, removed by FEN1 (5 nM and 10 nM) in the presence of 5 nM wild-type pol β (Figure 1.4B, compare lanes 23-24 with lanes 21-22). The results indicate that the weak DNA synthesis of pol β R137Q resulted in weaker FEN1 cleavage of TNRs than that from wild-type pol β by creating a shorter repeat flap.

Figure 1.4. Pol βR137Q leads to weak FEN1 cleavage during BER in the context of (CAG)₂₀ repeats [13].

FEN1 flap cleavage activity at 5 nM and 10 nM on the (CAG)₂₀ substrate with a THF located at the first repeat (**A**) or tenth repeat (**B**) was measured in the presence of 1 nM and 5 nM pol β. Lanes 1 and 13 represent the substrate only. Lanes 2 and 14 represent reactions with APE1 alone. Lanes 3-4 and lanes 15-16 represent reactions with FEN1 alone. Lanes 5-6 and 17-18 represent reactions with FEN1 in the presence of 1 nM pol β WT. Lanes 7-8 and 19-20 illustrate reactions with FEN1 in the presence of 1 nM pol βR137Q. Lanes 9-10 and 21-22 represent reactions with FEN1 in the presence of 5 nM pol β WT. Lanes 11-12 and 23-24 illustrate reactions with FEN1 in the presence of 5 nM of pol βR137Q variant. Substrates were ³²P-labeled at the 3'-end of the damaged strand. The red circles superimposed in the gels indicate the locations of FEN1 cleavage products.

Pol βR137Q variant leads to small TNR deletions during BER of an abasic site in a TNR tract.

To further determine whether the weak DNA synthesis of pol βR137Q variant may affect trinucleotide repeat instability, we reconstituted BER of an abasic lesion in the context of a TNR tract with the pol β variant and the (CAG)₂₀ or (CTG)₂₀ substrates containing the lesion at the 5'-end or in the middle of the repeat tract and measured the

length change of the repeats (Figure 1.5). We found that 1 nM and 5 nM wild-type pol β and pol β R137Q led to repaired products during BER of an abasic lesion in the context of a (CAG)₂₀ (Figure 1.5A, lanes 3-4 and lanes 9-10 for wild-type pol β ; Figure 1.5A lanes 5-6 and lanes 11-12 for pol β R137Q) or (CTG)₂₀ repeat tract (Figure 1.5B, lanes 3, 5, 11 and 13 for wild-type pol β ; Figure 1.5B lanes 6, 8, 14 and 16 for pol β R137Q). Further analysis on the repeat size of the repaired products showed that BER reconstituted with 1 nM pol β R137Q or wild-type pol β resulted in a product with one repeat deletion with the CAG and CTG substrates containing the base lesion at the 5'-end (Figure 1.5C and 1.5D, left panel, panel b and c), whereas 5 nM pol β R137Q or wild-type pol β led to full length repaired products (Figure 1.5C and 1.5D, left panel, panel d and e). BER reconstituted with 1 nM pol β R137Q or wild-type pol β and the substrates with the lesion in the middle, resulted in the repaired products with one or two repeat deletion (Figure 1.5C and 1.5D, right panel, panel b and c). 5 nM pol β R137Q and wild-type enzyme decreased the amount of the deletion product (Figure 1.5C and 1.5D, right panel, panel d and e). The results indicate that pol β R137Q exhibited the same ability as wild-type pol β enzyme to modulate TNR instability by causing one or two repeat deletions.

Figure 1.5. Effects of pol β R137Q on the instability of (CAG)₂₀ and (CTG)₂₀ repeats [13].

BER was reconstituted with 10 nM of APE1, 10 nM FEN1, 5 nM LIG I and increasing concentrations of pol β WT or pol β R137Q (1 nM and 5 nM on (CAG)₂₀ substrate and 1 nM, 2.5 nM and 5 nM on (CTG)₂₀ substrate) along with 25 nM substrates. **(A)** BER reconstituted with the (CAG)₂₀ substrate with a THF located at the first repeat (left panel) or the tenth repeat (right panel). Lane 1 and lane 7 represent substrates only. Lane 2 and 8 represent reactions with APE1 only. Lanes 3-4 and 9-10 represent reactions with pol β WT. Lanes 5-6 and 11-12 represent reactions in the presence of pol β R137Q. **(B)** BER reconstituted with the (CTG)₂₀ substrate with a THF located at the first repeat (left panel) and the tenth repeat (right panel). Lane 1 and lane 9 represent substrates only. Lane 2 and 10 represent reactions with APE1 only. Lanes 3-5 and 11-13 present reactions with pol β WT. Lanes 6-8 and 14-16 illustrate reactions with pol β R137Q. Repaired products were subjected to PCR and capillary electrophoresis and analyzed by DNA fragment analysis to determine the length. **(C)** DNA fragment analysis of the repaired products of from the substrate with a lesion located at the first (left panel) and the tenth of the (CAG)₂₀ substrate. **(D)** DNA fragment analysis of the repaired products of from the substrate with a lesion located at the first (left panel) and the tenth of the (CTG)₂₀ substrate. Panel a represent a (CAG)₂₀ or (CTG)₂₀ marker without damage. Panel b and c indicate the sizes of the repaired products in the presence of 1 nM pol β WT or 1 nM R137Q, respectively. Panel d and e represent the sizes of the repaired products in the presence of 5 nM pol β WT or 5 nM pol β R137Q, respectively.

Pol β R137Q variant skips over a small loop on the template strand of a TNR repeat tract

Our previous studies have shown that secondary structures such as hairpins and loops can form readily during BER in a TNR tract [1, 7, 210, 211], and pol β can readily skip over the structures on the template strand [211]. The “skip over” refers to a scenario where the polymerase encountered the secondary structures such as a hairpin or loop structure formed on the template strand, it performed DNA synthesis to bypass the secondary structures other than copying through the inside of the secondary structures. We have further demonstrated that pol β DNA synthesis modulates TNR expansion and deletion by altering the balance between the synthesis of TNRs and removal of the

repeats by FEN1 flap cleavage [210, 211]. Since the weak DNA synthesis activity of pol β R137Q resulted in weak FEN1 cleavage activity, it is possible that the pol β polymorphic variant can alter TNR instability. To test this possibility, a single-stranded DNA specific endonuclease *Aspergillus* S1 nuclease [237] was used to probe the secondary structures that formed on the template strand of the (CAG)₂₀ repeat substrate in the absence or presence of DNA synthesis of pol β at the time intervals of 1 min-15 min. We found that in the absence of pol β , S1 cleavage on the template strand of the (CAG)₂₀ substrate with a THF located at the 5'-end, resulted in the products of 18 nt, 19 nt, 20nt, 21 nt, 22 nt and 23 nt (Figure 1.6A, left panel, lanes 3-7). This indicated the formation of a small loop containing one CTG repeat adjacent to the 3'-end of the flanking region of the repeat tract (Figure 1.6A, the scheme below the gel). In the presence of DNA synthesis of wild-type pol β and pol β R137Q variant, S1 nuclease cleavage led to the products of 21 nt, 23 nt, 24 nt, 26 nt and 27 nt indicating the formation of a loop containing two CTG repeats (Figure 1.6A, the panels in the middle and on the right, lanes 3-7). The results indicated that in both the absence and presence of DNA synthesis of both pol β R137Q variant and wild-type pol β , a small (CTG)₂ loop formed on the template strand. This further indicated that the DNA synthesis of pol β R137Q variant and wild-type pol β only altered the position of the loop by shifting the loop toward the 5'-end of the template. We also observed that in the presence of DNA synthesis of wild-type pol β and pol β R137Q variant, the S1 nuclease cleavage products were much weaker than those generated in the absence of pol β DNA synthesis. This was because pol β DNA synthesis copied through the template strand, which converted the S1 nuclease-sensitive ssDNA loop region on the template strand to a S1 nuclease-resistance dsDNA region.

Furthermore, the pol β DNA synthesis displaced the downstream damaged strand, exposing the annealed template strand as a ssDNA region. This then allowed the ssDNA region on the template strand approximately 1 repeat to shift towards the 5'-end. This is consistent with the 1 repeat insertion observed in DNA synthesis (Figure 1.1A). Further analysis of S1 nuclease cleavage on the downstream strand of the damaged strand of the substrate showed that S1 nuclease resulted in the products of 64 nt, 67 nt, 70 nt, 73 nt and 76 nt in the absence and presence of both pol β R137Q and wild-type pol β indicating that the downstream strand formed a single-stranded (CAG)₅ flap (Figure 1.6B, lanes 3-7, and schemes below the gels), and DNA synthesis pol β R137Q variant and wild-type pol β did not alter the formation of the downstream flap.

Figure 1.6. S1 nuclease probing of loops formed in (CAG)₂₀ repeats with a THF located at the first CAG of the (CAG)₂₀ repeat substrate in the absence and presence of pol β WT or pol β R137Q DNA synthesis [13].

The S1 nuclease probing of a loop located at the template strand (A) and the downstream damage strand of the (CAG)₂₀ substrate. (B) Both the template strand and damage strand were labeled at the 3'-end. Substrates were incubated with 2 U or 25 U S1 nuclease. Panels on the left, in the middle and on the right, correspond to S1 nuclease cleavage results in the presence of APE1 alone, pol β WT and pol β R137Q, respectively. Lane 1 represents the substrate only. Lane 2 represents the reaction with APE1 alone. Lanes 3-7 represent reaction mixtures with S1 nuclease in the absence or the presence of pol β DNA synthesis at different time intervals. Lane 8 represents a synthesized marker with different sizes of DNA fragments. The schemes below the gels illustrate S1 nuclease cleavage patterns in the absence and presence of pol β WT and pol β R137Q. The red circles superimposed in the gels indicate the locations of the S1 nuclease digestion products.

For the (CAG)₂₀ substrate with a THF in the middle of the repeat tract, S1 cleavage on the template resulted in the products of 45 nt, 48 nt and 49 nt at the time interval of 1-15 min in the absence of pol β (Figure 1.7A, left panel, lanes 3-7) indicating the formation of a (CTG)₂ loop (Figure 1.7A, left panel, the scheme below the gel). In the presence of DNA synthesis of wild-type pol β or pol β R137Q variant, S1 nuclease cleavage on the template strand resulted in the products of 48 nt, 49 nt, 50 nt and 51 nt. This also indicated the formation of a (CTG)₂ loop that was shifted toward the 5'-end of the template strand (Figure 1.7A, the panels in the middle and on the right, lanes 3-7). The size of the loop is consistent with deletion of one or two repeats resulting from BER reconstituted by wild-type pol β and pol β R137Q with the substrate (Figure 1.5C and 1.5D, the left panel). This further suggested that pol β skipped over the loop structure and performed DNA synthesis to displace the downstream strand generating a flap. To test this, we then examined S1 nuclease cleavage on the upstream strand of the damage strand of the (CAG)₂₀ substrate (Figure 1.7B). We found that in the absence of pol β , S1 nuclease cleavage resulted in the products of 47 nt, 46 nt and 45 nt that are shorter than

the APE1 cleavage product of 50 nt (Figure 1.7B, left panel, lanes 3-7), indicating the formation of an upstream flap containing two CAG repeats (Figure 1.7B, the scheme below the left panel). In the presence of S1 nuclease and DNA synthesis of wild-type pol β or pol β R137Q, a product of 53 nt, which is longer than the APE1 cleavage product of 50 nt, was detected (Figure 1.7B, the panels in the middle and on the right, lanes 3-7), indicating that both wild-type pol β and pol β R137Q efficiently pushed the upstream flap to re-anneal to the template strand and extended the upstream strand (Figure 1.7B, the schemes below the panels in the middle and on the right). This further indicated that pol β DNA synthesis displaced the downstream strand and created a flap. This can be confirmed by the FEN1 cleavage results showing that FEN1 cleaved more nucleotides in the presence of wild-type pol β or pol β R137Q variant than the ones in the absence of pol β (Figure 1.4B, compare lanes 18, 20, 22 and 24 to lane 16). The cleavage of the downstream strand of the substrate in the absence and presence of wild-type pol β or pol β R137Q variant resulted in the product of 39 nt, 42 nt, 45 nt and 48 nt (Figure 1.7C, left, middle and right panels, lanes 3-7). However, S1 nuclease cleaved more on the substrate in the absence of pol β and fewer on the substrate in the presence of pol β (Figure 1.7C, compare lanes 3-7 from the panels in the middle and on the right with lanes 3-7 in the panel on the left). This suggests that pol β bound to the strand break intermediate generated by APE1 5'-incision of the THF, which protected the 5'-flap downstream ssDNA from S1 nuclease cleavage. The results indicate that the weak DNA synthesis activity of pol β R137Q still exhibited the similar ability as wild-type pol β in skipping-over of a small loop as well as in performing strand-displacement synthesis to create a downstream flap during BER in the context of CAG repeats.

Figure 1.7. S1 nuclease probing of loops formed in (CAG)₂₀ repeats with a THF located at the tenth CAG of the (CAG)₂₀ repeat substrate in the absence and presence of pol β WT or pol β R137Q DNA synthesis [13].

S1 nuclease probing of a loop on the template strand (A), the upstream damaged strand (B) and the downstream damaged strand (C) of the (CAG)₂₀ substrate with a THF at the tenth repeat were conducted. The template strand and downstream strand of the damaged strand were labeled at the 3'-end, and the upstream damaged strand was labeled at the 5'-end. Substrates were incubated with 5 U, 3 U and 25 U S1 nuclease in the absence and presence of pol β WT and pol β R137Q. Panels on the left, in the middle and on the right, correspond to reactions with S1 nuclease and APE1 alone, in the presence of pol β WT and in the presence of pol β R137Q, respectively. Lane 1 represents the substrate only. Lane 2 represents the reaction with APE1 alone. Lanes 3-7 represent reaction mixtures with S1 nuclease and APE1 at different time intervals. Lane 8 represents synthesized markers. The schemes below the gels illustrate S1 nuclease cleavage pattern on the different strands of the substrate. The red circles superimposed in the gels indicate the S1 nuclease digestion products.

In order to exclude the possibility that the digestion bands may be the result of the “breathing effect”, we performed the control experiments with the nick substrates digested by S1 nuclease, where the nicks were located after the first CAG or after the tenth CAG of the (CAG)₂₀ substrates. Under the same experimental condition with Figure 1.6A and Figure 1.7A, no S1 digestion products that indicate the S1 cleavage at the site opposite the nick on the template strand detected (Figure 1.8A and 1.8B, lanes 2-7). This indicates that S1 cleavage did not result from the “breathing effect”. This further suggests that S1 nuclease cleavage products were the result of the skipping over of the template loop by the polymerases (Figures 1.6A and 1.7A, the panels in the middle and on the right, lanes 3-7).

Figure 1.8. S1 nuclease digestion of the (CAG)₂₀ substrates with a nick located after first CAG or after the tenth CAG [13].

Both substrates were labeled at the 5'-end of the template strand. (A) S1 nuclease digestion of the (CAG)₂₀ substrate with a nick located after first CAG. Substrates were incubated with 2 U S1 nuclease. (B) S1 nuclease digestion of (CAG)₂₀ substrate with a nick located after tenth repeat. Substrates were incubated with 5 U S1 nuclease. Lane 1 represents the substrate only. Lane 2 represents the reaction with APE1 alone. Lanes 3-7 represent reaction mixtures with S1 nuclease and APE1 at different time intervals. Lane 8 represents synthesized markers.

DISCUSSION

Recent studies have shown that the polymorphism of several DNA repair proteins is associated with TNR diseases and the age of disease onset. This includes the polymorphisms in OGG1 (Rs1052133) (Ser326Cys) [238, 239], MSH3 (Rs26279) [220, 239], XPC [239] and ERCC6 (rs2228528)[224]. This suggests that the polymorphisms of proteins of BER, mismatch repair (MMR) and nucleotide excision repair (NER) pathways can modulate TNR instability. In this study, for the first time, we showed that pol β R137Q exhibited weaker DNA synthesis activity than wild-type pol β in the context

of TNRs (Figure 1.1). However, the DNA synthesis activity of pol β R137Q was not affected by PCNA during BER (Figure 1.3) although it is reported that this pol β variant has lost the interaction with PCNA [61]. The weak DNA synthesis of the pol β polymorphic variant further led to weak FEN1 cleavage of TNR flaps (Figure 1.4). Yet we found that pol β R137Q exhibited the similar ability to that of wild-type pol β in mediating deletion of one or two TNRs during BER (Figure 1.5). We demonstrated that this was because the polymorphic variant exhibited similar capability to that of wild-type pol β of skipping over a small template loop structure during BER (Figure 1.6 and Figure 1.7).

A previous study has shown that pol β R137Q exhibits 30% of wild-type enzymatic activity in the context of random sequence, which in turn reduces BER capacity and increases cellular sensitivity to alkylating DNA damaging agents as well as promotes apoptosis [61]. This effect has been further confirmed in a pol β R137Q transgenic mouse model *in vivo* in a recent study [62]. The results of the studies suggest that the pol β polymorphic variant exhibits significantly reduced DNA synthesis activity in random DNA sequence compared with wild-type pol β , and this may result in genome instability. However, our results showed that pol β R137Q polymorphic variant exhibited similar one-nucleotide gap-filling synthesis and strand displacement synthesis (Figure 1.2) and slightly weaker DNA synthesis compared with the wild-type enzyme in the context of TNRs. Thus, pol β R137Q variant exhibited similar ability in skipping over a template loop (Figure 1.7) with the wild-type enzyme, thereby resulting in the same size of small repeat deletions as wild-type pol β . Since the dynamic TNR tracts readily form secondary structures such as hairpins and loops on the damaged and template strand,

which would not form on random DNA sequence during BER [145, 207, 210, 211], Our results further indicate that secondary structures formed on the damaged and template strands of a TNR tract facilitated pol β R137Q to skip over a secondary structure on the template strand to readily perform DNA strand-displacement synthesis, thereby leading to the same effect on TNR instability as the wild-type enzyme.

A previous study has shown that pol β R137Q variant exhibits the impaired interaction with PCNA [61]. Our results showed that PCNA did not affect gap-filling synthesis activity of both wild-type pol β and pol β R137Q variant in the context of a TNR tract (Figure 1.3). This indicates that the impaired interaction between pol β R137Q variant and PCNA did not affect pol β activity and its resulted TNR instability during BER. However, our results cannot rule out a possibility that PCNA with post-translational modifications may alter the activity of wild-type pol β and pol β R137Q variant to modulate TNR instability during BER. Moreover, it is possible that the interaction between PCNA and other DNA polymerases can still modulate TNR instability by stimulating the activities of the DNA polymerases via cooperation with pol β DNA synthesis during BER. Previous studies have shown that ubiquitinated PCNA functions as a platform for polymerase switching between replicative DNA polymerases such as pol δ and translesion DNA polymerases such as pol η when encountered a DNA base lesion [240, 241]. Thus, it is conceivable that polymerase switching between replication DNA polymerases and translesion DNA polymerases that is mediated by ubiquitinated PCNA may occur when a hairpin or loop structure formed in the template strand during DNA replication and repair in the context of TNRs. It is also possible that PCNA-mediated polymerase switching between pol β and translesion DNA polymerases may occur during

BER in a TNR tract in the postmitotic cells. It is of interest to further elucidate the roles of the interaction between ubiquitinated PCNA and pol β and its polymorphic variants as well as translesion polymerases in modulating TNR instability during BER.

In summary, in this study, we provided the first evidence that pol β R137Q polymorphic variant exhibited weaker DNA synthesis than wild-type pol β in the context of a TNR tract. We showed that the pol β polymorphic variant led to a weak FEN1 cleavage of TNR flaps compared with wild-type pol β during BER, and PCNA did not affect pol β DNA synthesis. Pol β R137Q polymorphic variant exhibited the same ability to skip over a template loop structure, thereby leading to the same TNR deletion as wild-type pol β during BER. Since pol β R137Q variant exhibited the similar activity to that of wild-type pol β in the context of TNRs, our results suggest that human carriers of the pol β polymorphic variant may not exhibit a higher risk than the individuals bearing wild-type pol β in developing TNR expansion diseases.

CHAPTER 2: WEAK DNA SYNTHESIS BY DNA REPAIR POLYMERASES PROMOTES TRINUCLEOTIDE REPEAT DELETIONS DURING BASE EXCISION REPAIR

ABSTRACT

Trinucleotide repeat (TNR) expansion is associated with more than 40 neurodegenerative diseases including Huntington's disease, for which no effective treatment is available. New treatments need to be developed to attenuate the progression of TNR expansion diseases. Previous studies have shown that DNA polymerase β and base excision repair (BER) pathway play a critical role in modulating somatic TNR instability, suggesting that BER may be developed as a new therapeutic target for TNR expansion diseases. A recent study has shown that a pol β variant, pol β Y265C, with impaired DNA synthesis activity can suppress CGG repeats expansion in the fragile X syndrome mice, suggesting that weak DNA synthesis of pol β promotes TNR deletions through BER. However, it remains enigmatic if the DNA synthesis activity of repair DNA polymerases could be utilized as a potential therapeutic target that shortens expanded TNRs. In the present study, we show that the both pol β Y265C and a translesion DNA polymerase (pol ν) with weak DNA synthesis predominantly promoted TNR deletions through BER. The inefficient DNA synthesis of pol β Y265C and pol ν leads to efficient removal of TNRs by FEN1, resulting in TNR deletion. We further demonstrate that TNR deletions can be induced by intentionally inhibiting DNA synthesis activity of wild-type pol β with a pol β inhibitor during BER. Our results indicate that weak DNA synthesis activity of repair DNA polymerases may be developed as a potential therapeutic target for TNR expansion-mediated neurodegenerative disorders.

INTRODUCTION

Huntington's disease (HD), one of the most devastating fatal inherited neurological disorders [242, 243], occurs in human population at a prevalence of 1 in 10,000 [242, 243]. It manifests as gradually losing control of movement and cognition resulted from progressively neuronal breakdowns [244-246]. The average age of the onset is 30-50 years old, yet the symptoms can also occur in patients with age below 20 years old [244-246]. The disease patients have a mean life expectancy of 20 years, and no effective treatment has been developed to cure the disease except for the symptomatic management of the disease for improvement of the quality life of the patients [244-246]. The neuronal death and the limited methods in managing the disease symptoms lead to high mortalities [244-246]. Therefore, it is urgent to develop new treatment that can attenuate and cure the disease.

Huntington's disease is caused by the expansion of CAG repeats (more than 36 CAG repeats) in Huntingtin gene that is located in the Chromosome 4p16.3 [243]. The expanded CAG repeats encode a long stretches of glutamines in the Huntingtin protein, which can aggregate as amyloid-like fibril abundant in β -sheet in the neuron cells, leading to neuronal toxicity, protein degradation, and neuronal cell death [247, 248]. In addition to HD, other TNR expansion neurological disorders, such as Friedreich's ataxia (GAA/TTC), myotonic dystrophy (CTG/GAC), fragile X syndrome (CGG/GCC), etc., are all caused by TNR expansions [88, 89, 92]. Trinucleotide repeat expansions can be induced by the formation of secondary structures including hairpins, loops, and G4 structures during DNA replication, repair, recombination, and gene transcription [7, 14, 92, 199-202]. However, the fact that substantial TNR expansions occur in somatic cells

indicates that DNA repair is actively involved in TNR expansions in somatic cells [92]. Among the different DNA repair pathways, the base excision repair (BER) pathway has been shown to play a critical role in regulating TNR expansions and deletions in somatic cells since TNR tracts contain a long stretch of guanines that form hotspots to oxidative DNA damage. Base excision repair mediates TNR expansions is further supported by the findings showing that oxidative stress and increasing amounts of 8-oxoguanine (8-oxoG) are associated with TNR expansions in the neuronal cells of Huntington's transgenic mice [14, 123, 131], and core BER enzymes, OGG1, APE1, pol β , and FEN1 play critical roles in mediating TNR expansions [14].

During BER in context of TNRs, pol β performs multiple nucleotides gap-filling synthesis, whereas FEN1 removes a short flap, which allows addition of more repeats than those removed by FEN1, causing ligation of a hairpin on the damaged strand and TNR expansions [1, 5, 14]. Furthermore, analysis on the correlations between the levels of BER enzymes and TNR expansions in neuron cells isolated from both striatum and cerebellum of HD mice showed that TNR expansions preferentially occur in striatum, where the neuronal death occurs, and a high ratio of pol β to FEN1 was identified in striatum compared to the ratio in cerebellum [122, 249]. The studies suggest that pol β plays a critical role in BER-mediated TNR expansions. The repeat expansion may be further mediated by the crosstalk between pol β and the mismatch repair protein complex, MSH2-MSH3, which stimulates pol β DNA synthesis and inhibits FEN1 cleavage of TNRs [2]. Thus, pol β and its DNA synthesis activity play a central role in mediating TNR expansions during BER.

On the other hand, pol β is also actively involved in BER-mediated TNR deletions during the repair of DNA base lesions induced by environmental oxidative DNA damage agents, bromate and chromate, and alkylating DNA agent, temozolomide [5, 141]. During BER, pol β can promote TNR deletions by directly bypassing TNR hairpins or loops on the template strand, resulting in fewer TNRs synthesized by pol β than removed by FEN1 [5, 142, 143]. Also, weak DNA synthesis by pol β can prevent repeat expansion, which is supported by the fact that a pol β mutant, pol β Y265C variant, with impaired DNA synthesis activity has been reported to reduce the repeat expansion frequency in both sperm and brain cells from heterozygous pol β Y265C fragile X-syndrome mice [146, 151]. This further suggests that TNR deletions can be induced by weakening or inhibiting DNA synthesis activity of pol β . However, the mechanisms of how the reduced DNA synthesis activity of pol β promotes TNR deletions remains to be elucidated, and whether a weak DNA synthesis activity of pol β can be employed as a potential therapeutic target for TNR expansion diseases needs to be explored. In the current study, we demonstrate that DNA synthesis activity plays a vital role in regulating TNR instability during BER. We find that the weak DNA synthesis of pol β Y265C and a translesion polymerase (pol ν) promote TNR deletions. We also discover that the sequence of TNR modulates the DNA synthesis activities of pol θ , thereby leading to a different outcome of TNR instability. We further demonstrate that by intentionally inhibiting DNA synthesis activity of wild-type pol β with a pol β inhibitor, NSC666719, TNR deletions are preferentially induced. We show that the weak DNA synthesis activity of pol β does not significantly affect the instability of GT dinucleotides and telomere repeats that are associated with cancers and aging-related diseases, suggesting that inhibition of repair

DNA polymerase activity does not confer a significant adverse effect on normal cellular function. Thus, here we provide the first evidence that DNA synthesis activity of polymerases may be utilized as a potential therapeutic target for TNR expansion neurodegenerative diseases through shortening of expanded TNRs.

MATERIALS AND METHODS

Materials and oligonucleotide substrates

DNA oligonucleotide substrates were synthesized by Integrated DNA Technologies (Coralville, IA, USA). Terminal deoxynucleotidyl transferase, T4 polynucleotide kinase, and deoxynucleoside 5'-triphosphates (dNTPs) were purchased from Thermo Fisher Scientific (Waltham, MA). The radionucleotides γ - ^{32}P ATP (6000 mCi/mmol) and Cordycepin 5'-triphosphate 3'- α - ^{32}P (5000 mCi/mmol) were purchased from PerkinElmer (Boston, MA). Micro Bio-Spin 6 Columns were purchased from Bio-Rad Laboratories (Hercules, CA). Pierce Avidin Agarose resin was purchased from Thermo Fisher Scientific (Waltham, MA). All other standard chemical reagents were from Sigma-Aldrich (St. Louis, MO) and Thermo Fisher Scientific (Waltham, MA). Purified enzymes including wild-type (WT) pol β , APE1, FEN1, and LIG I were made according to the procedures of protein purification described previously [51, 145]. Pol β Y265C mutant protein was provided by Dr. Joann Sweasy from Yale University. Translesion polymerases (pol ν and pol θ) were provided by Dr. Yang Wei from National Institute of Diabetes, Digestive and Kidney Diseases/National Institutes of Health. The pol β inhibitor NSC666719 was provided by Dr. Satya Narayan from University of Florida.

Oligonucleotide substrates

The 100 nt oligonucleotide substrates contain (CAG)₂₀, (CTG)₂₀, (GAA)₂₀, and (CGG)₂₀ repeats with a tetrahydrofuran (THF), an analog of a modified abasic site. The THF residue substituted the first G either at the first, fifth, tenth, or fifteenth repeat unit of the (CAG)₂₀, (CTG)₂₀, (GAA)₂₀, (CGG)₂₀, (GT)₃₀, and (TTAGGG)₁₀ containing substrates, which mimics the scenario that the damage occurs either at the 5'-end or in the middle of the repeat tracts. Substrates were constructed by annealing the damaged strand with the template strand at a molecular ratio of 1:2. A strand of DNA fragment containing (CAG)₂₀, (CTG)₂₀, (GAA)₂₀, (CGG)₂₀, (GT)₃₀, and (TTAGGG)₁₀ without damage was used as the size marker. The sequences of substrates are listed in Table 2.

Table 2. Oligonucleotide Sequences

Oligonucleotide	nt	Sequence (5'-3')
(CAG)₂₀ substrate		
CAG-0 (Undamaged strand)	100	CGA GTC ATC TAG CAT CCG TA CAG TA CGT AGA CTT ACT CAT TGC
CAG-1 (THF 1 st repeat)	99	CGA GTC ATC TAG CAT CCG TA CA THF CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG TA CGT AGA CTT ACT CAT TGC
CAG-10 (THF 10 th repeat)	99	CGA GTC ATC TAG CAT CCG TA CAG CAG CAG CAG CAG CAG CAG CAG CAG CA THF CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG TA CGT AGA CTT ACT CAT TGC
CAG-T (Template strand)	100	GCA ATG AGT AAG TCT ACG TA CTG TA CGG ATG

		CTA GAT GAC TCG
(CTG)₂₀ substrate		
CTG-0 (Undamaged strand)	100	CGA GTC ATC TAG CAT CCG TA CTG TA CGT AGA CTT ACT CAT TGC
CTG-1 (THF 1 st repeat)	99	CGA GTC ATC TAG CAT CCG TA CT THF CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG TA CGT AGA CTT ACT CAT TGC
CTG-10 (THF 10 th repeat)	99	CGA GTC ATC TAG CAT CCG TA CTG CTG CTG CTG CTG CTG CTG CTG CTG CT THF CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG TA CGT AGA CTT ACT CAT TGC
CTG-T (Template strand)	100	GCA ATG AGT AAG TCT ACG TAC AG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG TA CGG ATG CTA GAT GAC TCG
(CGG)₂₀ substrate		
CGG-0 (Undamaged strand)	100	CGA GTC ATC TAG CAT CCG TA CGG TA CGT AGA CTT ACT CAT TGC
CGG-10 (THF 10 th repeat)	99	CGA GTC ATC TAG CAT CCG TA CGG CGG CGG CGG CGG CGG CGG CGG CGG CG THF CGG CGG CGG CGG CGG CGG CGG CGG CGG CGG TA CGT AGA CTT ACT CAT TGC
CGG-T (Template strand)	100	GCA ATG AGT AAG TCT ACG TA CCG TA CGG ATG CTA GAT GAC TCG
(GAA)₂₀ substrate		
GAA-0 (Undamaged strand)	100	CGA GTC ATC TAG CAT CCG TA GAA TA CGT AGA

GAA-10 (THF 10 th repeat)	99	CTT ACT CAT TGC CGA GTC ATC TAG CAT CCG TA GAA GAA GAA GAA GAA GAA GAA GAA GAA THF AA GAA GAA GAA GAA GAA GAA GAA GAA GAA GAA TA CGT AGA CTT ACT CAT TGC
GAA-T (Template strand)	100	GCA ATG AGT AAG TCT ACG TA TTC TA CGG ATG CTA GAT GAC TCG
<hr/> (GT)₃₀ substrate		
GT-0 (Undamaged strand)	100	CGA GTC ATC TAG CAT CCG TA GT TA CGT AGA CTT ACT CAT TGC
GT-15 (THF 15 th repeat)	99	CGA GTC ATC TAG CAT CCG TA GT GT GT GT GT GT GT GT GT GT GT GT GT THF T GT GT GT GT GT GT GT GT GT GT GT GT GT GT TA CGT AGA CTT ACT CAT TGC
GT-T (Template strand)	100	GCA ATG AGT AAG TCT ACG TA AC TA CGG ATG CTA GAT GAC TCG
<hr/> (TTAGGG)₁₀ substrate		
TTAGGG-0 (Undamaged strand)	100	CGA GTC ATC TAG CAT CCG TA TTAGGG TTAGGG TTAGGG TTAGGG TTAGGG TTAGGG TTAGGG TTAGGG TTAGGG TTAGGG TA CGT AGA CTT ACT CAT TGC
TTAGGG-5 (THF 5 th repeat)	99	CGA GTC ATC TAG CAT CCG TA TTAGGG TTAGGG TTAGGG TTAGGG TTAGG THF TTAGGG TTAGGG TTAGGG TTAGGG TTAGGG TA CGT AGA CTT ACT CAT TGC
TTAGGG-T (Template strand)	100	GCA ATG AGT AAG TCT ACG TA CCCTAA CCCTAA CCCTAA CCCTAA CCCTAA CCCTAA CCCTAA CCCTAA CCCTAA CCCTAA TA CGG ATG CTA GAT GAC TCG

Universal PCR primers for DNA fragment analysis				
Forward primer		20	CGA GTC ATC TAG CAT CCG TA	
Reverse primer	(6'-FAM labeled)	19	6-FAM-CA ATG AGT AAG TCT ACG TA	

***In vitro* polymerase DNA synthesis assay and FEN1 cleavage assay**

The DNA synthesis activity was determined by incubating various concentrations of WT pol β , pol β Y265C, pol ν , and pol θ separately in the presence of APE1 with 25 nM (CAG)₂₀, (CTG)₂₀, (GAA)₂₀, (CGG)₂₀, (GT)₃₀, or (TTAGGG)₁₀ substrates that contain a THF residue in reaction buffer containing 50 mM Tris-HCl, pH 7.5, 50 mM KCl, 0.1 mM EDTA, 0.1 mg/ml bovine serum albumin (BSA), and 0.01% Nonidet P-40 (NP-40) along with 50 μ M dNTPs and 5 mM Mg²⁺ at 37°C for 15 min. FEN1 cleavage activity on the substrates was measured in the absence and presence of wild-type pol β , pol β Y265C variant, pol ν , or pol θ in the reaction buffer containing 50 mM Tris-HCl, pH 7.5, 50 mM KCl, 0.1 mM EDTA, 0.1 mg/ml BSA, NP-40, and 5 mM Mg²⁺ at 37°C for 15 min. The 20 μ l-reaction mixture was assembled on ice, and reactions were quenched by addition of 20 μ l 2X stopping buffer containing 95% formamide and 10 mM EDTA and incubation at 95°C for 10 min. Substrates and products were separated in a 15% urea-denaturing polyacrylamide gel (PAGE) and detected by a Pharos FX Plus PhosphorImager (Bio-Rad Laboratory, CA).

Reconstitution of BER

Reconstituted BER was performed by incubating purified APE1, one of the DNA polymerases (WT pol β , pol β R137Q variant, pol ν , or pol θ), FEN1, and LIG I along

with (CAG)₂₀, (CTG)₂₀, (GAA)₂₀, or (CGG)₂₀ substrates (25 nM) with a THF residue in a reaction mixture (20 µl) that contained 5 mM MgCl₂, 50 µM dNTPs, 2 mM ATP, 50 mM Tris-HCl (pH 7.5), 50 mM KCl, 0.1 mg/ml BSA, 0.1 mM EDTA, and 0.01% NP-40. Reaction mixtures were assembled on ice and incubated at 37°C for 15 min. Reactions were terminated by addition of 20-µl 2X stopping buffer with incubation at 95°C for 10 min. Substrates and products were then separated in 15% urea-denaturing PAGE and were detected by the PhosphorImager.

Polymerase inhibition assay

The WT pol β was pre-incubated with pol β inhibitor NSC666719 in the buffer condition of 50 mM Tris-HCl (pH 7.5), 50 mM KCl, 0.1 mg/ml BSA, 0.1 mM EDTA, and 0.01% NP-40 at 37°C for 30 min with rotation. The (CAG)₂₀-containing substrate (25 nM) with an abasic lesion in the middle of the repeat tract was pre-incubated with APE1 at 37°C for 15 min. The pre-cut substrate was then added to the reaction mixture for additional 15 min at 37°C with the presence of 5 mM MgCl₂ and 50 µM dNTPs.

Determination of repeat lengths of BER products by DNA fragment analysis

To isolate a repaired strand, the template strands of all TNR-containing substrates were tagged by a biotin residue at the 5'-end. Reconstituted BER reactions were terminated with 1 µl of 100 mM EDTA, and reaction mixtures were incubated with 50-µl avidin agarose beads (Pierce-Thermo Scientific, Rockford, IL) for 2 hrs with rotation, allowing the binding of avidin beads to the biotin on the template strand. Separation of repaired strands from the template strands was then performed by incubating the reaction

mixtures with 0.15 M NaOH at room temperature for 15 min with rotation. This was followed by 2-min centrifugation at 5000 rpm pelleting the template strands bound by avidin beads. After clearing the template strands, the separated repaired strands in the supernatant were precipitated with ethanol, and subsequently dissolved in TE buffer for PCR amplification and size analysis. Repaired products were amplified through PCR with the AmpliTaq Gold 360 DNA polymerase Kit (Applied Biosystems, Foster City, CA) at the following condition: denaturation at 95°C for 30s, annealing at 52°C for 30s, extension at 72°C for 90s for 35 cycles, and a final extension at 72°C for 1 hr. Both a forward primer and a reverse primer were used for amplification of the repaired strands. All the reverse primers were tagged with 6-carboxyfluorescein (6-FAM). The primer sequences of different repeat substrates are listed in Table 2. PCR products mixed with the size marker MapMarker 500 (Bioventures, Murfreesboro, TN) were then subject to capillary electrophoresis via an ABI 3130XL Genetic Analyzer (Applied Biosystems, Foster City, CA) at Florida International University DNA sequencing core facility. The sizes of PCR products were determined by DNA fragment analysis with the Peak Scanner version 1.0 software (Applied Biosystems, Foster City, CA).

RESULTS

The weak DNA synthesis of pol β Y265C promotes CAG repeat deletions independent of damage location

A polymorphic pol β R137Q variant with slightly weaker DNA synthesis activity than WT pol β does not significantly promote repeat deletions compared to the wild-type pol β [13]. However, pol β Y265C, which has much weaker DNA synthesis activity than

WT pol β , significantly reduces TNR expansion frequency in fragile X syndrome mice [146, 151]. In addition, since the locations of DNA damage also plays a role in governing pol β -mediated TNR expansions or deletions [5], we initially compared the DNA synthesis activity of pol β Y265C with that of WT pol β with (CAG)₂₀-containing substrates with a THF located at the 5'-side or in the middle of the repeat tract. We found that WT pol β incorporated up to 2 repeats, 3 repeats, and 6 repeats at the concentrations of 1 nM, 2 nM, and 5 nM when the damage located at the 5'-side (Figure 2.1A, lanes 3-5). While at the same concentrations, pol β Y265C incorporated up to 1 nt, 1 repeat, and 2 repeats, respectively (Figure 2.1A, lanes 4-6). Similarly, when the damage located in the middle of the repeat tract, WT pol β inserted up to 1 repeat, 2 repeats, and 5 repeats (Figure 2.1B, lanes 3-5) compared to 1 nt, 2 nt, and 4 nt synthesized by pol β Y265C at the concentrations of 1 nM, 2 nM, and 5 nM, respectively (Figure 2.1B, lanes 4-6). The difference in the sizes of the synthesized products by WT pol β and pol β Y265C indicates that pol β Y265C synthesized fewer TNRs than WT pol β during BER.

Figure 2.1. Pol βY265C exhibits weak DNA synthesis activity during BER of a base lesion in the context of (CAG)₂₀ repeats.

DNA synthesis activity of WT pol β or pol βY265C at concentrations of 1 nM, 2 nM, and 5 nM was measured with (CAG)₂₀ containing substrates with a THF residue located at the first repeat (**A**). Lanes 1 represents substrate only. Lane 2 represents APE1 cleavage products. Lanes 3-5 represent WT pol β synthesized products. Lanes 6-8 represent pol βY265C synthesized products. Lanes 9-14 represent BER reconstitution with additional 5 nM FEN1 and 10 nM LIG I. Lanes 9-11 represent BER reconstitution in the presence of WT pol β. Lanes 12-14 represent BER reconstitution in the presence of pol βY265C. Substrates were ³²P-labeled at the 5'-end of the damaged strand.

Figure 2.1. Pol βY265C exhibits weak DNA synthesis activity during BER of a base lesion in the context of (CAG)₂₀ repeats.

DNA synthesis activity of WT pol β or pol βY265C at concentrations of 1 nM, 2 nM, and 5 nM was measured with (CAG)₂₀ containing substrates with a THF residue located at the tenth repeat (**B**). Lanes 1 represents substrate only. Lane 2 represents APE1 cleavage products. Lanes 3-5 represent WT pol β synthesized products. Lanes 6-8 represent pol βY265C synthesized products. Lanes 9-14 represent BER reconstitution with additional 5 nM FEN1 and 10 nM LIG I. Lanes 9-11 represent BER reconstitution in the presence of WT pol β. Lanes 12-14 represent BER reconstitution in the presence of pol βY265C. Substrates were ³²P-labeled at the 5'-end of the damaged strand.

To further test whether the weak DNA synthesis activity of pol β could promote repeat deletions during BER, we reconstituted the BER in the presence of pol βY265C with the (CAG)₂₀-containing substrates with a THF located at the 5'-end or in the middle of the repeat tract. We observed the repaired products in the presence of both WT pol β and pol βY265C when the lesion located either at the 5'-side or in the middle of the repeat tract (Figure 2.1A and 2.1B, lanes 9-14). We also observed the repeat deletion products in the presence of all concentrations of pol βY265C (Figure 2.1A and 2.1B,

lanes 12-14). Size analysis of the repaired products showed that pol β Y265C induced 1 repeat deletion products at the concentration of 1 nM and 2 nM, and it did not induce repeat deletions at 5 nM with the damage located at 5'-side of the repeat tract (Figure 2.2A, panels d, f, and h). Whereas, WT pol β at the same concentrations failed to induce repeat deletions (Figure 2.2A, panels c, e, and g). For the substrate with an abasic lesion in the middle of the repeat tract, pol β Y265C induced 1 repeat deletion products through BER at all three concentrations (1 nM, 2 nM, and 5 nM) compared to no repeat deletion products induced by WT pol β at the same concentrations (Figure 2.2B, compare panels d, f, and h to panels c, e, and g). The results indicated that the weak DNA synthesis activity of pol β Y265C predominantly promoted CAG repeat deletions independent of the locations of an abasic lesion. In addition, for both the damage located at the 5'-side or in the middle of the repeat tract, the amount of the repeat deletion products was decreased with increasing the concentration of pol β Y265C mutant protein, indicating that the DNA synthesis activity of pol β Y265C was increased at the a high concentration of the mutant protein.

Figure 2.2. Pol βY265 promotes (CAG)₂₀ repeat deletions.

BER was reconstituted with 5 nM APE1, 10 nM FEN1, 5 nM LIG I, and increasing concentrations (1 nM, 2 nM, and 5 nM) of WT pol β or pol βY265C along with (CAG)₂₀ substrates. The repaired products were subjected to PCR amplification and capillary electrophoresis and the repeat lengths were analyzed by DNA fragment analysis. (A) Length analysis of the repaired products from the substrate with a lesion located at the first repeat of the (CAG)₂₀ substrate. Panels a and b represent a (CTG)₂₀ marker without damage. Panels c, e, and g indicate the sizes of the repaired products in the presence of 1 nM, 2 nM, and 5 nM WT pol β. Panels b, f, and h represent the sizes of the repaired products in the presence of 1 nM, 2 nM, and 5 nM pol βY265C. Note: Panels a and b used the same data since in the experiment, all samples were run along with the same markers for comparison.

Figure 2.2. Pol βY265 promotes (CAG)₂₀ repeat deletions during BER.

BER was reconstituted with 5 nM APE1, 10 nM FEN1, 5 nM LIG I, and increasing concentrations (1 nM, 2 nM, and 5 nM) of WT pol β or pol βY265C along with (CAG)₂₀ substrates. The repaired products were subjected to PCR and capillary electrophoresis and the repeat lengths were analyzed by DNA fragment analysis. **(B)** Size analysis of the repaired products of from the substrate with a lesion located at the tenth repeat of the (CAG)₂₀ substrate. Panels a and b represent a (CTG)₂₀ marker without damage. Panels c, e, and g indicate the sizes of the repaired products in the presence of 1 nM, 2 nM, and 5 nM WT pol β. Panels b, f, and h represent the sizes of the repaired products in the presence of 1 nM, 2 nM, and 5 nM pol βY265C. Note: Panels a and b used the same data since in the experiment, all samples were run along with the same markers for comparison.

To further explore the mechanisms by which the weak DNA synthesis activity had induced repeat deletions, we then determined the FEN1 cleavage activity in the presence of pol β Y265C. When the damage located at the 5'-side, FEN1 alone removed up to 1 repeat (Figure 2.3A, lane 3). Flap endonuclease 1 removed mainly 4 nt in the presence of 1 nM pol β Y265C (Figure 2.3A, lane 7), and it removed up to 2 repeats when the pol β Y265C increased to 2 nM and 5 nM (Figure 2.3A, lanes 8-9). In the presence of WT pol β , FEN1 exhibited efficient cleavage of repeats, it removes up to 2 repeats, 3 repeats, and 4 repeats in corresponding to 1 nM, 2 nM, and 5 nM of WT pol β , respectively (Figure 2.3A, lanes 4-6). Similarly, when the damage was located in the middle of the repeat tract, FEN1 alone mainly removed 1 repeat (Figure 2.3B, lane 12). Flap endonuclease 1 removed up to 2 repeats in the presence of all three concentrations (1 nM, 2 nM, and 5 nM) of pol β Y265C (Figure 2.3B, lanes 16-18), whereas it removed up to 2 repeats, 3 repeats, and 5 repeats with the increasing concentrations (1 nM, 2 nM, and 5 nM) of WT pol β (Figure 2.3B, lanes 13-15). The results indicate that the presence of WT pol β stimulated FEN1 cleavage. However, the pol β Y265C did not significantly affect FEN1 cleavage activity. Comparing the major DNA synthesis products by pol β Y265C with the ones cleaved by FEN1 in the presence of the variant protein, the repeat flap removed by FEN1 is 1-repeat longer than the repeats synthesized by pol β variant (compare Figure 2.3A, lanes 7-9 and 2.3B, lanes 16-18 to Figure 2.1A and 2.1B, lanes 6-8 respectively), indicating that the weak DNA synthesis of pol β Y265C resulted in the loss of CAG repeats, leading to repeat deletions.

Figure 2.3. Pol βY265C leads to weak FEN1 cleavage during BER in the context of (CAG)₂₀ repeats.

FEN1 flap cleavage activity at 5 nM on the (CAG)₂₀ containing substrate with a THF located at the first repeat (**A**) or tenth repeat (**B**) was measured in the presence of 1 nM, 2 nM, and 5 nM of pol β. Lanes 1 and 10 represent the substrate only. Lanes 2 and 11 represent reactions with APE1 alone. Lanes 3 and lanes 12 represent reactions with FEN1 alone. Lanes 4-6 and 13-15 represent reactions with FEN1 in the presence of 1 nM, 2 nM, and 5 nM of WT pol β accordingly. Lanes 7-9 and 16-18 illustrate reactions with FEN1 in the presence of 1 nM, 2 nM, and 5 nM of pol βY265C. Substrates were ³²P-labeled at the 3'-end of the damaged strand.

Weak DNA synthesis activity of repair DNA polymerases promotes repeat deletions

There are 17 human DNA polymerases [161, 250, 251] with different DNA synthesis activities that may all participate in DNA synthesis on TNR tracts. For example, translation polymerase pol θ has deoxyribose phosphate lyase activity, indicating its role in BER since it may be able to remove a sugar residue efficiently [252, 253]. Another translesion polymerase, pol v, may also be involved in modulating TNR instability during BER. Since pol v is specialized in bypassing the thymine glycol lesion [251], it is

possible that it can bypass the thymine glycols formed in TNR tracts, such as (CAG)_n/(CTG)_n and (GAA)_n/(CTT)_n, which are rich of thymine. Moreover, translesion DNA polymerases adopt unique structures to interact with DNA, allowing them to accommodate DNA lesions on the template strand. The interaction between the translesion DNA polymerases and the TNR may result in different outcomes of TNR instability. Both pol θ and pol ν can cause a primer loop-out during DNA synthesis [253, 254], which may occur in TNR tracts and forms a basis for TNR expansion. To examine the roles of translesion DNA polymerases in modulating TNR instability during BER, we initially compared the DNA synthesis activity of pol ν and pol θ with pol β during BER of an abasic lesion in the middle of a (CAG)₂₀ repeat tract. We found that Pol ν exhibited very weak DNA synthesis activity and mainly inserted only 1 nt, 1 nt, 1 repeat at the concentrations of 1 nM, 2 nM, and 5 nM (Figure 2.4A, lanes 6-8), respectively, compared to 1 repeat, 2 repeats, and 6 repeats inserted mainly by pol β at the corresponding concentrations (Figure 2.4A, lanes 3-5). The weak DNA synthesis activity of the translesion DNA polymerase also resulted in the products with one repeat deletion (Figure 2.4A, lanes 12-14 and Figure 2.4B, panels b, c, and d). The repeat deletion products decreased when the concentration of pol ν increased (Figure 2.4B, compare the peak heights among panels b, c, and d), indicating that the increased amount of pol ν facilitated the synthesis on repeat tract and thus prevented repeat deletion. On the other hand, FEN1 cleaved up to 2 repeats in the presence of 1 nM, 2 nM of pol ν , and 3 repeats in the presence of 5 nM of pol ν (Figure 2.4C, lanes 7-9), compared to 1 repeat removed by FEN1 alone (Figure 2.4C, lane 3). The results suggest that weak DNA synthesis of pol ν caused more repeats removed by FEN1, which led to TNR deletions.

Figure 2.4. Weak DNA synthesis of pol v promotes (CAG)₂₀ repeat deletions.
(A) Measurement of DNA synthesis activity and BER reconstitution were conducted with the (CAG)₂₀-containing substrate with a THF residue located at tenth repeat. Lane 1 represents substrate only. Lane 2 represents APE1 cleavage products. Lanes 3-5 represent the DNA synthesis products by pol β at the concentration of 1 nM, 2 nM, and 5 nM. Lanes 6-8 represent the DNA synthesis products of pol v at the concentration of 1 nM, 2 nM, and 5 nM. Lanes 9-14 represent BER reconstituted with additional 5 nM FEN1 and 10 nM LIG I. Lanes 9-11 represent BER reconstituted in the presence of pol β (1 nM, 2 nM, and 5 nM). Lanes 12-14 represent BER reconstituted in the presence of pol βY265C (1 nM, 2 nM, and 5 nM).

Figure 2.4. Weak DNA synthesis of pol v promotes (CAG)₂₀ repeat deletions. (B) Size analysis of the repaired products in the presence of 1 nM, 2 nM, and 5 nM of pol v. Panel a represents the length of the (CAG)₂₀ marker. Panels b, c, and d represent the size change of the repaired products in the presence of 1 nM, 2 nM, and 5 nM of pol v, respectively. **(C)** FEN1 (5 nM) cleavage activities in the presence of 1 nM, 2 nM, and 5 nM of pol β or pol v. Lane 1 represents substrate only. Lane 2 represents APE1 cleavage products. Lane 3 represents the reaction with FEN1 only. Lanes 4-6 represent reactions with FEN1 in the presence of 1 nM, 2 nM, and 5 nM of WT pol β. Lanes 7-9 illustrate reactions with FEN1 in the presence of 1 nM, 2 nM, and 5 nM of pol v. Substrates were labeled with ³²P-labeled either at the 5'- end or at the 3'-end of the damaged strand as illustrated in the figure.

The pol θ at the concentration of 2 nM, 5 nM, and 10 nM exhibited stronger DNA synthesis activity than pol β. It inserted 4 nt, 8 repeats, and 50 nt (Figure 2.5A, lanes 6-8) compared to 2 repeats, 5 repeats, and 10 repeats synthesized by pol β at the corresponding concentrations (Figure 2.5A, lanes 3-5). The strong DNA synthesis activity resulted in the maintenance of TNR stability (Figure 2.5B, compare panels b, c, and d to panel a). The results suggest that the relatively stronger DNA synthesis of pol θ prevented repeat deletion. Thus, it appears that only weak DNA synthesis activity of repair DNA polymerase facilitated TNR deletions during BER.

Figure 2.5. Pol θ does not induce (CAG)₂₀ repeat deletions.

(A) DNA synthesis activity of pol θ at 2 nM, 5 nM, and 10 nM was measured with the (CAG)₂₀-containing substrate with a THF residue located at tenth repeat. Lane 1 represents substrate only. Lane 2 represents APE1 cleavage products. Lanes 3-5 represent the DNA synthesis products of pol β at 2 nM, 5 nM, and 10 nM. Lanes 6-8 represent the DNA synthesized products of pol θ at the concentration of 2 nM, 5 nM, and 10 nM. (B) Sizing analysis of the repaired products in the presence of 2 nM, 5 nM, and 10 nM of pol θ. Panel a represents the length of the (CAG)₂₀ marker. Panel b, c, and d represent the sizes of the repaired products in the presence of 2 nM, 5 nM, and 10 nM of pol θ.

DNA sequences alter TNR instability by modulating pol β DNA synthesis activity

Previous studies indicated that different repeat sequences can facilitate the formation of different types of secondary structures [7]. For example, CAG, CTG, and CGG repeats may facilitate the formation of hairpin structures, and the same size of hairpin structures formed by different types of TNRs can have different thermostability and exhibit different dynamics [7]. Hairpins with different thermostability can result in different outcomes of repeat instability. Moreover, DNA polymerases may have different sequence preference that may result in the different efficiency of DNA synthesis on

different TNR tracts. To further identify whether the weak DNA synthesis activity of pol β Y265C and pol ν can also promote repeat deletions in the context of different repeated sequences, we examined the DNA synthesis activity of pol β , pol β Y265C, and pol ν in modulating the instability of (CTG)₂₀ and (CGG)₂₀ tracts.

We initially tested the DNA synthesis activity of pol β Y265C and its roles in modulating CTG repeat instability during BER of an abasic lesion located at the 5'-side or in the middle of a (CTG)₂₀ repeat tract. Pol β Y265C at 1 nM, 2 nM, 5 nM, and 10 nM performed weak DNA synthesis on the (CTG)₂₀ repeat tracts. For a lesion located at 5'-side of the repeat tract, pol β Y265C incorporated up to 1 nt, 1 repeat, 2 repeats, and 4 repeats (Figure 2.6A, lanes 7-10), whereas the same concentrations of WT pol β incorporated 1 repeat, 2 repeats, 5 repeats, and 8 repeats (Figure 2.6A, lanes 3-6). For a lesion located in the middle of the repeat tract, pol β Y265C at the concentrations of 2 nM, 5 nM, and 10 nM inserted 1 nt, 3 nt, and 7 nt (Figure 2.6B, lanes 6-8), whereas the same concentrations of WT pol β inserted 2 repeats, 4 repeats, and 10 repeats (Figure 2.6B, lanes 3-5). The weak DNA synthesis of pol β Y265C led to repeat deletion on the (CTG)₂₀ repeat tract. Both WT pol β and pol β Y265C at the concentration of 1 nM induced 1 repeat deletion (Figure 2.6C and 2.5D, panels c and d). However, pol β Y265C-mediated BER resulted in more repeat deletion products than WT pol β (compare panels d to panels c in Figure 2.6C and 2.6D). At higher concentrations of 2 nM and 5 nM, only pol β Y265C promoted repeat deletions (compare panels f to e, and panels h to g in Figure 2.6C and 2.6D, respectively).

Figure 2.6. Weak DNA synthesis of pol βY265C promotes CTG repeat deletion during BER in the context of (CTG)₂₀ repeats.

(A) DNA synthesis activity and BER reconstitution was measured with (CTG)₂₀ containing substrates with a THF residue located at first repeat. Lane 1 represents substrate only. Lane 2 represents APE1 cleavage products. Lanes 3-6 represent the DNA synthesized products of WT pol β (1 nM, 2 nM, 5 nM, and 10 nM). Lanes 7-10 represent the DNA synthesized products of pol βY265C (1 nM, 2 nM, 5 nM, and 10 nM). Lanes 11-14 represent BER reconstituted in the presence of WT pol β at 1 nM, 2 nM, 5 nM, and 10 nM. Lanes 15-18 represent BER reconstituted in the presence of pol βY265C at 1 nM, 2 nM, 5 nM, and 10 nM).

Figure 2.6. Weak DNA synthesis of pol βY265C promotes CTG repeat deletion during BER in the context of (CTG)₂₀ repeats.

(B) Measurement of DNA synthesis activity and BER reconstitution was conducted with the (CTG)₂₀-containing substrates with a THF residue located at the tenth repeat. Lane 1 represents substrate only. Lane 2 represents APE1 cleavage products. Lanes 3-5 represent the DNA synthesis products of WT pol β at 2 nM, 5 nM, and 10 nM. Lanes 6-8 represent the DNA synthesis products of pol βY265C at 2 nM, 5 nM, and 10 nM. Lanes 9-11 represent BER reconstituted in the presence of WT pol β at 2 nM, 5 nM, and 10 nM. Lanes 12-14 represent BER reconstituted in the presence of pol βY265C at 2 nM, 5 nM, and 10 nM.

Figure 2.6. Weak DNA synthesis of pol βY265C promotes CTG repeat deletion during BER in the context of (CTG)₂₀ repeats.

(C) Sizing analysis of the repaired products in the presence of 1 nM, 2 nM, and 5 nM of pol βY265C with the (CTG)₂₀-containing substrates with a THF residue located at the first repeat. Panels a and b represent the length of the (CTG)₂₀ marker. Panels c, e, and g indicate the sizes of the repaired products in the presence of 1 nM, 2 nM, and 5 nM WT pol β. Panels b, f, and h represent the size analysis of the repaired products in the presence of 1 nM, 2 nM, and 5 nM pol βY265C. All samples in panels a and b were run along with the same size markers.

Figure 2.6. Weak DNA synthesis of pol βY265C promotes CTG repeat deletion during BER in the context of (CTG)₂₀ repeats.

(D) Sizing analysis of the repaired products in the presence of 1 nM, 2 nM, and 5 nM of pol βY265C with the (CTG)₂₀ containing substrates with a THF residue located at the tenth repeat. Panels a and b represent the length of the (CTG)₂₀ marker. Panels c, e, and g indicate the sizes of the repaired products in the presence of 1 nM, 2 nM, and 5 nM WT pol β. Panels b, f, and h represent the sizes of the repaired products in the presence of 1 nM, 2 nM, and 5 nM pol βY265C. All samples in panels a and b were run along with the same size markers.

On the other hand, pol v performed weaker DNA synthesis activity on CGG repeats than it did on CAG repeats. It mainly inserted 1 nt at the concentrations of 2 nM and 5 nM. It inserted up to 1 repeat when the concentrations increased to 10 nM (Figure 2.7A, lanes 6-8), while WT pol β inserted up to 1 nt, 3 repeats, and 6 repeats at the respective concentrations (Figure 2.7A, lanes 3-5). The weak DNA synthesis of pol v also resulted in the products with 1 repeat deletion (Figure 2.7B, panels d, f, h), whereas WT pol β did not promote the formation of deletion products (Figure 2.7B, panels c, e, g). Analysis on the FEN1 cleavage with these substrates showed that FEN1 cleaved up to 2 repeats in the presence of 2 nM, 5 nM, and 10 nM pol v (Figure 2.7C, lanes 7-9). The

results indicate that weak DNA synthesis of pol v promoted repeat deletions independent of the sequences of TNRs.

Figure 2.7. Weak DNA synthesis of pol v promotes (CGG)₂₀ repeat deletions.

(A) Measurements of DNA synthesis activity and BER reconstitution were conducted with (CGG)₂₀-containing substrates with a THF residue located at the tenth repeat. Lane 1 represents substrate only. Lane 2 represents APE1 cleavage products. Lanes 3-5 represent the DNA synthesis products from WT pol β at 2 nM, 5 nM, and 10 nM. Lanes 6-8 represent the DNA synthesis products from pol v at 2 nM, 5 nM, and 10 nM. Lanes 9-11 represent BER reconstitution in the presence of WT pol β. Lanes 12-14 represent BER reconstitution in the presence of pol v. Substrates were ³²P-labeled at the 5'-end of the damaged strand.

Figure 2.7. Weak DNA synthesis of pol v promotes (CGG)₂₀ repeat deletions.

(B) Sizing analysis of the repaired products from pol v with the (CGG)₂₀-containing substrate with a THF residue located at the tenth repeat. Panels a and b represent the length of the (CGG)₂₀ marker. Panels c, e, and g indicate the sizes of the repaired products in the presence of 2 nM, 5 nM, and 10 nM WT pol β. Panels b, f, and h represent the sizes of the repaired products in the presence of 2 nM, 5 nM, and 10 nM pol v.

Figure 2.7. Weak DNA synthesis of pol ν promotes (CGG)₂₀ repeat deletions.

(C) FEN1 (5 nM) cleavage activities in the presence of 2 nM, 5 nM, and 10 nM of WT pol β or pol ν . Lane 1 represents substrate only. Lane 2 represents APE1 cleavage products. Lane 3 represents reaction with FEN1 only. Lanes 4-6 represent reactions with FEN1 in the presence of 2 nM, 5 nM, and 10 nM of WT pol β . Lanes 7-9 illustrate the reactions with FEN1 in the presence of 2 nM, 5 nM, and 10 nM pol ν . Substrates were labeled with ³²P-labeled at the 3'-end of the damaged strand.

In addition, we further compared the DNA synthesis activity of pol θ on GAA repeats to WT pol β at 1 nM, 2 nM, and 5 nM and found that pol θ performed weaker DNA synthesis activity than pol β on GAA repeats. Pol θ mainly inserted 1 nt at 1 nM, 2 nM, and 5 nM (Figure 2.8A, lanes 6-8), whereas the same concentrations of pol β inserted 1 nt, 1 repeat, and 2 repeats (Figure 2.8A, lanes 3-5). Analysis on the sizes of the repaired products showed that pol β resulted in a small amount products with 1-2 GAA repeat deletions (Figure 2.8B, panels c, e, and g), whereas pol θ resulted in a large amount of 1 repeat deletion products (Figure 2.8B, panels d, f, and h). The results indicate that the DNA sequences of TNRs affected DNA synthesis activity and thus modulated TNR instability.

Figure 2.8. Pol θ induces (GAA)₂₀ repeat deletions.

(A) DNA synthesis activity of 1 nM, 2 nM, and 5 nM of pol θ was measured with the (GAA)₂₀-containing substrate with a THF residue located at tenth repeat. Lane 1 represents substrate only. Lane 2 represents APE1 cleavage products. Lanes 3-5 represent the DNA synthesis products from pol β at 1 nM, 2 nM, and 5 nM. Lanes 6-8 represent the DNA synthesis products from pol θ at 1 nM, 2 nM, and 5 nM.

Figure 2.8. Pol θ induces (GAA)₂₀ repeat deletions.

(B) Sizing analysis of the repaired products in the presence of 1 nM, 2 nM, and 5 nM of pol θ. Panels a and b represent the length of the (GAA)₂₀ marker. Panels c, e, and g indicate the sizes of the repaired products from BER in the presence of 1 nM, 2 nM, and 5 nM pol β. Panels b, f, and h represent the sizes of the repaired products in the presence of 1 nM, 2 nM, and 5 nM pol θ. Panels a and b are the same marker which was run along the samples.

Inhibition of pol β significantly promotes TNR repeat deletions

To further determine whether the DNA synthesis activity of DNA polymerases can be utilized as a potential therapeutic target that shortens the expanded TNRs, we then used a pol β inhibitor NSC666719, which inhibits the strand displacement DNA synthesis activity of pol β [255], to modulate TNR instability through BER. We reason that inhibition of the strand displacement DNA synthesis of pol β during BER in TNR tracts may reduce the formation of secondary structures such as TNR hairpins and loops on the downstream strand. This is because pol β strand displacement DNA synthesis facilitates separation of the downstream strand from the template strand. Formation of TNR hairpins and loops on the downstream damage strand inhibits FEN1 5'-flap endonucleolytic activity, resulting in FEN1 alternative cleavage of a shorter flap and causing repeat expansion [1, 5]. Thus, reduction of DNA strand displacement synthesis of pol β by inhibiting pol β polymerase activity may promote TNR deletions.

We found that the pol β inhibitor significantly reduced the DNA synthesis activity of pol β on the (CAG)₂₀ repeat tract with a lesion in the middle of the repeat tract. Pol β at 1 nM mainly inserted 1-2 repeats (Figure 2.9, lane 2). However, pol β mainly inserted 1 repeat with the presence of 1 μ M, 3 μ M, and 5 μ M of the inhibitor (Figure 2.9, lanes 4-6). When the concentrations of the inhibitor increased to 8 μ M, 10 μ M, and 25 μ M, the DNA synthesized products of pol β was reduced to 2 nt, 1 nt, and 0, respectively (Figure 2.9, lanes 7-9). On the other hand, 10 nM pol β synthesized up to 8 repeats in the absence of the inhibitor (Figure 2.9, lane 3). The sizes of pol β synthesis products of was reduced to 7 repeats, 6 repeats, and 4 repeats, respectively, in the presence of 1 μ M, 3 μ M, and 5 μ M of the inhibitor (Figure 2.9, lanes 10-12). The DNA synthesis activity of pol β on the

repeat tract was further reduced when the concentrations of the inhibitor was increased to 8 μ M, 10 μ M, and 25 μ M as pol β only incorporated 1 repeat, 1 repeat, and 1 nt at the concentrations of the inhibitor (Figure 2.9, lanes 13-15).

Figure 2.9. DNA synthesis activity of pol β is inhibited by a pol β inhibitor, NSC666719.

The inhibition of DNA synthesis activity of pol β at 1 nM and 10 nM by NSC666719 was measured by incubating with increasing concentrations of NSC666719 (1 μ M, 3 μ M, 5 μ M, 8 μ M, 10 μ M, and 25 μ M) with the (CAG)₂₀-containing substrate with a THF residue located at the tenth repeat. Lane 1 represents APE1 pre-cut substrate. Lane 2 and lane 3 represent DNA synthesis products 1 nM and 10 nM of pol β , respectively. Lanes 4-9 represent the DNA synthesis products from 1 nM pol β in the presence of increasing concentrations of NSC666719. Lanes 10-15 represent the DNA synthesis products of 10 nM pol β in the presence of increasing concentrations of NSC666719.

To determine whether the inhibition of the DNA synthesis activity of pol β promotes TNR deletions, we measured the sizes of the repeat tracts in the repaired products. In the absence of the inhibitor, neither 1 nM nor 10 nM WT pol β promoted repeat deletions (Figure 2.10A, lanes 3 and 6; Figure 2.10B and 2.10C, panels b). The inhibitor at 8 μ M promoted repeat deletions in the presence of both 1 nM and 5 nM pol β (Figure 2.10A, lanes 5 and 8; Figure 2.10B, panel d; and Figure 2.10C, panel c), and 5

μM inhibitor also promoted repeat deletions in the presence of 1 nM WT pol β (Figure 2.10A, lane 4; and Figure 2.10B, panel c). Interestingly, with 8 μM inhibitor, it completely converted all full size repaired products into the products with 1 repeat deletion (Figure 2.10B, panel d). The results suggested that inhibition of the DNA synthesis activity of pol β significantly promoted TNR deletions through BER.

Figure 2.10. Inhibition of DNA synthesis activity of pol β promotes (CAG)₂₀ repeat deletions

(A) BER reconstituted in the presence of NSC666719 with the (CAG)₂₀-containing substrates with a THF residue located at the tenth repeat. Lane 1 represents substrate only. Lane 2 represents APE1 cleavage products. Lanes 3 and 6 represent the repair products in the presence of 1 nM and 10 nM pol β , respectively. Lanes 4-5 and lanes 7-8 represent the repair products in the presence of 1 nM pol β and 10 nM pol β with 5 μM or 8 μM of NSC666719.

Figure 2.10. Inhibition of DNA synthesis activity of WT pol β promotes (CAG)₂₀ repeat deletions

(B) Sizing analysis of the repaired products with the inhibition of 1 nM pol β. Panel a illustrates the (CAG)₂₀ size marker. Panel b illustrates the repaired products without inhibitor. Panel c and d illustrates the sizes of the repaired products with the inhibition of DNA synthesis activity of 1 nM pol β by NSC666719 (5 μM and 8 μM). **(C)** Sizing analysis of the repaired products with the inhibition of 10 nM pol β. Panel a illustrates the (CAG)₂₀ size marker. Panel b illustrates the repaired products without the inhibitor. Panel c illustrates the sizes of the repaired products with the inhibition of DNA synthesis activity of 10 nM pol β by NSC666719 (8 μM).

Weak DNA synthesis activity of pol βY265C does not significantly affect the instability of dinucleotide GT repeats and telomere repeats

Microsatellite DNA also consists of dinucleotide repeats and hexanucleotide repeats [91]. The repeat deletions are associated with many human cancer and aging-related diseases [184, 256]. Deletions of GT repeats are specifically associated with colorectal cancers [184]. Deletions of telomere repeats, TTAGGG repeats, are associated with aging-related diseases, such as diabetes, Alzheimer's disease, and hypertension [256]. Although inducing repeat deletions on the expanded TNRs is beneficial for the

individuals who are prone to TNR expansion that mediates neurodegenerative diseases, deletions on GT repeats and telomere repeats may lead to cancer development and aging-related diseases. Therefore, we determined whether the weak DNA synthesis activity of pol β Y265C would promote deletions of GT repeats and telomere repeats.

First, we determined the DNA synthesis and repeat instability in the presence of pol β Y265C with the substrate containing (GT)₃₀ with a lesion located in the middle. The results showed that both WT pol β and pol β Y265C exhibited more efficient DNA synthesis activities on GT repeats than those on CAG or CTG repeats (Compare Figure 2.11 with Figure 2.1 or Figure 2.6). Pol β Y265C exhibited weaker DNA synthesis activity than WT pol β . It incorporated 10 nt at the concentration of 2 nM (Figure 2.11A, lane 4) compared to 14 nt incorporated by WT pol β (Figure 2.11A, lane 3). Analysis on sizes of the repeats in the repaired products showed that both WT pol β and pol β Y265C induced 1 repeat deletion products (Figure 2.11B, panels b and c). However, the weaker DNA synthesis activity of pol β Y265C on GT repeats did not significantly alter the repeat instability during BER (compare panel c to panel b in Figure 2.11B).

Figure 2.11. Pol βY265C does not significantly alter (GT)₃₀ repeat instability.

(A) Measurement of DNA synthesis activity and BER reconstitution were conducted measured with the (GT)₃₀ containing substrate with a THF residue located at the fifteenth repeat. Lane 1 represents substrate only. Lane 2 represents APE1 cleavage products. Lanes 3 and 4 represent the DNA synthesis products of 2 nM WT pol β and pol βY265C, respectively. Lanes 5 and 6 represent the repaired products in the presence of 2 nM WT pol β and pol βY265C, respectively. (B) Sizes measurement of sizes of the repeats in the repaired products. Panel a represents the (GT)₃₀ marker. Panels b and c represent the lengths of the repaired products in the presence of 2 nM WT pol β and pol βY265C, respectively.

We further evaluated the effect of weak DNA synthesis activity of pol βY265C on telomere repeat instability through BER. With the (TTAGGG)₁₀-containing substrate with the lesion located at the last G of the fifth repeat, the DNA synthesis activity of WT pol β and pol βY265C did not exhibit a significant difference. WT pol β inserted 1 repeat, 2 repeats, and 3 repeats (Figure 2.12A, lanes 3-5), whereas pol βY265C incorporated 2/3 repeat, 1.5 repeats, and 2.5 repeats at the concentrations of 2 nM, 5 nM, and 10 nM (Figure 2.12A, lanes 6-8), respectively. However, neither WT pol β nor pol βY265C resulted in instability of telomere repeats (compare panels b and c to panel a in Figure 2.12B). The results indicated that the weak DNA synthesis did not significantly alter instability of GT repeats and telomere repeats. The results further suggested that

shortening the expanded TNRs by inhibiting pol β DNA synthesis may not confer a significantly high risk of development of cancers and aging-related diseases, which are associated with GT repeat or telomere repeat deletions.

Figure 2.12. Pol β Y265C does not significantly alter (TTAGGG)₁₀ repeat instability.

(A) Measurement of DNA synthesis activity and BER reconstitution were conducted with the (TTAGGG)₁₀-containing substrates with a THF residue located at the fifth repeat. Lane 1 represents substrate only. Lane 2 represents APE1 cleavage products. Lanes 3-5 and lanes 6-8 represent the DNA synthesis products of pol β and pol β Y265C at the concentrations of 2 nM, 5 nM, and 10 nM, respectively. Lanes 9-11 and lanes 12-14 represent the repaired products in the presence of WT pol β and pol β Y265C at the concentrations of 2 nM, 5 nM, and 10 nM, respectively. (B) Sizing analysis of the repaired products by DNA fragment analysis. Panel a represents the (TTAGGG)₁₀ marker. Panels b and c represent the lengths of the repaired products in the presence of 2 nM WT pol β and pol β Y265C.

DISCUSSION

Trinucleotide repeat expansion diseases including Huntington's disease, Friedreich's ataxia, and myotonic dystrophy are fatal inherited neurological disorders [88, 89, 92]. Despite decades of studies, no cure has been developed for these diseases, and only symptomatic managements are available for the patients [244-246]. Current studies have mainly focused on alleviating the symptoms of patients, such as reduction of toxic RNAs resulted from the transcription of expanded TNRs in the noncoding regions [257-

259], decreation of the toxicity caused by aggregations of polyglutamine proteins encoded by expanded CAG repeats [260-262], and suppression of toxic gene expression via RNA interference [263, 264]. However, since expanded TNRs are increasingly expanded, repeat expansions and cellular toxicity, therefore, accumulate at an accelerated rate. These approaches do not provide a long-term solution for the patients, nor do they provide an effective treatment targeting the root of TNR expansion diseases. Our studies suggest that a novel strategy of shortening expanded TNR tracts can be achieved by inhibiting the DNA synthesis activity of pol β through BER, provideing a treatment of TNR expansion diseases by eradicating their roots. This may effectively reverse and attenuate the courses of the diseases.

Previous studies have demonstrated that BER plays an essential role in mediating TNR expansions in somatic cells, and pol β plays a critical role in BER-mediated TNR expansions [1, 14, 122, 144]. In the striatum of HD mice that has high frequency of CAG repeat expansions, a high level of pol β and a low level of FEN1 is observed [122, 249]. On the other hand, CAG repeat expansions barely occur in the cerebellum of the disease mice, and the levels of pol β and FEN1 are balanced [122, 249]. This suggests that pol β DNA synthesis predominantly promotes TNR expansions through BER [2]. In this study, we demonstrated that the weak synthesis by pol β Y265C and the weak synthesis by inhibiting WT pol β DNA could both lead to TNR deletion. This indicates that TNR expansions can be shifted to deletions by regulating the pol β DNA synthesis activity the outcome from the to TNR deletions.

Our study further suggested a general role of DNA synthesis in regulating TNR instability. Through characterizing the effects of the translesion polymerases, pol ν and

pol θ on TNR instability, we discovered that pol ν , which exhibited weak DNA synthesis activity resulted in CAG repeat deletions (Figure 2.4), whereas pol θ , which exhibited strong DNA synthesis, did not significantly alter CAG repeat instability (Figure 2.5). The results suggest a fundamental role of weak DNA synthesis in promoting TNR deletions through BER. During the weak DNA synthesis-mediated TNR deletions, weak DNA synthesis of pol β Y265C and pol ν resulted in addition of fewer repeats than those removed by FEN1 (Figure 2.3 and Figure 2.4). This further suggests that regulation of TNR instability through BER is governed by the balance between the synthesis of TNRs by the polymerases and the cleavage of TNRs by the nucleases such as FEN1 and Mus81/Eme1. Our results further demonstrate that translesion DNA synthesis can actively involved in modulating TNR instability during DNA replication and repair.

In our current study, we also demonstrated that the sequence of TNR tracts may play a role in modulating TNR instability by altering DNA synthesis activity of repair DNA polymerases. TNRs such as (GAA/CTT)_n, (CAG/CTG)_n, and (CGG/GCC)_n have increasing thermostabilities because their GC pairs in each repeat unit increases, and the repeats with higher thermostabilities are more challenging for DNA polymerases to separate them to perform the strand displacement synthesis. This may reduce the efficiency of TNR synthesis by pol β and pol ν , leading to weak synthesis of TNRs during BER. We found that pol β Y265C performed similar level of DNA synthesis on (CAG/CTG)_n and (CTG/CAG)_n tracts, resulting similar deletions of TNRs during BER. Pol ν exhibited weak DNA synthesis activity and promoted more severe deletions of (CGG/GCC)_n than it did on (CAG/CTG)_n (compare the height of the peaks in Figure 2.7B to that in Figure 2.4B with the concentrations of 2 nM and 5 nM of pol ν). The sequence

differences in DNA synthesis and repeat instability may attribute to the higher thermostability of (CGG/GCC)_n than that of (CAG/CTG)_n. On the other hand, pol θ exhibited a marked difference in DNA synthesis on the (GAA/CTT)_n from that on the (CAG/CTG)_n, causing different outcomes of repeat instabilities. On the (CAG/CTG)_n, pol θ exhibited similar DNA synthesis activity as pol β and did not alter the instability of (CAG/CTG)_n (Figure 2.5). However, it performed weaker DNA synthesis activities on (GAA/CTT)₂₀ tract and resulted in GAA repeat deletion (Figure 2.8). The results indicate that pol θ may not be able to efficiently bind to a more dynamic upstream 3'-OH in the upstream strand of the GAA repeat tract during BER of a base lesion located in the middle of the repeat tract. In addition, DNA sequences may also affect DNA synthesis activities of polymerases on the TNRs: polymerases have stronger DNA synthesis on their preferred template. This is supported by the fact that the translesion polymerase, pol η, binds more tightly to the sequence that can form G-quadruplex and performs more efficiently DNA synthesis than that of the random sequence, whereas pol ε has a reduced binding and synthesis on the same sequence that can form G-quadruplex [265]. Moreover, pol ι has a preference to incorporate nucleotides opposite to a template A [266], and pol ν has preferentially misincorporates dTTP to a template G [267]. The substrate preferences and different DNA synthesis fidelity of DNA polymerases may further explain the differences of DNA synthesis activities on different repeat sequence, particularly with the weaker DNA synthesis of pol ν on the CGG repeats than on the CAG repeats.

Since deletions of GT and telomere repeats are associated with cancers and aging-related diseases [184, 256], here, we also evaluated how the weak DNA synthesis activity

can affect instability of GT repeats and telomere repeats (TTAGGG repeats) during BER. Our results showed that WT pol β and pol β Y265C did not exhibit significant differences in synthesizing GT repeats and TTAGGG repeats as well as in modulating the repeat instability (compare Figure 2.11 to Figure 2.1-2 and Figure 2.6). The similar amount of GT repeat deletion products induced by WT pol β and pol β Y265C indicates that the slightly weaker DNA synthesis activity of pol β Y265C on GT repeats than that of WT pol β Y265C does not alter the outcome of GT repeat instability. On the other hand, the telomere sequence requires at least 6 nucleotides for misalignments, which make it difficult for the formation of secondary structures, especially on the template strand that can facilitate repeat instability during BER. This is supported by our results showing that neither WT pol β nor pol β Y265C failed to cause repeat instability (Figure 2.12). The results further suggest that inhibition of DNA synthesis of pol β can be employed to contract expanded TNRs without significantly affecting the stability of other repeats that are associated with development of cancers and aging-related diseases.

Taken together, our study has demonstrated that weak DNA synthesis activity of repair DNA polymerases can lead to TNR deletion through BER. We found that repeat sequence of TNRs can modulate altering DNA synthesis activities of polymerases and thus contributes to a different outcome of TNR instability. We demonstrate that TNR deletions can be induced by inhibiting DNA synthesis activity of pol β during BER, and the inhibition of DNA synthesis does not affect instability of GT repeats and telomere repeats. Our study further suggests that transient inhibition of DNA synthesis activity can be potentially developed as a therapeutic strategy for treating TNR expansion neurological disorders.

CHAPTER 3: MONO-UBIQUITINATION OF PROLIFERATING CELL NUCLEAR ANTIGEN MODULATES TRINUCLEOTIDE REPEAT INSTABILITY BY COORDINATING WITH FANCONI ANEMIA-ASSOCIATED NUCLEASE 1 AND FLAP ENDONUCLEASE 1

ABSTRACT

Trinucleotide repeat (TNR) instability is associated over 40 neurodegenerative diseases including Huntington's disease. Previous studies have indicated that PCNA promotes FEN1 5'-flap cleavage activity to prevent TNR expansion. Yet the role of post-translational modification of PCNA in TNR instability remains to be elucidated. Recent studies have indicated that monoubiquitinated PCNA (ub-PCNA) can physically interact with the Fanconi anemia-associated nuclease 1 (FAN1), which is a 5'-endo/exonuclease and a genetic modifier of polyglutamine diseases including Huntington's disease. This suggests that ub-PCNA may modulate TNR instability by cooperating with FAN1. In our current study, we discover that ub-PCNA stimulated 5'-3' exonucleolytic activity of FAN1 in processing a TNR hairpin, which destabilized a hairpin, converting it into a flap that was then cleaved by FAN1 5'-endonuclease activity or FEN1. The coordination among ub-PCNA, FAN1, and FEN1 led to removal of TNR hairpins, preventing TNR expansion. Our study provides the first evidence that the cooperation among ub-PCNA, FAN1, and FEN1 plays a critical role in preventing TNR expansions. Our study further suggests that an essential role of post-translation modification of PCNA in maintaining TNR stability by coordinating nucleases in removing secondary structures during BER on TNRs.

INTRODUCTION

Trinucleotide repeat instability is associated with more than 40 neurodegenerative diseases including Huntington's disease (CAG)_n, fragile X syndrome (CGG)_n, Friedreich's ataxia (GAA)_n, and myotonic dystrophy (CTG)_n [88, 89, 92]. Trinucleotide repeat expansions can occur during maternal transmission or after cellular differentiation [92]. However, recent studies show that BER is responsible for TNR expansions in post-mitotic neuron cells [2, 14, 122]. Base excision repair mediates TNR expansions is supported by the fact that oxidative DNA damage facilitates TNR expansions in somatic cells through multiple rounds of BER in TNR tracts [14, 123, 131]. Base excision repair contributes to TNR expansions is further supported by the fact that TNR expansions consistently occur in the neuronal cells of Huntington's disease transgenic mice through the repair of 8-oxoG [14]. The mechanism of BER mediated TNR instability has been proposed that the formation of secondary structures, such as hairpins, loops, and G4-quadruplex, which disrupts enzymatic coordination during BER, lead to extra repeat synthesized by DNA polymerases or removal of extra repeats by nucleases, resulting in repeat expansions or deletions [1, 5, 7, 142]. It has been found that inefficient cleavage from flap endonuclease 1 (FEN1) promotes TNR expansions through BER [1, 5]. The hairpins inhibit FEN1 5'-flap cleavage and force the enzyme to adopt an alternate flap cleavage of a short flap, resulting in ligation of a TNR hairpin and repeat expansions [1, 5]. In addition, the mismatch repair protein complex, MSH2-MSH3, inhibits FEN1 cleavage by binding and stabilizing a TNR hairpin or loop that formed on downstream damage strand, while it stimulates the synthesis of TNRs by pol β [2]. Moreover, a low

level of FEN1 and high level of pol β has also been observed in the neuronal cells of striatum of Huntington's disease (HD) mice that exhibits an abundance of CAG repeat expansions, suggesting that FEN1 plays a critical role in preventing TNR expansion by efficiently removing the repeats during BER [122, 144]. Facilitation of the removal of TNRs by other nucleases can also prevent TNR expansion and promote TNR deletions [9, 145]. The 3'-5' exonucleolytic activity of AP endonuclease 1 (APE1) can remove TNRs from the upstream strand exonucleolytically during BER of a TNR tract [9]. The Mus81/Eme1 can remove upstream 3'-flap endonucleolytically [145]. In cooperating with FEN1, APE1 and Mus81/Eme1 facilitate the removal of TNRs, thereby suppressing TNR expansions [9, 145]. Other DNA nucleases including Exonuclease 1 (EXO1), Xeroderma pigmentosum complementation group G endonuclease (XPG), and Fanconi anemia-associated nuclease 1 (FAN1) [170, 268] may also facilitate the removal of TNRs and prevent repeat expansions. In particular, FAN1 may participate in suppressing TNR expansions since single nucleotide polymorphisms of FAN1 have been associated with early or late onset of HD [173, 174]. Fanconi anemia-associated nuclease 1, which possesses both 5'-3' exonucleolytic activity [269] and 5'-flap endonucleolytic activity [270], may cooperate with FEN1 in removing TNR hairpins and prevent repeat expansions.

Proliferating cell nuclear antigen is a DNA replication and repair cofactor, which acts as a central regulator to facilitate DNA replication and DNA repair through protein-protein interactions [77]. At least 200 proteins contain the PCNA-interacting protein (PIP) box, a conserved 8-amino-acid motif that can physically interact with PCNA [78]. During BER in TNRs, PCNA assists in resolving a hairpin structure by stimulating FEN1

5'-flap endonucleolytic activity and LIG I catalytic activity and thus facilitates shortening of the expanded TNR [149]. PCNA can also regulate its binding partners through the monoubiquitination of PCNA [4]. Monoubiquitination of PCNA can be induced by DNA damage agents such as methyl methanesulfonate (MMS) and UV irradiation [167]. Repair of MMS-induced alkylating DNA damage is subjected to the BER pathway [168], indicating that PCNA can be monoubiquitinated during repair of base lesions. In fact, ub-PCNA can physically interact with FAN1 with a high binding affinity via the ubiquitin-binding zinc domain (UBZ) and a non-canonical PIP box located at the N-terminal domain of FAN1 [159, 160], suggesting a potential role of FAN1 in BER-mediated TNR instability that can be regulated by ub-PCNA. To test this possibility, we characterized the effects of ub-PCNA on FAN1 during BER and find that ub-PCNA stimulated the 5'-3' exonucleolytic activity of FAN1. The stimulation allowed FAN1 to resolve a CAG repeat hairpin structure by destabilizing the hairpin through its 5'-3'-exonuclease leading to the removal of the hairpin through BER. Our study provides the first evidence that the coordination between ub-PCNA and FAN1 plays a critical role in suppressing TNR expansions during BER. Our study further suggests that an essential role of post-translation modification of PCNA in maintaining TNR stability by coordinating DNA nucleases in removing secondary structures during BER.

MATERIALS AND METHODS

Materials and oligonucleotide substrates

DNA oligonucleotide substrates were synthesized by Integrated DNA Technologies (Coralville, IA, USA). The Cordycepin (3'-deoxyadenosine) 5'-

triphosphate 3'- α - ^{32}P (5000 mCi/mmol) and γ - ^{32}P ATP (6000 mCi/mmol) were purchased from PerkinElmer (Boston, MA). Terminal deoxynucleotidyl transferase and T4 polynucleotide kinase were purchased from Thermo Fisher Scientific (Waltham, MA). Micro Bio-Spin 6 Columns were purchased from Bio-Rad Laboratories (Hercules, CA). All other standard chemical reagents were purchased from Sigma-Aldrich (St. Louis, MO) or Thermo Fisher Scientific (Waltham, MA). Purified human enzymes including APE1, wild-type pol β , FEN1, and DNA LIG I were made according to the protein purification procedures described previously [51, 145]. Human ubiquitin was purchased from Boston Biochem (Boston, MA).

Oligonucleotide substrates

The nicked flap substrate containing a (CAG)₁₅ was made by annealing the upstream primer, downstream primer, and the template strand with a molar ratio of 4:1:2. The hairpin substrate containing a (CAG)₁₁ hairpin and a nicked (CAG)₁ 5'-flap was made by annealing the upstream primer, downstream primer, and the template strand with a molar ratio of 4:1:2. The sequences of substrates are listed in Table 3.

Table 3. Oligonucleotide Sequences

Oligonucleotide	nt	Sequence (5'-3')
5'-flap substrate
Upstream strand-(CAG) ₁₅	65	CGA GTC ATC TAG CAT CCG TA CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG

Downstream strand-(CAG) ₁₅	65	CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG TA CGT AGA CTT ACT CAT TGC
Template strand-(CAG) ₂₀	99	GCA ATG AGT AAG TCT ACG TA CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG TA CGG ATG CTA GAT GAC TCG
Hairpin substrate

Upstream strand	38	CGA GTC ATC TAG CAT CCG TA GCA GTC CTC TAG TCG TAG
Downstream strand	65	CAG CTC GAG TGG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAGTA CGT AGA CTT ACT CAT TGC
Template strand	100	GCA ATG AGT AAG TCT ACG TA CCA CTC GAG CTA CGA CTA GAG GAC TGC TAC GGA TGC TAG ATG ACT CG

Protein purifications

The recombinant plasmids of UBE1 (Ube1/pET21d) and UbCH5c S22R (UbCH5c S22R pET28a) were purchased from Addgene (Cambridge, MA). The expression vector with UBE1 and UbCH5c S22R were then transformed into *E. coli* BL21DE3 purchased from Thermo Fisher Scientific (Waltham, MA). The recombinant FAN1/pGEX-2T plasmid was a gift from Dr. Josef Jiricny from the University of Zurich.

Purification of FAN1 is followed by the procedures described previously [271-273].

The purification of human PCNA was followed by the procedures described previously [274] with modifications. The PCNA/pT-7 was transformed and expressed in

E. coli BL21DE3 (Agilent Technologies, Santa Clara, CA), which was sequentially inoculated and expressed in 6 L LB with OD researched to 0.6 with addition of 1 mM IPTG. Cell pellets were resuspended in the lysis buffer with the pH at 7.5 that contains 0.1 M KCl, 20 mM Tris-HCl (pH 7.5), 30 mM spermidine, 10% sucrose, 0.5 mM ethylenediaminetetraacetic acid (EDTA), 2 mM dithiothreitol (DTT), and 0.2 M phenylmethylsulfonyl fluoride (PMSF). The PCNA was purified followed by a 7-column procedure. Briefly, soluble proteins and cell debris were separated by centrifugation at 12,000 rpm, 4°C for 30 min. The supernatant was subjected to 40 ml P11 column from Sigma-Aldrich (St. Louis, MO) for purification. Proteins were eluted by elution buffer containing 0.1 M KCl, 20 mM Tris-HCl (pH 7.5), 0.5 mM EDTA, 10% glycerol, 2 mM DTT, and 0.2 mM PMSF. The flowthrough was collected and applied to the 12-ml DEAE sepharose column from Thermo Fisher Scientific (Waltham, MA). The protein was eluted with buffer containing 0.5 M KCl, 20 mM Tris-HCl (pH 7.5), 0.5 mM EDTA, 10% glycerol, 2 mM DTT, and 0.2 mM PMSF. Peak fractions were collected and dialyzed into buffer that contains 1 M ammonium sulfate, 50 mM KCl, 20 mM Tris-HCl (pH 7.5), 0.5 mM EDTA, 10% glycerol, 2 mM DTT, and 0.2 mM PMSF. The hPCNA was eluted with buffer containing 50 mM KCl, 20 mM Tris-HCl (pH 7.5), 0.5 mM EDTA, 10% glycerol, 2 mM DTT, and 0.2 mM PMSF. The peak fractions were then combined and dialyzed into the buffer A containing 0.1 M KCl, 20 mM Tris-HCl (pH 7.5), 0.5 mM EDTA, 10% glycerol, 2 mM DTT, and 0.2 mM PMSF. Proteins were then separated by the 10-ml Q sepharose column (GE Healthcare Bio-Sciences, Pittsburgh, PA) with the elution buffer containing 0.8 M KCl, 20 mM Tris-HCl (pH 7.5), 0.5 mM EDTA, 10% glycerol, 2 mM DTT, and 0.2 mM PMSF. The peak fractions were combined and dialyzed into the buffer

A. Proteins were then applied to Heparin-Sepharase column (GE Healthcare Bio-Sciences, Pittsburgh, PA) and the flowthroughs were then dialyzed into the buffer containing 0.2 M KCl, 50 mM sodium acetate (pH 5.5), 0.5 mM EDTA, 10% glycerol, 2 mM DTT, and 0.2 mM PMSF. The flowthroughs were then collected for final purification with Mono Q (GE Healthcare Bio-Sciences, Pittsburgh, PA) with the elution buffer containing 1 M KCl, 20 mM Tris-HCl (pH 7.5), 0.5 mM EDTA, 10% glycerol, 2 mM DTT, and 0.2 mM PMSF. Peak fractions were combined and dialyzed into the storage buffer containing 150 mM KCl, 20 mM Tris-HCl (pH 7.5), 0.5 mM EDTA, 10% glycerol, 2 mM DTT, and 0.2 mM PMSF. The hPCNA was then aliquoted and froze at -80°C for storage.

The UBE1 was expressed in *E. coli* BL21DE3 AI strain (Agilent Technologies, Santa Clara, CA). Cells transformed with UBE1/pET21 plamids were inoculated into 1 L LB with OD researched to 0.6, UBE1 was expressed with addition of 1 mM IPTG and 0.1% arabinose. Cell pellets were resuspended in the lysis buffer that contains 50 mM NaH₂PO₄ (Ph7.8), 300 mM NaCl, 30 mM imidazole, 10 mM EDTA, 10 mM DTT, 1 mM PMSF, 1 mM benzamidin, 1 µg/ml leupeptin, and 1 µg/ml pepstatin A. Soluble proteins and cell debris were then separated by centrifugation at 12,000 rpm, 4°C for 30 min. The supernatant was subjected to Ni-NTA agarose column from Qiagen (Hilden, Germany) for purification. Proteins were eluted with elution buffer containing 50 mM NaH₂PO₄ (pH7.8), 300 mM NaCl, 500 mM imidazole, and 1 mM PMSF. Peak fractions were collected and dialyzed into buffer that contains 50 mM NaCl, 20 mM Tris-HCl (pH 7.8), and 1 mM PMSF. Proteins were separated by Mono Q column (GE Healthcare Bio-Science) with the elution buffer containing 400 mM NaCl, 20 mM Tris-HCl (7.8), and 1

mM PMSF. Eluted peak fractions were combined, aliquoted, and froze at -80°C for storage.

The Ubch5c S22R (Ubch5c S22R pET28a) plasmids were transformed in *E. coli* BL21DE3 (Agilent Technologies, Santa Clara, CA). A bacterial colony was sequentially inoculated into 1 L LB with OD researched to 0.8. The Ubch5c S22R was expressed with addition of 1 mM IPTG. The Ubch5c S22R was initially purified with the Ni-NTA column and then purified with SP column. Cell pellets were resuspended in the lysis buffer that contains 50 mM NaCl, 20 mM Bis-tris (pH 6.0), 1 mM EDTA, 1 mM PMSF, 1 mM benzamidin, 1 µg/ml leupeptin, and 1 µg/ml pepstatin A. Soluble proteins and cell debris were then separated by centrifugation at 12,000 rpm, 4°C for 30 min. The supernatant was subjected to SP column (GE Healthcare Bio-Science) for purification. Proteins were eluted with elution buffer containing 1 M NaCl, 20 mM Bis-tris (pH 6.0), 1 mM EDTA, and 1 mM PMSF. Peak fractions were collected and dialyzed into buffer that contains 50 mM NaCl, Tris-HCl (pH 7.8), and 1 mM PMSF. Proteins were separated by Mono Q column (GE Healthcare Bio-Science, Uppsala, Sweden) with the elution buffer containing 400 mM NaCl, 20 mM Tris-HCl (7.8), and 1 mM PMSF. Eluted peak fractions were combined and dialyzed into the storage buffer that contains 25 sodium phosphate, 150 mM NaCl, and 20% glycerol at pH 7.0. Ubch5c S22R was then quoted and froze at -80°C for storage.

***In vitro* construction and purification of ub-PCNA**

The *in vitro* ubiquitination of PCNA was performed according to the conditions described in a previous study [275]. The reaction-condition was optimized, and the final

reaction was assembled on ice in a 500- μ l reaction mixture containing 160 nM Ube1, 32 μ M UbcH5c S22R, 32 μ M ubiquitin, and 16 μ M PCNA with the buffer conditions of 50 mM MMT (pH 9.0), 25 mM NaCl, 3 mM MgCl₂, 0.5 mM TCEP, and 3 mM ATP. The reactions were incubated in the water bath at 37 °C for 2 h.

The reaction-mixtures were further dialyzed into the buffer A, and the ub-PCNA was then purified by a 3-column procedure. The proteins were then subjected to the 10-ml Q sepharose column (GE Healthcare Bio-Science) with the elution buffer containing 0.8 M KCl, 20 mM Tris-HCl (pH 7.5), 0.5 mM EDTA, 10% glycerol, 2 mM DTT, and 0.2 mM PMSF. The peak fractions were combined and then dialyzed into 1 M ammonium sulfate, 50 mM KCl, 20 mM Tris-HCl (pH 7.5), 0.5 mM EDTA, 10% glycerol, 2 mM DTT, and 0.2 mM PMSF. The ub-PCNA protein was eluted out with 50 mM KCl, 20 mM Tris-HCl (pH 7.5), 0.5 mM EDTA, 10% glycerol, 2 mM DTT, and 0.2 mM PMSF. The peak fractions were pooled and dialyzed into the buffer A. The proteins were subject to the Mono Q column with the same elution buffer for the PCNA purification. The ub-PCNA was then dialyzed into the same storage buffer as PCNA.

***In vitro* enzymatic assays**

FEN1 and FAN1 cleavage assays were performed by incubating FEN1 and FAN1 separately with 100 nM hairpin substrate and flap substrate in reaction buffer that contains 50 mM Tris-HCl, pH 7.5, 50 mM KCl, 0.1 mM EDTA, 0.1 mg/ml BSA, and 0.01% Nonidet P-40 along with 5 mM Mg²⁺ at 37°C for 15 min. The PCNA and the ub-PCNA stimulation assays were performed by incubating FEN1 or FAN1 in the presence of PCNA or ub-PCNA in the same condition described earlier. The 10- μ l reaction mixture

was assembled on ice, and the reaction-mixtures were quenched with addition of 10 μ l stopping buffer containing 95% formamide and 10 mM EDTA with incubation at 95°C for 10 min. Substrates and products were further separated in a 15% urea-denaturing polyacrylamide gel (PAGE) and detected by a Pharos FX Plus PhosphorImager (Bio-Rad Laboratory, CA).

Reconstitution of BER

Reconstituted BER was performed by incubating purified FEN1 or FAN1 along with LIG I with the CAG repeat hairpin- or flap-containing substrates (100 nM) in the reaction mixture (10 μ l) with the same reaction condition described earlier in the Materials and Methods. Reaction mixtures were assembled on ice and incubated at 37°C for 15 min. Reactions were then terminated by addition of 2X stopping buffer with 10-min incubation at 95°C. Substrates and products were then separated in 15% urea-denaturing PAGE and detected by a PhosphorImager.

RESULTS

PCNA purification

PCNA is a homotrimer protein with the total molecular weight (MW) of 87 kDa. The isoelectric point (pI) of PCNA is 4.53. Our previously purified PCNA with N-terminal His-tag was not correctly assembled as homotrimer. Since the His-tag may affect the protein folding, we used PCNA/pT 7 plasmid, which generates a non-tagged PCNA. Since the challenge of the purification of un-tagged PCNA is to eliminate the nuclease and Pol I contaminations, so in each step we have tested the nuclease and Pol I

contaminations in addition to SDS-PAGE to verify the purified protein and to determine the contamination of fractions during purification. Initially, we have verified the expression of PCNA. Since the SDS-PAGE is a denaturing gel, the denatured PCNA protein is shown as a monomer with the MW of 27 kDa, after verification of the expression level, the purification was followed by a 7-column purification. Our strategy in purifying the un-tagged PCNA focused utilization of the charge statuts and hydrophobicity of PCNA protein.

For the first column, we used the P11 column at pH of 7.5. This column was efficient in removing the DNA that bound to the PCNA. Since the PCNA is a DNA binding protein, the negatively charged resins bound to PCNA, which helped to remove DNA from cell lysates. Since PCNA was negatively charged at pH 7.5, the protein was eluted out into the flowthrough (Figure 3.1). After verifying the size of PCNA with SDS-PAGE, the peak fractions containing PCNA were then pooled and applied directly to the 10-ml DEAE column (Figure 3.2). DEAE resins were positively charged at pH 7.5, so the negatively charged PCNA bound to the DEAE resins along with other proteins. After elution, the fractions were tested with SDS-PAGE, and protein concentrations were measured. Majority of the contaminants were removed during the flowthrough and the wash steps (Figure 3.2B-D). The peak fractions (E5-E9) containing PCNA were pooled and dialyzed into the buffer for the Phenyl Sepharose column.

Figure 3.1. P11 phosphocellulose cation exchange chromatography of PCNA.

A. Chromatography profile of fractions after DEAE chromatography. The X-axis represents all the flowthrough and wash fractions. The Y-axis indicates the total amount of protein in each fraction.

Figure 3.2. DEAE ion-exchange chromatography of PCNA.

A. Chromatography profile of fractions after DEAE chromatography. The X-axis represents the elution fractions. The Y-axis indicates the total protein in each fraction. **B, C and D.** Coomassie stained PAGE analysis of selective fractions after DEAE chromatography of PCNA (29 kDa). Selected fractions from flowthrough (F), wash (W) and elution (E) were analyzed in determining the cleaner peak fractions that contains PCNA for purification with next column.

The third column is the Phenyl Sepharose column (Figure 3.3), and the purification was based on the hydrophobicity of different proteins. PCNA was eluted out with ammonium sulfate concentration between 0.2- 0.8 M. Since the PCNA was eluted out with a very broad range of salt concentrations, we have tested every fraction with the SDS-PAGE. Not surprisingly, the PCNA was in each fraction (Figure 3.3B-E). We further tested the nucleases and Pol I contaminations. The fractions of E8-E17 have the minimum amount of the nucleases contamination (Figure 3.3G, lanes 5-14). On the other hand, the fractions of E11-E27 were highly contaminated with Pol I (Figure 3.3F, lanes 6-16). However, based on our previous PCNA purification, nuclease contamination was more challenging to be eliminated than that of Pol I contamination. We combined the fraction E8-E17 for dialysis and prepared for the 4th column: Q column (5 ml) (Figure 3.4). After verified with SDS-PAGE, PCNA was eluted out between fractions of E18-E24 (Figure 3.4B and 3.4C). With further verification of Pol I and nuclease contamination, we observed that Pol I was eliminated in the flowthrough and early wash steps (Figure 3.4D, lanes 2-3). Since fraction E19 and later fractions have the minimal nuclease contamination (Figure 3.4F, lanes 18-24), we combined and dialyzed those fractions for the purification with the Heparin-Sepharose column (10 ml).

Figure 3.3. Phenyl sepharose chromatography of PCNA.

A. Chromatography profile of fractions after phenyl sepharose chromatography. The X-axis represents the elution fractions. The Y-axis indicates the total protein in each fraction. **B.** Coomassie stained PAGE analysis of selective fractions after phenyl sepharose chromatography of PCNA (29 kDa). OP depicts as on-put, which is the protein before loading to phenyl sepharose chromatography. Selected fractions from flowthrough (F), wash (W) and elution (E) were analyzed in determining the cleaner peak fractions that contains PCNA for purification with next column.

Figure 3.3. Phenyl sepharose chromatography of PCNA.

C, D and E. Coomassie stained PAGE analysis of selective fractions after phenyl sepharose chromatography of PCNA (29 kDa). OP depicts as on-put, which is the protein before loading to phenyl sepharose chromatography. Selected fractions from flowthrough (F), wash (W) and elution (E) were analyzed in determining the cleaner peak fractions that contains PCNA for purification with next column.

Figure 3.3. Phenyl sepharose chromatography of PCNA.

F. Pol I contamination test with selected fractions after phenyl sepharose chromatography. **G.** Nuclease contamination test with selected fractions after phenyl sepharose chromatography. Sub depicts as substrate only. OP depicts as on-put, which is the protein before loading to phenyl sepharose chromatography. Selected fractions from flowthrough (F), wash (W) and elution (E) were analyzed in determining the cleaner peak fractions that contains PCNA for purification with next column.

Figure 3.4. Q sepharose anion exchange chromatography of PCNA.

A. Chromatography profile of fractions after Q sepharose anion exchange chromatography. The X-axis represents the elution fractions. The Y-axis indicates the total protein in each fraction. **B and C.** Coomassie stained PAGE analysis of selective fractions after after Q sepharose anion exchange chromatography of PCNA (29 kDa). OP depicts as on-put, which is the protein before loading to this column. Selected fractions from flowthrough (F), wash (W) and elution (E) were analyzed in determining the cleaner peak fractions that contains PCNA for purification with next column.

Figure 3.4. Q sepharose anion exchange chromatography of PCNA.

D. Pol I contamination test with selected fractions after Q sepharose anion chromatography. **E.** Nuclease contamination test with selected fractions after Q sepharose anion exchange chromatography. Sub depicts as substrate only. OP depicts as on-put, which is the protein before loading to this column. Selected fractions from flowthrough (F), wash (W) and elution (E) were analyzed in determining the cleaner peak fractions that contains PCNA for purification with next column.

Heparin Sepharose is a resin that can be used to separate proteins relying on the binding affinity for heparin (GE handbook) and can be used to purify coagulation factors, DNA binding proteins, lipoproteins. Previous studies have successfully separated prototype foamy virus integrase from bacterial nucleases by using heparin column [276], indicating Heparin Sepharose might be a reliable method to eliminate nuclease contamination. In our purification, the PCNA was directly eluted out from the flowthrough step (Figure 3.5), and the fractions were then pooled and dialyzed with the buffer at pH 5.5 for the second round of Heparin Sepharose purification (Figure 3.6). After the second round the Heparin Sepharose column, the fractions from the flowthrough and wash step were then combined and dialyzed for the final purification with Mono Q column (Figure 3.7).

Figure 3.5. The first Heparin chromatography of PCNA.

Coomassie stained PAGE analysis of selective fractions after the first heparin affinity chromatography of PCNA (29 kDa). OP depicts as on-put, which is the protein before loading to this column. All the fractions from flowthrough (F), wash (W) were analyzed in determining the cleaner peak fractions that contains PCNA for purification with next column.

Figure 3.6. The second Heparin chromatography of PCNA.

A. Chromatography profile of fractions after the second heparin affinity chromatography. The X-axis represents the flowthrough, wash and elution fractions. The Y-axis indicates the total protein in each fraction. **B and C.** Coomassie stained PAGE analysis of selective fractions after after the second heparin affinity chromatography of PCNA (29 kDa). OP depicts as on-put, which is the protein before loading to this column. All the fractions from flowthrough (F), wash (W) were analyzed in determining the cleaner peak fractions that contains PCNA for purification with next column.

Mono Q resin was positive charge at pH 7.5, so it concentrated PCNA in the column. The PCNA was eluted out during the elution step (Figure 3.7B and 3.7C). The Pol I and nuclease contaminations were tested. No Pol I contamination was observed (Figure 3.7D) and only very limited nuclease contamination existed in the peak fractions of E22-E24 (Figure 3.7E, lanes 17-19). The peak fractions (E22-E32) were also tested for the functional stimulation of FEN1 5'-flap endonuclease activity. Compared to the FEN1 alone and BSA control (Figure 3.7F, lanes 2-3), all PCNA from the peak fractions stimulated FEN1 cleavage (Figure 3.7, lanes 6-16). Therefore, the fractions of E25-E32 were combined and dialyzed into the storage buffer for the future experiments.

Figure 3.7. Mono Q anion exchange chromatography of PCNA.

A. Chromatography profile of fractions after Mono Q anion exchange chromatography. The X-axis represents the elution fractions. The Y-axis indicates the total protein in each fraction. **B and C.** Coomassie stained PAGE analysis of selective fractions after Mono Q anion exchange chromatography of PCNA (29 kDa). Selected fractions from flowthrough (F), wash (W) and elution (E) were analyzed in determining the cleaner and functional peak PCNA for storage.

Figure 3.7. Mono Q anion exchange chromatography of PCNA.

D. Pol I contamination test with selected fractions after Mono Q anion exchange chromatography. **E.** Nuclease contamination test with selected fractions after Mono Q anion exchange chromatography. OP depicts as on-put, which is the protein before loading to this column. **F.** Functional stimulation of FEN1 nuclease cleavage activity by PCNA. Lane 1 represents substrate only. Lane 2 represents FEN1 alone. Lane 3 illustrates the negative control of BSA in stimulating FEN1 cleavage. Lane 4 illustrates the positive control of PCNA (Enzymax, Lexington, KY) in stimulating FEN1 cleavage. Selected fractions from flowthrough (F), wash (W) and elution (E) were analyzed in determining the cleaner and functional peak PCNA for storage.

UBE1, UbcH5c S22R purification

The UBE1 has the MW of 118 kDa with carboxyl-terminal polyhistidine-tag. The pI of UBE1 is 4.53. The UBE1 did not express well in the BL21 DE3 strain. The expression of UBE1 in BL21 DE3 AI strain gave a higher yield (Figure 3.8). The purification of UBE1 has subjected to 2 columns: the Ni-NTA column and the Mono Q column. The UBE1 bound well with the Ni-NTA resins as it was eluted during the elution

step. The peak fractions were combined and dialyzed, which were then subjected to the Mono Q column (Figure 3.9).

Figure 3.8. Expression of UBE1 and UbchH5c S22R.

Coomassie stained PAGE analysis of expression of UBE1 (118 kDa) and UbchH5c S22R (17.5 kDa). The protein without and with the induction of IPTG were depicted as –IPTG and +IPTG, respectively.

Figure 3.9. The Mono Q anion exchange chromatography of UBE.

Coomassie stained PAGE analysis of selective elution fractions after the first heparin affinity chromatography of UBE1 (118 kDa).

On the other hand, the UbchH5c S22R has the MW of 17.5 kDa with C-terminal His-tag. The pI of UBE1 is 7.76. The E2 was efficiently expressed by *E. coli* BL21 DE3

(Figure 3.8). However, E2 protein did not bind to the Ni-NTA resin. We then optimized the purification procedures and purified the protein with SP column (Figure 3.10). The peak fractions of E12-E17 were pooled and dialyzed and later concentrated with Mono Q column.

Figure 3.10. The SP sepharose cation exchange chromatography of UbcH5c S22R.

Coomassie stained PAGE analysis of selective elution fractions after the SP sepharose cation exchange chromatography of UBE1 (17.5 kDa).

***In vitro* construction of ub-PCNA and ub-PCNA purification**

Ubiquitination of PCNA *in vivo* requires the E1 ubiquitination activating enzymes, E2 ubiquitination conjugases, and E3 ubiquitin ligases. These three enzymes coordinate the transportation of ubiquitin to the substrate protein [277, 278]. The E1 protein initiates the activation of ubiquitin by covalently attaching the ubiquitin to its cysteine residue [277, 278]. The activated ubiquitin can then be conjugated to the cysteine residue of E2 [279]. The E3 determines the target specificities by transferring the ubiquitin from E2 to the target protein [280].

Ubiquitination of PCNA *in vivo* is initiated by the RAD6-RAD18 (E2/E3 complex)-mediated monoubiquitination of lysine164 residue of PCNA during DNA repairs [162, 163]. Rad 6 is capable of catalyzing formation of ubiquitin chains, while the RAD18 competitively binds to the ubiquitin, which constrains only one ubiquitin molecule covalently attached to Lys164 of PCNA [162].

According to the conditions described previously [275], we have successfully generated monoubiquitinated PCNA at lysine164 by using E1 (UBE1) and E2 (UbcH5c S22R) *in vitro*. However, only 60% of PCNA was ubiquitinationated (Figure 3.11A, 1-4). In addition, the impurities in the factions of ub-PCNA may compromise our experiments and complicate the purification of ub-PCNA since the ub-PCNA does not have significantly difference of pI, hydrophobicity, and molecular weight from PCNA. To solve the problem, we systematically tested different conditions for ubiquitination of PCNA. The optimized condition as described in the Materials and Methods led to 95%-100% ubiquitination of PCNA (Figure 3.11B, 1-6).

Figure 3.11. Mono-ubiquitination of PCNA.

Coomassie stained PAGE analysis of the yield of the monoubiquitinated PCNA (37.5 kDa). **A.** Mono-ubiquitination with low yield. Lanes 1-4 depict the ubiquitination reactions in a 10- μ l reaction. PCNA, UB, UBE1 and UbcH5c illustrate the control of each protein. **B.** Monoubiquitination of PCNA under an optimal mono-ubiquitination.

To remove the bacterial contaminants in the ub-PCNA fractions, we performed the purification of ub-PCNA with additional 3 columns to eliminate contaminations of Pol I and bacterial nuclease as well as to remove the extra ubiquitin, E1, and E2, which were introduced during the ubiquitination reaction. The ub-PCNA was initially subjected to Q column (Figure 3.12) at pH 7.5. The ubiquitin and E2 were separated with the ub-PCNA during the flowthrough and wash steps (Figure 3.12B). The eluted ub-PCNA fractions (37.5 kDa) were fairly clean (Figure 3.12C). Contaminations of bacterial Pol I and nucleases in the ub-PCNA fractions were tested. The results showed that Pol I was successfully removed and went into the fractions from flowthrough and wash steps (Figure 3.12D). However, nucleases were detected in the peak fraction of E13-14 (Figure 3.12E, lanes 11-12). We further tested the stimulation of pol η DNA synthesis activity by ub-PCNA fractions. A stimulatory effect on pol η DNA synthesis by the peak ub-PCNA fractions (E13-E16) was observed compared to pol η alone or to the BSA control (Figure 3.12F, compare lanes 6-9 to lanes 2-3). The peak fractions of ub-PCNA were combined, dialyzed, and subjected to the Phenyl Sepharose column (Figure 3.13). The ub-PCNA was eluted, and the peak fractions (E12-E19) were combined and subjected to the next column (Figure 3.13A and 3.13C).

Figure 3.12. Q sepharose anion exchange chromatography of ub-PCNA.

A. The chromatography profile of fractions of Q sepharose anion exchange chromatography. The X-axis represents the flowthrough, wash and elution fractions. The Y-axis indicates the total protein in each fraction. **B and C.** Coomassie strained PAGE analysis of selective fractions after after Q sepharose anion exchange chromatography of ub-PCNA (37.5 kDa). Selected fractions from flowthrough (F) and elution (E) were analyzed. OP depicts as on-put, which is the protein before loading to this column.

Figure 3.12. Q sepharose anion exchange chromatography of ub-PCNA.

D. Pol I contamination of selected fractions from Q sepharose anion chromatography was tested. **E.** Nuclease contamination of selected fractions from Q sepharose anion exchange chromatography was examined. **F.** Functional stimulation of pol η (10 nM) DNA synthesis activity by ub-PCNA. Lane 1 represents substrate only. Lane 2 represents pol η alone. Lane 3 illustrates the negative control of BSA in stimulating pol η DNA synthesis activity. Selected fractions from flowthrough (F) and elution (E) were analyzed. Sub depicts as substrate only. OP depicts as on-put, which is the protein before loading to this column.

Figure 3.13. Phenyl sepharose chromatography of ub-PCNA.

A. The profile of phenyl sepharose chromatography of ub-PCNA fractions. The X-axis represents the flowthrough, wash and elution fractions. The Y-axis indicates the total protein in each fraction. **B** and **C.** Coomassie stained PAGE analysis of selective fractions from phenyl sepharose chromatography of ub-PCNA (37.5 kDa). OP depicts as on-put..Selected fractions from flowthrough (F), wash (W) and elution (E) were analyzed in determining the peak fractions that contain ub-PCNA and its purity.

The ub-PCNA fractions were finally subjected to MonoQ column (Figure 3.14), and the fractions were tested for the contamination of bacterial Pol I and nucleases. Neither Pol I (Figure 3.14C) nor the nucleases were detected (Figure 3.14D).

Figure 3.14. Mono Q anion exchange chromatography of ub-PCNA.

A. The chromatography profile of Mono Q column of ub-PCNA fractions. The X-axis represents the flowthrough, wash and elution fractions. The Y-axis indicates the total protein in each fraction. **B.** Coomassie stained PAGE analysis of selective ub-PCNA fractions from Mono Q sepharose anion exchange chromatography of ub-PCNA (37.5 kDa). OP depicts as on-put, which was loaded on the column.

Figure 3.14. Mono Q anion exchange chromatography of ub-PCNA

C. Pol I contamination of selected fractions from Q sepharose anion chromatography was examined. **D.** Nuclease contamination of selected fractions from Q sepharose anion exchange chromatography was examined. Sub depicts as substrate only. OP depicts as on-put, which was loaded onto the column.

FEN1 and FAN1 have different substrate specificities on TNRs

Trinucleotide repeat expansions are mediated by the formation of the secondary structures such as hairpins and loops [14]. TNR hairpins and loops are susceptible to oxidative DNA damage since the exposure of the guanines towards outside of DNA backbone in the hairpin region can increase the probability of oxidation compared to the guanines that are buried inward when they basepair with cytosine [152, 153]. After the OGG1 removes the 8-oxoG and APE1 generates a strand break, it forms a double flap structure with 3'-flap and 5'-flap structures, and this double flap structure is dynamic and can anneal back with the template strand, forming a big 5'-flap or a hairpin structure (Figure 3.15) that can be cleaved by a 5'-flap endonuclease such as FEN1 and FAN1.

Thus, we initially compared the cleavage activity of FEN1 and FAN1 on the substrates with a 5'-(CAG)₁₅ flap or with a downstream 5'-(CAG)₁₁ hairpin, which mimic the intermediates during BER of a base lesion on TNR tracts.

Figure 3.15. The 5'-flap and hairpin structures formed during BER in the context of TNR tracts.

With the flap substrate, FEN1 at 5 nM, 10 nM, 50 nM, 100 nM, and 200 nM removed the 5'-(CAG)₁₅ flap (Figure 3.16A, lanes 4-8). FEN1 also removed the shorter 5'-(CAG)₁ and 5'-(CAG)₂ flaps, suggesting that FEN1 performed the alternate flap cleavage in removing shorter 5'-flaps attached to the transient hairpins. In contrast, FAN1 at the same concentrations only removed the shorter 5'-(CAG)₁ flap that was attached to the transient hairpins (Figure 3.16A, lanes 12-15). The results indicated that FEN1 cleaved both the (CAG)₁₅ flap and short flaps attached to a CAG repeat hairpin, whereas FAN1 only cleaved a short CAG flap attached to a CAG repeat hairpin. Both FEN1 and FAN1 alone failed to remove the transient hairpins.

Figure 3.16. FAN1 and FEN1 have different cleavage patterns in cleaving TNR.

A. FAN1 and FEN1 cleavage activities were examined by incubating increasing concentrations (0.1 nM, 1 nM, 5 nM, 10 nM, 50 nM, 100 nM, and 200 nM) of FEN1 and FAN1 along with the substrate containing a 5'-(CAG)₁₅-flap (illustrated above the gel). Lane 1 represents the substrate only. Lanes 2-8 illustrate the FEN1 cleavage products with increasing concentrations of FEN1. Lanes 9-15 illustrates the FAN1 cleavage products with the increasing concentrations of FAN1.

To further characterize the cleavage of FEN1 or FAN1 on a short flap that was attached to a CAG repeat hairpin, we created a hairpin substrate containing a (CAG)₁₁ hairpin and a 5'-(CAG)₁ flap. With this hairpin substrate, both FEN1 (0.1 nM, 1 nM, 5 nM, 10 nM, 50 nM, 100 nM, and 200 nM) and FAN1 (5 nM, 10 nM, 50 nM, 100 nM, and 200 nM) efficiently removed the 5'-(CAG)₁ flap (Figure 3.16B, lanes 2-8 and lanes 11-15). The results further confirmed that FEN1 and FAN1 performed the alternate flap cleavage activity during BER.

Figure 3.16. FAN1 and FEN1 have different patterns in cleaving TNRs.

B. FAN1 and FEN1 cleavage activities were examined by incubating increasing concentrations (0.1 nM, 1 nM, 5 nM, 10 nM, 50 nM, 100 nM, and 200 nM) of FEN1 and FAN1 with the hairpin-containing substrate (illustrated in the figure). Lane 1 represents the substrate only. Lanes 2-8 illustrate the FEN1 cleavage products. Lanes 9-15 illustrates the FAN1 cleavage products.

Ub-PCNA does not stimulate FAN1 5'-endonucleolytic activity, but significantly enhances FAN1 5'-3' exonucleolytic activity

Previous studies show that ub-PCNA can physically interact with FAN1 with a high binding affinity via UBZ and a non-canonical PIP box located at N-terminal domain of FAN1 [159, 160]. The FAN1 possesses both 5'-endo and exonucleolytic activity. Yet it remains unknown whether interaction between FAN1 and PCNA or ub-PCNA can affect the 5'-end and exonucleolytic activity of FAN1. To test the possibility, we initially determined the effects of PCNA and ub-PCNA on FAN1 5'-endonucleolytic activity by using the 5'-(CAG)₁₅ flap substrate. PCNA or ub-PCNA at 100 nM and 500 nM did not

affect FAN1 cleavage of the (CAG)₁₅ flap or the short (CAG)₁ flap attached to a transient hairpin (Figure 3.17A, compare lanes 2-5 to lanes 6-7). The results indicate that PCNA and ub-PCNA did not stimulate FAN1 5'-endonucleolytic activity.

Figure 3.17. Ub-PCNA stimulates FAN1 5'-3' exonucleolytic activity, but not its 5'-endonucleolytic activity.

The effects of ub-PCNA on the 5'-flap endonucleolytic activity of FAN1 (**A**) and 5'-3' exonucleolytic activity of FAN1 (**B**). FAN1 (2 nM) was incubated with PCNA or ub-PCNA along and 100 nM 5'-(CAG)₁₅-containing substrate and the substrate containing a (CAG)₁ flap attached to a (CAG)₁₁ hairpin. Lanes 1 and 8 represent the substrate only. Lanes 6 and 9 represents FAN1 only. Lanes 2 and 10 represent FAN1 cleavage products in the presence of 100 nM of PCNA. Lanes 4 and 12 represent FAN1 cleavage products in the presence of 500 nM PCNA. Lanes 3 and 11 represent FAN1 cleavage products in the presence of 100 nM of ub-PCNA. Lanes 5 and 13 represent FAN1 cleavage products in the presence of 500 nM ub-PCNA. Lane 7 illustrates the BSA control.

We further determined whether PCNA and ub-PCNA could stimulate FAN1 5'-3' exonucleolytic activity by using the hairpin-containing substrate. We found that FAN1

cleavage generated 1-9 repeat cleavage products (Figure, 3.17B, lane 9), which reached to 5'-end of the hairpin (the intermediates were illustrated on the right of the gel). The result suggests that FAN1 5'-3' exonucleolytic activity removed the short (CAG)₁ flap and continued to progressively cleave the annealed region until it reached to the 5'-end of the hairpin. Surprisingly, we found that both PCNA and ub-PCNA significantly stimulated FAN1 5'-3' exonucleolytic activity (Figure 3.17B, lanes 10-13). However, ub-PCNA stimulated FAN1 cleavage at a lower concentration (100 nM) than PCNA (500 nM) (Figure 3.17B, compare lane 11 to lane 10). The results indicated that both PCNA and ub-PCNA stimulate FAN1 5'-3' exonucleolytic activity.

We then determined the stimulatory effects of PCNA on FEN1 cleavage of the (CAG)₁₅ flap and found that PCNA significantly stimulated FEN1 cleavage of the (CAG)₁₅ flap as well as the alternate flap cleavage of the short (CAG)₁ flap at all the concentrations (10 nM, 25 nM, 50 nM, 75 nM, 100 nM, 200 nM, and 500 nM) of PCNA (Figure 3.18, lanes 3-9). The result is consistent with our previous study that PCNA stimulates FEN1 cleavage of TNRs during BER [149].

Figure 3.18. PCNA stimulates FEN1 5'-flap endonucleolytic activity on TNRs. The effects of PCNA on the 5'-flap endonucleolytic activity of FEN1 (1 nM) at the increasing concentrations (10 nM, 25 nM, 50 nM, 75 nM, 100 nM, 200 nM, and 500 nM). Lane 1 represents the substrate only. Lane 2 illustrates FEN1 only. Lanes 3-9 represent FEN1 cleavage products in the presence of increasing concentrations of PCNA. The red circles illustrate number of the repeats that were removed.

FEN1 and FAN1 alone do not suppress TNR expansions

To further test whether FEN1 and FAN1 could suppress TNR expansion by removing a TNR hairpin, we reconstituted BER with FAN1 or FEN1 and LIG I (5 nM). To determine whether the stimulation of the FEN1 cleavage and the 5'-3' exonucleolytic activity of FAN1 by PCNA and ub-PCNA could facilitate suppression of expanded TNRs, we also reconstituted BER in the absence and presence of PCNA or ub-PCNA.

In the presence of 0.1 nM FEN1 and various concentrations of PCNA or ub-PCNA (100 nM and 500 nM), 5 nM LIG I resulted in ligation of a hairpin (Figure 3.19B, lanes 5-8), suggesting that FEN1 alone failed to suppress the expanded TNRs. On the other hand, in the presence of 10 nM of FAN1 and the PCNA or ub-PCNA (100 nM and

500 nM), 5 nM LIG I also resulted in ligation of a small amount of hairpins (Figure 3.19A, lanes 1-4). However, additional experiments need to be performed to verify whether the stimulation of FAN1 5'-3' exonuclease activity by PCNA and ub-PCNA facilitates repeat deletions since ligation efficiency was low in the absence of DNA polymerases that can fill in the gaps generated by 5'-3' exonucleolytic cleavage of FAN1 in the presence of PCNA or ub-PCNA.

Figure 3.19. Effects of FAN1 and FEN1 on TNR instability.

The effects of FAN1 (10 nM) (**A**) and FEN1 (0.1 nM) (**B**) on TNR expansions mediated by hairpin ligation. The reactions were performed by incubating FAN1 or FEN1 with 100 nM PCNA or ub-PCNA together with 5 nM LIG I and 100 nM hairpin containing substrates. Lanes 1 and 3 represent the ligation products in the presence of FAN1 and PCNA (100 nM and 500 nM). Lanes 2 and 4 represent the ligation products in the presence of FAN1 and ub-PCNA (100 nM and 500 nM). Lanes 5 and 7 represent the ligation products in the presence of FEN1 and PCNA (100 nM and 500 nM). Lanes 6 and 8 represent the ligation products in the presence of FEN1 and ub-PCNA (100 nM and 500 nM).

Ub-PCNA regulates a switch between FEN1 and FAN1 in removing a TNR hairpin

Since FAN1 and FEN1 exhibited different substrate specificity during BER in TNRs, i.e., FAN1 5'-3' exonuclease activity directly removed repeats on CAG repeat hairpin and FEN1 removed 5'-CAG repeat flap, we then tested whether FAN1 and FEN1 could cooperate in removing a CAG repeat hairpin. We sequentially incubated the hairpin substrate with FAN1 and FEN1. Pre-incubation of 10 nM FAN1 with the substrate resulted in removal of up to 9 repeats (Figure 3.20A, lane 1). Subsequent addition of 5 nM FEN1 resulted in removal of 14 repeats (Figure 3.20A, lane 3). The results suggested a cooperative cleavage between FEN1 and FAN1 in removing the big hairpin, thereby suppressing TNR expansions during BER.

Since the requirement of FAN1 can be initiated by ub-PCNA and our results indicate that ub-PCNA facilitated FAN1 5'-3' exonucleolytic activity on the hairpins, we further tested whether the stimulation of FAN1 5'-3' exonucleolytic activity on the hairpin would destabilize the hairpin to convert it into a flap, which could be further removed by FEN1, leading to suppression of expanded TNRs. We found that 10 nM FAN1 and 250 nM ub-PCNA, FAN1 removed up to 14 repeats (Figure 3.20B, lane 4). Subsequent addition of 5 nM FEN1 also increased the amount of cleavage products with 9-14 repeats (Figure 3.20B, lane 6). This further suggested that ub-PCNA regulated FEN1 and FAN1 activities in removing a big hairpin and suppressed the TNR expansions.

Figure 3.20. Ub-PCNA facilitates FAN1 and FEN1 in removing a big hairpin.

The cooperative FAN1 and FEN1 cleavage was measured by sequentially incubating 10 nM FAN1 and 5 nM FEN1 in the absence (A) and the presence (B) of 250 nM ub-PCNA along with 100 nM hairpin substrate. Lanes 1 and 4 represent the FAN1 alone in the absence and the presence of ub-PCNA, respectively. Lanes 2 and 5 represent the FEN1 alone in the absence and the presence of ub-PCNA, respectively. Lanes 3 and 6 represent sequential cleavage of FAN1 and FEN1 in the absence and the presence of ub-PCNA. The red circles illustrate number of the repeats have been removed.

DISCUSSION

Trinucleotide repeat instability is caused by the formation of secondary structures, such as hairpins and loops that disrupt the coordination of BER enzymes and cofactors, can lead to imbalanced synthesis of TNRs by DNA polymerases and cleavage of TNRs by nucleases [1, 5, 7, 142]. The BER cofactors, including the scaffold proteins [281, 282] such as XRCC1 and PCNA, provide a platform for protein recruitment via protein-protein interactions, thereby facilitating the BER coordination [68, 78, 281]. As a cofactor of DNA replication and repair, PCNA can be involved in preventing TNR expansions by

stimulating both FEN1 cleavage of TNRs and LIG I catalytic activity to promote TNR deletion through physically interacting with both enzymes [147-149]. The studies demonstrate a vital role of PCNA in modulating TNR stability. Proliferating cell nuclear antigen can also regulate the activities of its binding partners through post-translational modifications [159-161]. However, how the post-translational modifications of PCNA may affect the TNR instability remains unknown. Here, we provided the first evidence that monoubiquitination of PCNA plays an important role in preventing TNR expansion by regulating the coordination between FEN1 and FAN1 to remove a TNR hairpin during BER.

Faconi anemia-associated nuclease 1 has been implicated to play an important role in modulating CAG repeat instability since the single nucleotide polymorphisms of FAN1 are associated with early or late onset of HD and other polyglutamine diseases. Our results show that FAN1 removed a small flap of (CAG)₁, and continue to perform its 5'-exonucleolytic cleavage of CAG repeats, thereby leading to the removal of a CAG repeat hairpin. Since FEN1 cannot remove the downstream hairpin, the ability of FAN1 to destabilize the hairpin cooperated with FEN1 to remove a TNR hairpin, thus maintaining TNR stability. Our results showed that FAN1 used its 5'-3' exonucleolytic activity to progressively cleave TNRs and destabilized a downstream CAG repeat hairpin. The cleavage by FAN1 converted the hairpin to a flap, which was then efficiently removed by FEN1.

FAN1 can be recruited to DNA damage sites through binding to ub-PCNA or ub-FANCD2 via its UBZ [160, 283]. However, we found that cleavage of the (CAG)₁₄ flap in the presence of both FAN1 and FEN1 did not increase compared with the condition

with FAN1 alone (compare lane 6 of Figure 3.20B to lane 3 of Figure 3.20A). Moreover, FEN1 cleavage of the flap was also decreased. The decreased of FEN1 cleavage products may result from the weaker binding affinity of FEN1 to ub-PCNA. Because FAN1 contains a UBZ, and the binding affinity between FAN1 and the ub-PCNA is much higher than that between PCNA and FEN1, which depends solely on the PIP box interaction. In addition, the structural alignment of ub-PCNA with the PCNA-FEN1 complex reveals that ubiquitin on the ub-PCNA overlaps with the active site of FEN1 in the PCNA-FEN1 complex [4]. This suggests that ub-PCNA interfere with FEN1 flap cleavage. Therefore, the interaction between FAN1 and ub-PCNA makes the 5'-TNR flap inaccessible to FEN1, leading to a weaker FEN1 cleavage of the (CAG)₁₄ flap. It is possible that ub-PCNA coordinated FAN1 and FEN1 by initially stimulating FAN1 cleavage to destabilize the hairpin, which convert the hairpin into a 5'-flap that was subsequently removed by FEN1. On the other hand, de-ubiquitination of ub-PCNA can restore the interaction between PCNA and FEN1, which can promote the cleavage of the 5'-CAG repeat flap. Therefore, the mono-ubiquitination of PCNA and de-ubiquitination of ub-PCNA may act as a switch in coordinating FAN1 and FEN1 cleavage to eliminate the challenging secondary structures formed on TNR tracts during BER.

Our results also support a working model for the cooperation among FAN1, ub-PCNA, and FEN1 in preventing CAG repeat expansion. FAN1 functions as two modes, it forms as a monomer-DNA complex or a head-to-tail dimer-DNA complex [175, 270]. FAN1 monomer is responsible for the 5'-3' exonuclease activity, whereas FAN1 dimer is responsible for cleaving a long 5'-flap (30 nt) [175, 270]. Although our results did not provide the direct evidence showing FAN1 adopts a monomer or dimer when it interacts

with ub-PCNA, the stimulatory effect of ub-PCNA on FAN1 exonuclease activity (Figure 3.17) suggests that ub-PCNA interacted with a FAN1 monomer. Moreover, since the dimer model requires both FAN1 units to contact with DNA via backbone phosphate interaction that is often associated with a conformational change of the protein [175, 270], it is likely that only one FAN1 monomer interacts with ub-PCNA as ub-PCNA would be sandwiched by the two FAN1 monomers. Therefore, ub-PCNA interacted with FAN1 monomer to specifically stimulate its 5'-3' exonuclease activity (Figure 3.17).

Taken together, our study has identified a critical role of mono-ubiquitination of PCNA in suppressing TNR expansions by regulating the nuclease switching between FAN1 and FEN1. The regulation allows stimulation of FAN1 5'-3' exonucleolytic activity, which directly destabilizes a CAG repeat hairpin, converting the hairpin into a flap that can be removed efficiently by FEN1. Therefore, through the cooperation among ub-PCNA, FAN1, and FEN1, a challenging TNR hairpin formed during BER can be removed.

SUMMARY

Base excision repair (BER) is one of the essential DNA repair pathways that safeguards genomic stability and fidelity in both dividing and non-dividing cells. It repairs the most prevalent DNA damage, e.g., DNA base lesions including oxidized and alkylated base lesions, AP lesions, and single strand DNA breaks. Repair of the base lesions in the post-mitotic neuronal cells by BER can promote TNR expansions and deletions, which are associated with more than 40 neurological disorders and cancer. Trinucleotide repeat expansion diseases including Huntington's disease, Friedreich's ataxia, and myotonic dystrophy are devastating genetic neurological disorders. However, no effective cures have been developed as yet. Currently, only symptomatic managements are available for alleviating the disease symptoms of patients. Therefore, understanding the molecular mechanisms that underlie the diseases is critically important for the development of new therapeutic targets for prevention, diagnosis, and treatment of such diseases. It has been shown that the formation of secondary structures, such as hairpins, loops, and G4-quadruplex, can interfere with the coordination among BER enzymes and cofactors, which can lead to extra repeats synthesized by polymerases or removed by nucleases, resulting in repeat expansions or deletions. It appears that the synthesis of TNRs and removal of TNRs, which are fulfilled by DNA polymerases and nucleases, govern TNR stability. However, there are 17 human polymerases and a variety of nucleases, how they mediate TNR instability remains to be explored, and whether they can be utilized as therapeutic targets remains unknown. In addition, while many studies confirm that disruption of the enzymatic coordination during BER is responsible for BER-mediated TNR instability, the mechanisms of cellular regulation of the enzymatic

coordination in maintaining TNR stability are not well understood. Therefore, this Ph.D. dissertation research project focuses on exploring the role of DNA polymerases, nucleases, and their coordination with BER cofactors on TNR instability in search of potential therapeutic targets for repeat instability-mediated diseases.

Initially, we evaluated the effect of a germline polymorphic pol β R137Q variant in leading TNR instability. We find the variant exhibits the similar level of gap-filling synthesis as wild-type pol β . It performs less efficient strand-displacement DNA synthesis than wild-type pol β . We find that pol β R137Q variant exhibits slightly weaker overall DNA synthesis compared to WT pol β . However, it does not significantly affect TNR instability during BER, which may attribute to both pol β R137Q and WT pol β skip over a small loop on the template strand with an equal efficiency. This project provides the direct evidence that the pol β R137Q individual carriers do not have an increased risk of developing TNR instability-related cancers or neurodegenerative diseases.

In the second part of this research project, we determined the roles of DNA synthesis activities in modulating TNR instability. We find that both pol β Y265C mutant and the translesion DNA polymerase, pol ν , have much weaker DNA synthesis activity than WT pol β , whereas the translesion DNA polymerase, pol θ , exhibits similar DNA synthesis activities with the WT pol β . Both pol β Y265C and pol ν can induce large amount of (CAG)₁ repeat deletion, whereas pol θ does not alter the CAG repeat instability compared to that of WT pol β . The weak DNA synthesis of pol β mutant and pol ν results in the removal of more TNRs than synthesized, thus leading repeat deletion. We then demonstrate the effects of TNR sequence on TNR instability. We find that the pol

β Y265C mutant shows similar DNA synthesis activities on CAG and CTG and induces similar effect of repeat instability on both repeat tracts. The pol ν shows weaker DNA synthesis activity on the CGG-repeat tract compared to that on the CAG repeat tract. Therefore, it induces stronger CGG repeat deletion. Meanwhile, pol θ induces GAA repeat deletion, which attributes to its weaker DNA synthesis activity on the GAA repeats compared to that induced by WT pol β . Finally, we explored the effect of transient inhibition of pol β with the pol β inhibitor, NSC666719, on TNR instability; we find that the outcome of TNR instability is shifted towards repeat deletions. We further demonstrate the weak DNA synthesis of pol β Y265C does not significantly alter the instability of GT repeats and telomere repeats, suggesting that weak DNA synthesis by inhibiting pol β with pol β inhibitors may not increase the risks in development of cancers and aging-related diseases.

The first two parts focused on the effects of DNA polymerases and their DNA synthesis activities on the TNR instability and evaluated the probability of DNA synthesis activities of repair DNA polymerases as potential therapeutic targets for disease treatment. While the enzymatic activities can be modulated through mutations or enzyme inhibitors, the cellular regulation of enzymatic activities and enzymatic coordination can be modulated by BER cofactors and by the post-translational modification of BER enzymes and cofactors. In the last part of my dissertation research, we investigated how the monoubiquitinated PCNA (ub-PCNA), a type of post-translational modification of PCNA, regulates the coordination between FAN1 and FEN1 during BER in TNR to alter TNR instability. We find that ub-PCNA greatly enhanced the 5'-3' exonucleolytic activity, and this leads to destabilization of a CAG repeat hairpin, converting the hairpin to a

5'-flap. We further find that the 5'-flap is efficiently removed by FEN1, leading to resolution of the secondary structures and attenuation of the expanded CAG repeats. We demonstrate that ub-PCNA plays a critical role in suppressing TNR expansions by regulating the nuclease switching between FAN1 and FEN1. Our study also suggests a new mechanism that maintains TNR stability by regulating nucleases switching through post-translational modification of PCNA.

By exploring the roles of DNA polymerases, nucleases, and the post-translational modification of BER cofactors in modulating TNR instability, this research can advance our understanding of the molecular mechanisms of BER mediated TNR instability and the cellular pathways that suppress the TNR expansions. We demonstrate that the post-translational modifications of BER cofactors play an important role in maintaining TNR stabilities by regulating enzymatic activities of nucleases. By studying the effect of DNA synthesis activities of DNA polymerases, we further identify a potential therapeutic strategy in treating or preventing TNR expansion mediated neurodegenerative diseases. It should be noted that our studies are limited by using the synthesized oligonucleotides. In cells, TNRs are wrapped by histones and may exhibit different epigenetic features, such as methylation, histone modifications, and among others. All these factors may contribute to TNR instability through BER. Therefore, by dissecting the interplay among BER enzymes, cofactors, and repair proteins from other DNA repair pathways, it can help to formulate a comprehensive understanding of molecular mechanisms underlying BER-mediated TNR instability *in vivo*. With the understanding of the mechanisms, it can further facilitate the development of new therapeutic strategies for disease treatment and prevention.

REFERENCES

1. Liu Y, Prasad R, Beard WA, Hou EW, Horton JK, McMurray CT, et al. Coordination between polymerase beta and FEN1 can modulate CAG repeat expansion. *The Journal of biological chemistry*. 2009;284(41):28352-66. Epub 2009/08/14. doi: 10.1074/jbc.M109.050286. PubMed PMID: 19674974; PubMed Central PMCID: PMCPMC2788885.
2. Lai Y, Budworth H, Beaver JM, Chan NL, Zhang Z, McMurray CT, et al. Crosstalk between MSH2-MSH3 and polbeta promotes trinucleotide repeat expansion during base excision repair. *Nature communications*. 2016;7:12465. Epub 2016/08/23. doi: 10.1038/ncomms12465. PubMed PMID: 27546332; PubMed Central PMCID: PMCPMC4996945.
3. Jacobs AL, Schär P. DNA glycosylases: in DNA repair and beyond. *Chromosoma*. 2012;121(1):1-20. doi: 10.1007/s00412-011-0347-4. PubMed PMID: 22048164; PubMed Central PMCID: PMCPMC3260424.
4. Zhang Z, Zhang S, Lin SHS, Wang X, Wu L, Lee EYC, et al. Structure of monoubiquitinated PCNA: Implications for DNA polymerase switching and Okazaki fragment maturation. *Cell cycle (Georgetown, Tex)*. 2012;11(11):2128-36. doi: 10.4161/cc.20595. PubMed PMID: 22592530; PubMed Central PMCID: PMCPMC3368864.
5. Lai Y, Xu M, Zhang Z, Liu Y. Instability of CTG Repeats is Governed by the Position of a DNA Base Lesion through Base Excision Repair. *PLoS ONE*. 2013;8(2). doi: 10.1371/journal.pone.0056960. PubMed PMID: 23468897; PubMed Central PMCID: PMCPMC3582642.
6. Kumari D, Usdin K. Is Friedreich ataxia an epigenetic disorder? *Clinical epigenetics*. 2012;4(1):2. Epub 2012/03/15. doi: 10.1186/1868-7083-4-2. PubMed PMID: 22414340; PubMed Central PMCID: PMCPMC3305337.
7. Mirkin SM. Expandable DNA repeats and human disease. *Nature*. 2007;447(7147):932-40. Epub 2007/06/22. doi: 10.1038/nature05977. PubMed PMID: 17581576.
8. Liu Y, Wilson SH. DNA base excision repair: A mechanism of trinucleotide repeat expansion. *Trends in Biochemical Sciences*. 2012;37(4):162-72. doi: 10.1016/j.tibs.2011.12.002. PubMed PMID: 22285516; PubMed Central PMCID: PMCPMC3323758.
9. Beaver JM, Lai Y, Xu M, Casin AH, Laverde EE, Liu Y. AP endonuclease 1 prevents trinucleotide repeat expansion via a novel mechanism during base excision

repair. *Nucleic acids research*. 2015;43(12):5948-60. doi: 10.1093/nar/gkv530. PubMed PMID: 25990721; PubMed Central PMCID: PMC4499148.

10. Hoeijmakers JH. Genome maintenance mechanisms for preventing cancer. *Nature*. 2001;411(6835):366-74. Epub 2001/05/18. doi: 10.1038/35077232. PubMed PMID: 11357144.

11. Kim S, Kim KT. Therapeutic Approaches for Inhibition of Protein Aggregation in Huntington's Disease. *Experimental neurobiology*. 2014;23(1):36-44. Epub 2014/04/17. doi: 10.5607/en.2014.23.1.36. PubMed PMID: 24737938; PubMed Central PMCID: PMC4499148.

12. Beard WA, Wilson SH. Structure and mechanism of DNA polymerase Beta. *Chemical reviews*. 2006;106(2):361-82. Epub 2006/02/09. doi: 10.1021/cr0404904. PubMed PMID: 16464010.

13. Ren Y, Lai Y, Laverde EE, Lei R, Rein HL, Liu Y. Modulation of trinucleotide repeat instability by DNA polymerase beta polymorphic variant R137Q. *PLoS ONE*. 2017;12(5):e0177299. Epub 2017/05/06. doi: 10.1371/journal.pone.0177299. PubMed PMID: 28475635; PubMed Central PMCID: PMC5419657.

14. Kovtun IV, Liu Y, Bjoras M, Klungland A, Wilson SH, McMurray CT. OGG1 initiates age-dependent CAG trinucleotide expansion in somatic cells. *Nature*. 2007;447(7143):447-52. doi: 10.1038/nature05778. PubMed PMID: 17450122; PubMed Central PMCID: PMC2681094.

15. Bell SP, Dutta A. DNA replication in eukaryotic cells. *Annual review of biochemistry*. 2002;71:333-74. Epub 2002/06/05. doi: 10.1146/annurev.biochem.71.110601.135425. PubMed PMID: 12045100.

16. Sonenberg N, Hinnebusch AG. Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets. *Cell*. 2009;136(4):731-45. doi: 10.1016/j.cell.2009.01.042. PubMed PMID: 19239892; PubMed Central PMCID: PMC2681094.

17. Lee TI, Young RA. Transcription of eukaryotic protein-coding genes. *Annual review of genetics*. 2000;34:77-137. Epub 2000/11/28. doi: 10.1146/annurev.genet.34.1.77. PubMed PMID: 11092823.

18. Palli D, Saieva C, Grechi D, Masala G, Zanna I, Barbaro A, et al. DNA bulky adducts in a Mediterranean population correlate with environmental ozone concentration, an indicator of photochemical smog. *International journal of cancer*. 2004;109(1):17-23. Epub 2004/01/22. doi: 10.1002/ijc.11687. PubMed PMID: 14735463.

19. Topinka J, Milcova A, Libalova H, Novakova Z, Rossner P, Jr., Balascak I, et al. Biomarkers of exposure to tobacco smoke and environmental pollutants in mothers and

their transplacental transfer to the fetus. Part I: bulky DNA adducts. *Mutation research*. 2009;669(1-2):13-9. Epub 2009/05/13. doi: 10.1016/j.mrfmmm.2009.04.011. PubMed PMID: 19433098.

20. Lindahl T, Barnes DE. Repair of endogenous DNA damage. *Cold Spring Harbor symposia on quantitative biology*. 2000;65:127-33. Epub 2003/05/23. PubMed PMID: 12760027.

21. Ballmaier D, Epe B. DNA damage by bromate: mechanism and consequences. *Toxicology*. 2006;221(2-3):166-71. Epub 2006/02/24. doi: 10.1016/j.tox.2006.01.009. PubMed PMID: 16490296.

22. Valdiglesias V, Prego-Faraldo MV, Pásaro E, Méndez J, Laffon B. Okadaic Acid: More than a Diarrheic Toxin. *Marine Drugs*. 2013;11(11):4328-49. doi: 10.3390/md11114328. PubMed PMID: 24184795; PubMed Central PMCID: PMCPMC3853731.

23. Nishimura H, Yasui H, Sakurai H. Generation and distribution of reactive oxygen species in the skin of hairless mice under UVA: studies on in vivo chemiluminescent detection and tape stripping methods. *Experimental dermatology*. 2006;15(11):891-9. Epub 2006/09/28. doi: 10.1111/j.1600-0625.2006.00484.x. PubMed PMID: 17002686.

24. Lomax ME, Folkes LK, O'Neill P. Biological consequences of radiation-induced DNA damage: relevance to radiotherapy. *Clinical oncology (Royal College of Radiologists (Great Britain))*. 2013;25(10):578-85. Epub 2013/07/16. doi: 10.1016/j.clon.2013.06.007. PubMed PMID: 23849504.

25. De Bont R, van Larebeke N. Endogenous DNA damage in humans: a review of quantitative data. *Mutagenesis*. 2004;19(3):169-85. Epub 2004/05/05. PubMed PMID: 15123782.

26. Lindahl T, Nyberg B. Rate of depurination of native deoxyribonucleic acid. *Biochemistry*. 1972;11(19):3610-8. Epub 1972/09/12. PubMed PMID: 4626532.

27. Robertson AB, Klungland A, Rognes T, Leiros I. DNA repair in mammalian cells: Base excision repair: the long and short of it. *Cellular and molecular life sciences : CMLS*. 2009;66(6):981-93. Epub 2009/01/21. doi: 10.1007/s00018-009-8736-z. PubMed PMID: 19153658.

28. David SS, O'Shea VL, Kundu S. Base-excision repair of oxidative DNA damage. *Nature*. 2007;447(7147):941-50. Epub 2007/06/22. doi: 10.1038/nature05978. PubMed PMID: 17581577; PubMed Central PMCID: PMCPMC2896554.

29. Visnes T, Akbari M, Hagen L, Slupphaug G, Krokan HE. The rate of base excision repair of uracil is controlled by the initiating glycosylase. *DNA repair*.

2008;7(11):1869-81. Epub 2008/08/30. doi: 10.1016/j.dnarep.2008.07.012. PubMed PMID: 18721906.

30. Sung JS, Dimple B. Roles of base excision repair subpathways in correcting oxidized abasic sites in DNA. *The FEBS journal*. 2006;273(8):1620-9. Epub 2006/04/21. doi: 10.1111/j.1742-4658.2006.05192.x. PubMed PMID: 16623699.

31. McCulloch SD, Gu L, Li GM. Bi-directional processing of DNA loops by mismatch repair-dependent and -independent pathways in human cells. *The Journal of biological chemistry*. 2003;278(6):3891-6. Epub 2002/11/30. doi: 10.1074/jbc.M210687200. PubMed PMID: 12458199.

32. Jiricny J. The multifaceted mismatch-repair system. *Nature reviews Molecular cell biology*. 2006;7(5):335-46. Epub 2006/04/14. doi: 10.1038/nrm1907. PubMed PMID: 16612326.

33. San Filippo J, Sung P, Klein H. Mechanism of eukaryotic homologous recombination. *Annual review of biochemistry*. 2008;77:229-57. Epub 2008/02/16. doi: 10.1146/annurev.biochem.77.061306.125255. PubMed PMID: 18275380.

34. Chang HHY, Pannunzio NR, Adachi N, Lieber MR. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nature reviews Molecular cell biology*. 2017;18(8):495-506. Epub 2017/05/18. doi: 10.1038/nrm.2017.48. PubMed PMID: 28512351.

35. Kirkpatrick DT, Petes TD. Repair of DNA loops involves DNA-mismatch and nucleotide-excision repair proteins. *Nature*. 1997;387(6636):929-31. Epub 1997/06/26. doi: 10.1038/43225. PubMed PMID: 9202128.

36. Mouw KW, D'Andrea AD. Crosstalk between the nucleotide excision repair and Fanconi anemia/BRCA pathways. *DNA repair*. 2014;19:130-4. Epub 2014/04/29. doi: 10.1016/j.dnarep.2014.03.019. PubMed PMID: 24768451.

37. Ghosal G, Chen J. DNA damage tolerance: a double-edged sword guarding the genome. *Translational cancer research*. 2013;2(3):107-29. doi: 10.3978/j.issn.2218-676X.2013.04.01. PubMed PMID: 24058901; PubMed Central PMCID: PMC3779140.

38. Neeley WL, Essigmann JM. Mechanisms of formation, genotoxicity, and mutation of guanine oxidation products. *Chemical research in toxicology*. 2006;19(4):491-505. Epub 2006/04/13. doi: 10.1021/tx0600043. PubMed PMID: 16608160.

39. Gedik CM, Collins A. Establishing the background level of base oxidation in human lymphocyte DNA: results of an interlaboratory validation study. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*.

2005;19(1):82-4. Epub 2004/11/10. doi: 10.1096/fj.04-1767fje. PubMed PMID: 15533950.

40. Measurement of DNA oxidation in human cells by chromatographic and enzymic methods. *Free radical biology & medicine*. 2003;34(8):1089-99. Epub 2003/04/10. PubMed PMID: 12684094.

41. Steenken S, Jovanovic SV. How easily oxidizable is DNA? One-electron reduction potentials of adenosine and guanosine radicals in aqueous solution. *Journal of the American Chemical Society*. 1997;119(3):617-8. doi: 10.1021/ja962255b. PubMed PMID: WOS:A1997WD86700024.

42. Cheng KC, Cahill DS, Kasai H, Nishimura S, Loeb LA. 8-Hydroxyguanine, an abundant form of oxidative DNA damage, causes G----T and A----C substitutions. *The Journal of biological chemistry*. 1992;267(1):166-72. Epub 1992/01/05. PubMed PMID: 1730583.

43. Lyons DM, O'Brien PJ. Human base excision repair creates a bias toward -1 frameshift mutations. *The Journal of biological chemistry*. 2010;285(33):25203-12. Epub 2010/06/16. doi: 10.1074/jbc.M110.118596. PubMed PMID: 20547483; PubMed Central PMCID: PMC2919083.

44. Boiteux S, Guillet M. Abasic sites in DNA: repair and biological consequences in *Saccharomyces cerevisiae*. *DNA repair*. 2004;3(1):1-12. Epub 2003/12/31. PubMed PMID: 14697754.

45. Iyama T, Wilson DM. DNA repair mechanisms in dividing and non-dividing cells. *DNA repair*. 2013;12(8):620-36. doi: 10.1016/j.dnarep.2013.04.015. PubMed PMID: 23684800; PubMed Central PMCID: PMC3720834.

46. Tubbs A, Nussenzweig A. Endogenous DNA Damage as a Source of Genomic Instability in Cancer. *Cell*. 2017;168(4):644-56. Epub 2017/02/12. doi: 10.1016/j.cell.2017.01.002. PubMed PMID: 28187286.

47. Nelson SR, Dunn AR, Kathe SD, Warshaw DM, Wallace SS. Two glycosylase families diffusively scan DNA using a wedge residue to probe for and identify oxidatively damaged bases. *Proceedings of the National Academy of Sciences of the United States of America*. 2014;111(20):E2091-9. Epub 2014/05/07. doi: 10.1073/pnas.1400386111. PubMed PMID: 24799677; PubMed Central PMCID: PMC4034194.

48. Dyrkheeva NS, Lebedeva NA, Lavrik OI. AP Endonuclease 1 as a Key Enzyme in Repair of Apurinic/Apyrimidinic Sites. *Biochemistry Biokhimiia*. 2016;81(9):951-67. Epub 2016/09/30. doi: 10.1134/s0006297916090042. PubMed PMID: 27682167.

49. Beard WA, Wilson SH. Structure and mechanism of DNA polymerase beta. *Biochemistry*. 2014;53(17):2768-80. Epub 2014/04/11. doi: 10.1021/bi500139h. PubMed PMID: 24717170; PubMed Central PMCID: PMC4018062.
50. Prasad R, Shock DD, Beard WA, Wilson SH. Substrate Channeling in Mammalian Base Excision Repair Pathways: Passing the Baton. *The Journal of biological chemistry*. 2010;285(52):40479-88. doi: 10.1074/jbc.M110.155267. PubMed PMID: 20952393; PubMed Central PMCID: PMC4003346.
51. Liu Y, Beard WA, Shock DD, Prasad R, Hou EW, Wilson SH. DNA polymerase beta and flap endonuclease 1 enzymatic specificities sustain DNA synthesis for long patch base excision repair. *The Journal of biological chemistry*. 2005;280(5):3665-74. Epub 2004/11/25. doi: 10.1074/jbc.M412922200. PubMed PMID: 15561706.
52. Krokan HE, Bjørås M. Base Excision Repair. *Cold Spring Harbor Perspectives in Biology*. 2013;5(4). doi: 10.1101/cshperspect.a012583. PubMed PMID: 23545420; PubMed Central PMCID: PMC4038398.
53. Tell G, Quadrioglio F, Tiribelli C, Kelley MR. The Many Functions of APE1/Ref-1: Not Only a DNA Repair Enzyme. *Antioxidants & Redox Signaling*. 2009;11(3):601-19. doi: 10.1089/ars.2008.2194. PubMed PMID: 18976116; PubMed Central PMCID: PMC4011080.
54. Sakumi K, Tominaga Y, Furuichi M, Xu P, Tsuzuki T, Sekiguchi M, et al. Ogg1 knockout-associated lung tumorigenesis and its suppression by Mth1 gene disruption. *Cancer research*. 2003;63(5):902-5. Epub 2003/03/05. PubMed PMID: 12615700.
55. Sugo N, Aratani Y, Nagashima Y, Kubota Y, Koyama H. Neonatal lethality with abnormal neurogenesis in mice deficient in DNA polymerase β . *The EMBO Journal*. 2000;19(6):1397-404. doi: 10.1093/emboj/19.6.1397. PubMed PMID: 10716939; PubMed Central PMCID: PMC4035680.
56. Horton JK, Joyce-Gray DF, Pachkowski BF, Swenberg JA, Wilson SH. Hypersensitivity of DNA polymerase beta null mouse fibroblasts reflects accumulation of cytotoxic repair intermediates from site-specific alkyl DNA lesions. *DNA repair*. 2003;2(1):27-48. Epub 2003/01/02. PubMed PMID: 12509266.
57. Larsen E, Gran C, Sæther BE, Seeberg E, Klungland A. Proliferation Failure and Gamma Radiation Sensitivity of Fen1 Null Mutant Mice at the Blastocyst Stage. *Molecular and cellular biology*. 2003;23(15):5346-53. doi: 10.1128/mcb.23.15.5346-5353.2003. PubMed PMID: 12861020; PubMed Central PMCID: PMC40165721.
58. Zhang SH, Wang LA, Li Z, Peng Y, Cun YP, Dai N, et al. APE1 polymorphisms are associated with colorectal cancer susceptibility in Chinese Hans. *World Journal of Gastroenterology : WJG*. 2014;20(26):8700-8. doi: 10.3748/wjg.v20.i26.8700. PubMed PMID: 25024628; PubMed Central PMCID: PMC4093723.

59. Rossner P, Jr., Terry MB, Gammon MD, Zhang FF, Teitelbaum SL, Eng SM, et al. OGG1 polymorphisms and breast cancer risk. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2006;15(4):811-5. Epub 2006/04/15. doi: 10.1158/1055-9965.Epi-05-0659. PubMed PMID: 16614128.
60. Khanra K, Bhattacharya C, Bhattacharyya N. Association of a newly identified variant of DNA polymerase beta (polbetaDelta63-123, 208-304) with the risk factor of ovarian carcinoma in India. *Asian Pacific journal of cancer prevention : APJCP*. 2012;13(5):1999-2002. Epub 2012/08/21. PubMed PMID: 22901161.
61. Guo Z, Zheng L, Dai H, Zhou M, Xu H, Shen B. Human DNA polymerase beta polymorphism, Arg137Gln, impairs its polymerase activity and interaction with PCNA and the cellular base excision repair capacity. *Nucleic acids research*. 2009;37(10):3431-41. Epub 2009/04/02. doi: 10.1093/nar/gkp201. PubMed PMID: 19336415; PubMed Central PMCID: PMC2691839.
62. Pan F, Zhao J, Zhou T, Kuang Z, Dai H, Wu H, et al. Mutation of DNA Polymerase beta R137Q Results in Retarded Embryo Development Due to Impaired DNA Base Excision Repair in Mice. *Scientific reports*. 2016;6:28614. Epub 2016/07/01. doi: 10.1038/srep28614. PubMed PMID: 27358192; PubMed Central PMCID: PMC4928080.
63. Starcevic D, Dalal S, Sweasy JB. Is there a link between DNA polymerase beta and cancer? *Cell cycle (Georgetown, Tex)*. 2004;3(8):998-1001. Epub 2004/07/29. PubMed PMID: 15280658.
64. Dalal S, Chikova A, Jaeger J, Sweasy JB. The Leu22Pro tumor-associated variant of DNA polymerase beta is dRP lyase deficient. *Nucleic acids research*. 2008;36(2):411-22. Epub 2007/11/28. doi: 10.1093/nar/gkm1053. PubMed PMID: 18039710; PubMed Central PMCID: PMC2241857.
65. Eckenroth BE, Towle-Weicksel JB, Sweasy JB, Doubie S. The E295K cancer variant of human polymerase beta favors the mismatch conformational pathway during nucleotide selection. *The Journal of biological chemistry*. 2013;288(48):34850-60. Epub 2013/10/18. doi: 10.1074/jbc.M113.510891. PubMed PMID: 24133209; PubMed Central PMCID: PMC3843097.
66. Lang T, Dalal S, Chikova A, DiMaio D, Sweasy JB. The E295K DNA polymerase beta gastric cancer-associated variant interferes with base excision repair and induces cellular transformation. *Molecular and cellular biology*. 2007;27(15):5587-96. Epub 2007/05/29. doi: 10.1128/mcb.01883-06. PubMed PMID: 17526740; PubMed Central PMCID: PMC1952088.
67. Sun H, He L, Wu H, Pan F, Wu X, Zhao J, et al. The FEN1 L209P mutation interferes with long-patch base excision repair and induces cellular transformation.

Oncogene. 2017;36(2):194-207. Epub 2016/06/09. doi: 10.1038/onc.2016.188. PubMed PMID: 27270424; PubMed Central PMCID: PMC5140775.

68. London RE. The structural basis of XRCC1-mediated DNA repair. DNA repair. 2015;30:90-103. doi: 10.1016/j.dnarep.2015.02.005. PubMed PMID: 25795425; PubMed Central PMCID: PMC5580684.

69. Ko R, Bennett SE. Physical and functional interaction of human nuclear uracil-DNA glycosylase with proliferating cell nuclear antigen. DNA repair. 2005;4(12):1421-31. Epub 2005/10/12. doi: 10.1016/j.dnarep.2005.08.006. PubMed PMID: 16216562; PubMed Central PMCID: PMC5340124.

70. Marsin S, Vidal AE, Sossou M, Menissier-de Murcia J, Le Page F, Boiteux S, et al. Role of XRCC1 in the coordination and stimulation of oxidative DNA damage repair initiated by the DNA glycosylase hOGG1. The Journal of biological chemistry. 2003;278(45):44068-74. Epub 2003/08/23. doi: 10.1074/jbc.M306160200. PubMed PMID: 12933815.

71. Hegde ML, Hegde PM, Arijit D, Boldogh I, Mitra S. Human DNA Glycosylase NEIL1's Interactions with Downstream Repair Proteins Is Critical for Efficient Repair of Oxidized DNA Base Damage and Enhanced Cell Survival. Biomolecules. 2012;2(4):564-78. Epub 2013/08/09. doi: 10.3390/biom2040564. PubMed PMID: 23926464; PubMed Central PMCID: PMC3733129.

72. Vidal AE. XRCC1 coordinates the initial and late stages of DNA abasic site repair through protein-protein interactions. 2001;20(22):6530-9. doi: 10.1093/emboj/20.22.6530. PubMed PMID: 11707423; PubMed Central PMCID: PMC125722.

73. Dianova II, Sleeth KM, Allinson SL, Parsons JL, Breslin C, Caldecott KW, et al. XRCC1-DNA polymerase β interaction is required for efficient base excision repair. Nucleic acids research. 2004;32(8):2550-5. doi: 10.1093/nar/gkh567. PubMed PMID: 15141024; PubMed Central PMCID: PMC149455.

74. Cappelli E, Taylor R, Cevasco M, Abbondandolo A, Caldecott K, Frosina G. Involvement of XRCC1 and DNA ligase III gene products in DNA base excision repair. The Journal of biological chemistry. 1997;272(38):23970-5. Epub 1997/09/20. PubMed PMID: 9295348.

75. Petermann E, Keil C, Oei SL. Roles of DNA ligase III and XRCC1 in regulating the switch between short patch and long patch BER. DNA repair. 2006;5(5):544-55. Epub 2006/01/31. doi: 10.1016/j.dnarep.2005.12.008. PubMed PMID: 16442856.

76. Matsumoto Y. Molecular mechanism of PCNA-dependent base excision repair. Progress in nucleic acid research and molecular biology. 2001;68:129-38. Epub 2001/09/14. PubMed PMID: 11554292.

77. Moldovan GL, Pfander B, Jentsch S. PCNA, the maestro of the replication fork. *Cell*. 2007;129(4):665-79. Epub 2007/05/22. doi: 10.1016/j.cell.2007.05.003. PubMed PMID: 17512402.
78. Mailand N, Gibbs-Seymour I, Bekker-Jensen S. Regulation of PCNA-protein interactions for genome stability. *Nature reviews Molecular cell biology*. 2013;14(5):269-82. Epub 2013/04/19. doi: 10.1038/nrm3562. PubMed PMID: 23594953.
79. Xia L, Zheng L, Lee HW, Bates SE, Federico L, Shen B, et al. Human 3-methyladenine-DNA glycosylase: effect of sequence context on excision, association with PCNA, and stimulation by AP endonuclease. *Journal of molecular biology*. 2005;346(5):1259-74. Epub 2005/02/17. doi: 10.1016/j.jmb.2005.01.014. PubMed PMID: 15713479.
80. Sakurai S, Kitano K, Yamaguchi H, Hamada K, Okada K, Fukuda K, et al. Structural basis for recruitment of human flap endonuclease 1 to PCNA. *The EMBO Journal*. 2005;24(4):683-93. doi: 10.1038/sj.emboj.7600519. PubMed PMID: 15616578; PubMed Central PMCID: PMC549611.
81. Mayanagi K, Kiyonari S, Saito M, Shirai T, Ishino Y, Morikawa K. Mechanism of replication machinery assembly as revealed by the DNA ligase-PCNA-DNA complex architecture. *Proceedings of the National Academy of Sciences of the United States of America*. 2009;106(12):4647-52. Epub 2009/03/04. doi: 10.1073/pnas.0811196106. PubMed PMID: 19255439; PubMed Central PMCID: PMC2660782.
82. Stodola JL, Burgers PM. Resolving individual steps of Okazaki-fragment maturation at a millisecond timescale. *Nature structural & molecular biology*. 2016;23(5):402-8. Epub 2016/04/12. doi: 10.1038/nsmb.3207. PubMed PMID: 27065195; PubMed Central PMCID: PMC4857878.
83. Dovrat D, Stodola JL, Burgers PM, Aharoni A. Sequential switching of binding partners on PCNA during in vitro Okazaki fragment maturation. *Proceedings of the National Academy of Sciences of the United States of America*. 2014;111(39):14118-23. Epub 2014/09/18. doi: 10.1073/pnas.1321349111. PubMed PMID: 25228764; PubMed Central PMCID: PMC4191785.
84. Levin DS, McKenna AE, Motycka TA, Matsumoto Y, Tomkinson AE. Interaction between PCNA and DNA ligase I is critical for joining of Okazaki fragments and long-patch base-excision repair. *Current biology : CB*. 2000;10(15):919-22. Epub 2000/08/26. PubMed PMID: 10959839.
85. Zheng L, Dai H, Hegde ML, Zhou M, Guo Z, Wu X, et al. Fen1 mutations that specifically disrupt its interaction with PCNA cause aneuploidy-associated cancer. *Cell Research*. 2011;21(7):1052-67. doi: 10.1038/cr.2011.35. PubMed PMID: 21383776; PubMed Central PMCID: PMC3129403.

86. Wallace SS, Murphy DL, Sweasy JB. Base excision repair and cancer. *Cancer letters*. 2012;327(1-2):73-89. Epub 2012/01/19. doi: 10.1016/j.canlet.2011.12.038. PubMed PMID: 22252118; PubMed Central PMCID: PMC3361536.
87. Senejani AG, Liu Y, Kidane D, Maher SE, Zeiss CJ, Park HJ, et al. Mutation of POL B Causes Lupus in Mice. *Cell reports*. 2014;6(1):1-8. doi: 10.1016/j.celrep.2013.12.017. PubMed PMID: 24388753; PubMed Central PMCID: PMC3916967.
88. Pearson CE, Nichol Edamura K, Cleary JD. Repeat instability: mechanisms of dynamic mutations. *Nature reviews Genetics*. 2005;6(10):729-42. Epub 2005/10/06. doi: 10.1038/nrg1689. PubMed PMID: 16205713.
89. Paulson HL, Fischbeck KH. Trinucleotide repeats in neurogenetic disorders. *Annual review of neuroscience*. 1996;19:79-107. Epub 1996/01/01. doi: 10.1146/annurev.ne.19.030196.000455. PubMed PMID: 8833437.
90. Budworth H, McMurray CT. A Brief History of Triplet Repeat Diseases. *Methods in molecular biology (Clifton, NJ)*. 2013;1010:3-17. doi: 10.1007/978-1-62703-411-1_1. PubMed PMID: 23754215; PubMed Central PMCID: PMC3913379.
91. Ellegren H. Microsatellites: simple sequences with complex evolution. *Nature reviews Genetics*. 2004;5(6):435-45. Epub 2004/05/22. doi: 10.1038/nrg1348. PubMed PMID: 15153996.
92. McMurray CT. Mechanisms of trinucleotide repeat instability during human development. *Nature reviews Genetics*. 2010;11(11):786-99. Epub 2010/10/19. doi: 10.1038/nrg2828. PubMed PMID: 20953213; PubMed Central PMCID: PMC3175376.
93. Nelson KA, Witte JS. Androgen receptor CAG repeats and prostate cancer. *American journal of epidemiology*. 2002;155(10):883-90. Epub 2002/05/08. PubMed PMID: 11994226.
94. Lee DY, McMurray CT. Trinucleotide expansion in disease: why is there a length threshold? *Current opinion in genetics & development*. 2014;0:131-40. doi: 10.1016/j.gde.2014.07.003. PubMed PMID: 25282113; PubMed Central PMCID: PMC34252851.
95. Luo Z, Gao X, Lin C, Smith ER, Marshall SA, Swanson SK, et al. Zic2 is an enhancer-binding factor required for embryonic stem cell specification. *Molecular Cell*. 2015;57(4):685-94. Epub 2015/02/24. doi: 10.1016/j.molcel.2015.01.007. PubMed PMID: 25699711; PubMed Central PMCID: PMC35009765.
96. Watabe Y, Baba Y, Nakauchi H, Mizota A, Watanabe S. The role of Zic family zinc finger transcription factors in the proliferation and differentiation of retinal

- progenitor cells. *Biochemical and biophysical research communications*. 2011;415(1):42-7. Epub 2011/10/26. doi: 10.1016/j.bbrc.2011.10.007. PubMed PMID: 22024047.
97. Messaied C, Rouleau GA. Molecular mechanisms underlying polyalanine diseases. *Neurobiology of disease*. 2009;34(3):397-405. Epub 2009/03/10. doi: 10.1016/j.nbd.2009.02.013. PubMed PMID: 19269323.
 98. Lazaro L, Dubourg C, Pasquier L, Le Duff F, Blayau M, Durou MR, et al. Phenotypic and molecular variability of the holoprosencephalic spectrum. *American journal of medical genetics Part A*. 2004;129a(1):21-4. Epub 2004/07/22. doi: 10.1002/ajmg.a.30110. PubMed PMID: 15266610.
 99. Fan HC, Ho LI, Chi CS, Chen SJ, Peng GS, Chan TM, et al. Polyglutamine (PolyQ) diseases: genetics to treatments. *Cell transplantation*. 2014;23(4-5):441-58. Epub 2014/05/13. doi: 10.3727/096368914x678454. PubMed PMID: 24816443.
 100. Takahashi T, Katada S, Onodera O. Polyglutamine diseases: where does toxicity come from? what is toxicity? where are we going? *Journal of molecular cell biology*. 2010;2(4):180-91. Epub 2010/04/23. doi: 10.1093/jmcb/mjq005. PubMed PMID: 20410236.
 101. Polling S, Ormsby AR, Wood RJ, Lee K, Shoubbridge C, Hughes JN, et al. Polyalanine expansions drive a shift into alpha-helical clusters without amyloid-fibril formation. *Nature structural & molecular biology*. 2015;22(12):1008-15. Epub 2015/11/17. doi: 10.1038/nsmb.3127. PubMed PMID: 26571108.
 102. Nasrallah IM, Minarcik JC, Golden JA. A polyalanine tract expansion in Arx forms intranuclear inclusions and results in increased cell death. *The Journal of Cell Biology*. 2004;167(3):411-6. doi: 10.1083/jcb.200408091. PubMed PMID: 15533998; PubMed Central PMCID: PMC2172475.
 103. de Cristofaro T, Affaitati A, Feliciello A, Avvedimento EV, Varrone S. Polyglutamine-mediated aggregation and cell death. *Biochemical and biophysical research communications*. 2000;272(3):816-21. Epub 2000/06/22. doi: 10.1006/bbrc.2000.2843. PubMed PMID: 10860836.
 104. Dorsman JC, Pepers B, Langenberg D, Kerkdijk H, Ijszenga M, den Dunnen JT, et al. Strong aggregation and increased toxicity of polyleucine over polyglutamine stretches in mammalian cells. *Human molecular genetics*. 2002;11(13):1487-96. Epub 2002/06/05. PubMed PMID: 12045202.
 105. Albrecht A, Mundlos S. The other trinucleotide repeat: polyalanine expansion disorders. *Current opinion in genetics & development*. 2005;15(3):285-93. Epub 2005/05/27. doi: 10.1016/j.gde.2005.04.003. PubMed PMID: 15917204.

106. Botta A, Rinaldi F, Catalli C, Vergani L, Bonifazi E, Romeo V, et al. The CTG repeat expansion size correlates with the splicing defects observed in muscles from myotonic dystrophy type 1 patients. *Journal of medical genetics*. 2008;45(10):639-46. Epub 2008/07/10. doi: 10.1136/jmg.2008.058909. PubMed PMID: 18611984.
107. Thornton CA. Myotonic Dystrophy. *Neurologic clinics*. 2014;32(3):705-19. doi: 10.1016/j.ncl.2014.04.011. PubMed PMID: 25037086; PubMed Central PMCID: PMC4105852.
108. Patel PI, Isaya G. Friedreich Ataxia: From GAA Triplet–Repeat Expansion to Frataxin Deficiency. *American Journal of Human Genetics*. 2001;69(1):15-24. PubMed PMID: 11391483; PubMed Central PMCID: PMC41226030.
109. Pastore A, Puccio H. Frataxin: a protein in search for a function. *Journal of neurochemistry*. 2013;126 Suppl 1:43-52. Epub 2013/07/24. doi: 10.1111/jnc.12220. PubMed PMID: 23859340.
110. Silva AM, Brown JM, Buckle VJ, Wade-Martins R, Lufino MMP. Expanded GAA repeats impair FXN gene expression and reposition the FXN locus to the nuclear lamina in single cells. *Human molecular genetics*. 2015;24(12):3457-71. doi: 10.1093/hmg/ddv096. PubMed PMID: 25814655; PubMed Central PMCID: PMC4498154.
111. Bagni C, Tassone F, Neri G, Hagerman R. Fragile X syndrome: causes, diagnosis, mechanisms, and therapeutics. *The Journal of clinical investigation*. 2012;122(12):4314-22. Epub 2012/12/04. doi: 10.1172/jci63141. PubMed PMID: 23202739; PubMed Central PMCID: PMC3533539.
112. Martí E, Estivill X. Small non-coding RNAs add complexity to the RNA pathogenic mechanisms in trinucleotide repeat expansion diseases. *Frontiers in Molecular Neuroscience*. 2013;6. doi: 10.3389/fnmol.2013.00045. PubMed PMID: 24348326; PubMed Central PMCID: PMC3848198.
113. Bacolla A, Wells RD. Non-B DNA conformations as determinants of mutagenesis and human disease. *Molecular carcinogenesis*. 2009;48(4):273-85. Epub 2009/03/24. doi: 10.1002/mc.20507. PubMed PMID: 19306308.
114. Watson JD, Crick FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*. 1953;171(4356):737-8. Epub 1953/04/25. PubMed PMID: 13054692.
115. Moon J, Han JH, Kim DY, Jung MJ, Kim SK. Effects of deficient of the Hoogsteen base-pairs on the G-quadruplex stabilization and binding mode of a cationic porphyrin. *Biochemistry and biophysics reports*. 2015;2:29-35. Epub 2015/04/08. doi: 10.1016/j.bbrep.2015.03.012. PubMed PMID: 29124143; PubMed Central PMCID: PMC45668627.

116. Mitas M. Trinucleotide repeats associated with human disease. *Nucleic acids research*. 1997;25(12):2245-54. Epub 1997/06/15. PubMed PMID: 9171073; PubMed Central PMCID: PMCPMC146772.
117. Usdin K, Woodford KJ. CGG repeats associated with DNA instability and chromosome fragility form structures that block DNA synthesis in vitro. *Nucleic acids research*. 1995;23(20):4202-9. Epub 1995/10/25. PubMed PMID: 7479085; PubMed Central PMCID: PMCPMC307363.
118. Kang S, Jaworski A, Ohshima K, Wells RD. Expansion and deletion of CTG repeats from human disease genes are determined by the direction of replication in *E. coli*. *Nature genetics*. 1995;10(2):213-8. Epub 1995/06/01. doi: 10.1038/ng0695-213. PubMed PMID: 7663518.
119. Ohshima K, Wells RD. Hairpin formation during DNA synthesis primer realignment in vitro in triplet repeat sequences from human hereditary disease genes. *The Journal of biological chemistry*. 1997;272(27):16798-806. Epub 1997/07/04. PubMed PMID: 9201985.
120. Pelletier R, Krasilnikova MM, Samadashwily GM, Lahue R, Mirkin SM. Replication and Expansion of Trinucleotide Repeats in Yeast. *Molecular and cellular biology*. 2003;23(4):1349-57. doi: 10.1128/mcb.23.4.1349-1357.2003. PubMed PMID: 12556494; PubMed Central PMCID: PMCPMC141142.
121. Cleary JD, Nichol K, Wang YH, Pearson CE. Evidence of cis-acting factors in replication-mediated trinucleotide repeat instability in primate cells. *Nature genetics*. 2002;31(1):37-46. Epub 2002/04/23. doi: 10.1038/ng870. PubMed PMID: 11967533.
122. Goula AV, Berquist BR, Wilson DM, 3rd, Wheeler VC, Trottier Y, Merienne K. Stoichiometry of base excision repair proteins correlates with increased somatic CAG instability in striatum over cerebellum in Huntington's disease transgenic mice. *PLoS genetics*. 2009;5(12):e1000749. Epub 2009/12/10. doi: 10.1371/journal.pgen.1000749. PubMed PMID: 19997493; PubMed Central PMCID: PMCPMC2778875.
123. Jonson I, Ougland R, Klungland A, Larsen E. Oxidative stress causes DNA triplet expansion in Huntington's disease mouse embryonic stem cells. *Stem cell research*. 2013;11(3):1264-71. Epub 2013/09/18. doi: 10.1016/j.scr.2013.08.010. PubMed PMID: 24041806.
124. Gerhardt J, Bhalla A, Butler JS, Puckett J, Dervan PB, Rosenwaks Z, et al. Stalled DNA replication forks at the endogenous GAA repeats drive repeat expansion in Friedreich's ataxia cells. *Cell reports*. 2016;16(5):1218-27. doi: 10.1016/j.celrep.2016.06.075. PubMed PMID: 27425605; PubMed Central PMCID: PMCPMC5028224.

125. Samadashwily GM, Raca G, Mirkin SM. Trinucleotide repeats affect DNA replication in vivo. *Nature genetics*. 1997;17(3):298-304. Epub 1997/11/14. doi: 10.1038/ng1197-298. PubMed PMID: 9354793.
126. Krasilnikova MM, Mirkin SM. Replication stalling at Friedreich's ataxia (GAA)_n repeats in vivo. *Molecular and cellular biology*. 2004;24(6):2286-95. Epub 2004/03/03. PubMed PMID: 14993268; PubMed Central PMCID: PMCPMC355872.
127. Fouché N, Özgür S, Roy D, Griffith JD. Replication fork regression in repetitive DNAs. *Nucleic acids research*. 2006;34(20):6044-50. doi: 10.1093/nar/gkl757. PubMed PMID: 17071963; PubMed Central PMCID: PMCPMC1635326.
128. Lipe H, Bird T. Late Onset Huntington Disease: Clinical and Genetic Characteristics of 34 Cases. *Journal of the neurological sciences*. 2009;276(1-2):159-62. doi: 10.1016/j.jns.2008.09.029. PubMed PMID: 18977004; PubMed Central PMCID: PMCPMC3140172.
129. Koutsis G, Karadima G, Kladi A, Panas M. Late-onset Huntington's disease: diagnostic and prognostic considerations. *Parkinsonism & related disorders*. 2014;20(7):726-30. Epub 2014/04/12. doi: 10.1016/j.parkreldis.2014.03.017. PubMed PMID: 24721491.
130. Koeppen AH, Morral JA, McComb RD, Feustel PJ. The Neuropathology of Late-Onset Friedreich's Ataxia. *Cerebellum (London, England)*. 2011;10(1):96-103. doi: 10.1007/s12311-010-0235-0. PubMed PMID: 21128039; PubMed Central PMCID: PMCPMC4894659.
131. Cadet J, Wagner JR, Shafirovich V, Geacintov NE. One-electron oxidation reactions of purine and pyrimidine bases in cellular DNA. *International journal of radiation biology*. 2014;90(6):423-32. doi: 10.3109/09553002.2013.877176. PubMed PMID: 24369822; PubMed Central PMCID: PMCPMC4047668.
132. Bidichandani SI, Garcia CA, Patel PI, Dimachkie MM. Very late-onset Friedreich ataxia despite large GAA triplet repeat expansions. *Archives of neurology*. 2000;57(2):246-51. Epub 2000/02/19. PubMed PMID: 10681084.
133. Orth M, Schwenke C. Age-at-onset in Huntington disease. *PLoS Currents*. 3. doi: 10.1371/currents.RRN1258. PubMed PMID: 22453877; PubMed Central PMCID: PMCPMC3164295.
134. Hargreaves VV, Shell SS, Mazur DJ, Hess MT, Kolodner RD. Interaction between the Msh2 and Msh6 nucleotide-binding sites in the *Saccharomyces cerevisiae* Msh2-Msh6 complex. *The Journal of biological chemistry*. 2010;285(12):9301-10. Epub 2010/01/22. doi: 10.1074/jbc.M109.096388. PubMed PMID: 20089866; PubMed Central PMCID: PMCPMC2838348.

135. Owen BA, Yang Z, Lai M, Gajec M, Badger JD, 2nd, Hayes JJ, et al. (CAG)(n)-hairpin DNA binds to Msh2-Msh3 and changes properties of mismatch recognition. *Nature structural & molecular biology*. 2005;12(8):663-70. Epub 2005/07/19. doi: 10.1038/nsmb965. PubMed PMID: 16025128.
136. Hashem VI, Pytlos MJ, Klysik EA, Tsuji K. Chemotherapeutic deletion of CTG repeats in lymphoblast cells from DM1 patients. 2004;32(21):6334-46. doi: 10.1093/nar/gkh976. PubMed PMID: 15576360; PubMed Central PMCID: PMC535684.
137. Zhang Y, Monckton DG, Siciliano MJ, Connor TH, Meistrich ML. Detection of radiation and cyclophosphamide-induced mutations in individual mouse sperm at a human expanded trinucleotide repeat locus transgene. *Mutation research*. 2002;516(1-2):121-38. Epub 2002/04/12. PubMed PMID: 11943618.
138. Gomes-Pereira M, Monckton DG. Chemically induced increases and decreases in the rate of expansion of a CAG*CTG triplet repeat. *Nucleic acids research*. 2004;32(9):2865-72. Epub 2004/05/25. doi: 10.1093/nar/gkh612. PubMed PMID: 15155855; PubMed Central PMCID: PMC419613.
139. Hashem VI, Pytlos MJ, Klysik EA, Tsuji K, Khajavi M, Ashizawa T, et al. Chemotherapeutic deletion of CTG repeats in lymphoblast cells from DM1 patients. *Nucleic acids research*. 2004;32(21):6334-46. Epub 2004/12/04. doi: 10.1093/nar/gkh976. PubMed PMID: 15576360; PubMed Central PMCID: PMC535684.
140. Hashem VI, Sinden RR. Chemotherapeutically induced deletion of expanded triplet repeats. *Mutation research*. 2002;508(1-2):107-19. Epub 2002/10/16. PubMed PMID: 12379466.
141. Lai Y, Beaver JM, Lorente K, Melo J, Ramjagsingh S, AgoulNIK IU, et al. Base Excision Repair of Chemotherapeutically-Induced Alkylated DNA Damage Predominantly Causes Contractions of Expanded GAA Repeats Associated with Friedreich's Ataxia. *PLoS ONE*. 2014;9(4). doi: 10.1371/journal.pone.0093464. PubMed PMID: 24691413; PubMed Central PMCID: PMC3972099.
142. Xu M, Gabison J, Liu Y. Trinucleotide repeat deletion via a unique hairpin bypass by DNA polymerase β and alternate flap cleavage by flap endonuclease 1. *Nucleic acids research*. 2013;41(3):1684-97. doi: 10.1093/nar/gks1306. PubMed PMID: 23258707; PubMed Central PMCID: PMC3561997.
143. Xu M, Lai Y, Jiang Z, Terzidis MA, Masi A, Chatgililoglu C, et al. A 5', 8-cyclo-2'-deoxypurine lesion induces trinucleotide repeat deletion via a unique lesion bypass by DNA polymerase β . *Nucleic acids research*. 2014;42(22):13749-63. doi: 10.1093/nar/gku1239. PubMed PMID: 25428354; PubMed Central PMCID: PMC4267656.

144. Duffy CM, Hilbert BJ, Kelch BA. A Disease-Causing Variant in PCNA Disrupts a Promiscuous Protein Binding Site. *Journal of molecular biology*. 2016;428(6):1023-40. Epub 2015/12/22. doi: 10.1016/j.jmb.2015.11.029. PubMed PMID: 26688547.
145. Xu M, Lai Y, Torner J, Zhang Y, Zhang Z, Liu Y. Base excision repair of oxidative DNA damage coupled with removal of a CAG repeat hairpin attenuates trinucleotide repeat expansion. *Nucleic acids research*. 2014;42(6):3675-91. Epub 2014/01/16. doi: 10.1093/nar/gkt1372. PubMed PMID: 24423876; PubMed Central PMCID: PMC3973345.
146. Lokanga RA, Senejani AG, Sweasy JB, Usdin K. Heterozygosity for a Hypomorphic Pol β Mutation Reduces the Expansion Frequency in a Mouse Model of the Fragile X-Related Disorders. *PLoS genetics*. 2015;11(4). doi: 10.1371/journal.pgen.1005181. PubMed PMID: 25886163; PubMed Central PMCID: PMC4401650.
147. Lopez Castel A, Tomkinson AE, Pearson CE. CTG/CAG repeat instability is modulated by the levels of human DNA ligase I and its interaction with proliferating cell nuclear antigen: a distinction between replication and slipped-DNA repair. *The Journal of biological chemistry*. 2009;284(39):26631-45. Epub 2009/07/25. doi: 10.1074/jbc.M109.034405. PubMed PMID: 19628465; PubMed Central PMCID: PMC2785351.
148. Refsland EW, Livingston DM. 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*. 2005;171(3):923-34. Epub 2005/08/05. doi: 10.1534/genetics.105.043448. PubMed PMID: 16079237; PubMed Central PMCID: PMC1456850.
149. Beaver JM, Lai Y, Rolle SJ, Liu Y. Proliferating cell nuclear antigen prevents trinucleotide repeat expansions by promoting repeat deletion and hairpin removal. *DNA repair*. 2016;48:17-29. doi: 10.1016/j.dnarep.2016.10.006. PubMed PMID: 27793507; PubMed Central PMCID: PMC4533789.
150. Chatgililoglu C, D'Angelantonio M, Kciuk G, Bobrowski K. New insights into the reaction paths of hydroxyl radicals with 2'-deoxyguanosine. *Chemical research in toxicology*. 2011;24(12):2200-6. Epub 2011/09/24. doi: 10.1021/tx2003245. PubMed PMID: 21939266.
151. Senejani AG, Dalal S, Liu Y, Nottoli TP, McGrath JM, Clairmont CS, et al. Y265C DNA polymerase beta knockin mice survive past birth and accumulate base excision repair intermediate substrates. *Proceedings of the National Academy of Sciences of the United States of America*. 2012;109(17):6632-7. Epub 2012/04/12. doi: 10.1073/pnas.1200800109. PubMed PMID: 22493258; PubMed Central PMCID: PMC3340078.

152. Jarem DA, Wilson NR, Schermerhorn KM, Delaney S. 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*. 2011;10(8):887-96. Epub 2011/07/06. doi: 10.1016/j.dnarep.2011.06.003. PubMed PMID: 21727036; PubMed Central PMCID: PMC3146575.
153. Jarem DA, Wilson NR, Delaney S. Structure-dependent DNA damage and repair in a trinucleotide repeat sequence. *Biochemistry*. 2009;48(28):6655-63. Epub 2009/06/17. doi: 10.1021/bi9007403. PubMed PMID: 19527055.
154. Dou H, Mitra S, Hazra TK. Repair of oxidized bases in DNA bubble structures by human DNA glycosylases NEIL1 and NEIL2. *The Journal of biological chemistry*. 2003;278(50):49679-84. Epub 2003/10/03. doi: 10.1074/jbc.M308658200. PubMed PMID: 14522990.
155. Dyrkheeva NS, Khodyreva SN, Lavrik OI. Interaction of APE1 and other repair proteins with DNA duplexes imitating intermediates of DNA repair and replication. *Biochemistry Biokhimiia*. 2008;73(3):261-72. Epub 2008/04/09. PubMed PMID: 18393760.
156. Baple EL, Chambers H, Cross HE, Fawcett H, Nakazawa Y, Chioza BA, et al. Hypomorphic PCNA mutation underlies a human DNA repair disorder. *The Journal of clinical investigation*. 2014;124(7):3137-46. Epub 2014/06/10. doi: 10.1172/jci74593. PubMed PMID: 24911150; PubMed Central PMCID: PMC4071375.
157. Green CM, Baple EL, Crosby AH. PCNA mutation affects DNA repair not replication. *Cell cycle (Georgetown, Tex)*. 2014;13(20):3157-8. doi: 10.4161/15384101.2014.969994. PubMed PMID: 25485490; PubMed Central PMCID: PMC4613369.
158. Clark AB, Valle F, Drotschmann K, Gary RK, Kunkel TA. Functional interaction of proliferating cell nuclear antigen with MSH2-MSH6 and MSH2-MSH3 complexes. *The Journal of biological chemistry*. 2000;275(47):36498-501. Epub 2000/09/27. doi: 10.1074/jbc.C000513200. PubMed PMID: 11005803.
159. Shereda RD, Machida Y, Machida YJ. Human KIAA1018/FAN1 localizes to stalled replication forks via its ubiquitin-binding domain. *Cell cycle (Georgetown, Tex)*. 2010;9(19):3977-83. Epub 2010/10/12. doi: 10.4161/cc.9.19.13207. PubMed PMID: 20935496.
160. Porro A, Berti M, Pizzolato J, Bologna S, Kaden S, Saxer A, et al. FAN1 interaction with ubiquitylated PCNA alleviates replication stress and preserves genomic integrity independently of BRCA2. *Nature communications*. 2017;8(1):1073. Epub 2017/10/21. doi: 10.1038/s41467-017-01074-6. PubMed PMID: 29051491; PubMed Central PMCID: PMC5648898.

161. Yang W. An overview of Y-Family DNA polymerases and a case study of human DNA polymerase η . *Biochemistry*. 2014;53(17):2793-803. Epub 2014/04/11. doi: 10.1021/bi500019s. PubMed PMID: 24716551; PubMed Central PMCID: PMC4018060.
162. Hibbert RG, Huang A, Boelens R, Sixma TK. E3 ligase Rad18 promotes monoubiquitination rather than ubiquitin chain formation by E2 enzyme Rad6. *Proceedings of the National Academy of Sciences of the United States of America*. 2011;108(14):5590-5. Epub 2011/03/23. doi: 10.1073/pnas.1017516108. PubMed PMID: 21422291; PubMed Central PMCID: PMC3078399.
163. Hoege C, Pfander B, Moldovan GL, Pyrowolakis G, Jentsch S. RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. *Nature*. 2002;419(6903):135-41. Epub 2002/09/13. doi: 10.1038/nature00991. PubMed PMID: 12226657.
164. Davies AA, Huttner D, Daigaku Y, Chen S, Ulrich HD. Activation of Ubiquitin-Dependent DNA Damage Bypass Is Mediated by Replication Protein A. *Molecular Cell*. 2008;29(5-2):625-36. doi: 10.1016/j.molcel.2007.12.016. PubMed PMID: 18342608; PubMed Central PMCID: PMC2507760.
165. Tsuji Y, Watanabe K, Araki K, Shinohara M, Yamagata Y, Tsurimoto T, et al. Recognition of forked and single-stranded DNA structures by human RAD18 complexed with RAD6B protein triggers its recruitment to stalled replication forks. *Genes to cells : devoted to molecular & cellular mechanisms*. 2008;13(4):343-54. Epub 2008/03/28. doi: 10.1111/j.1365-2443.2008.01176.x. PubMed PMID: 18363965.
166. Zou Y, Liu Y, Wu X, Shell SM. Functions of Human Replication Protein A (RPA): From DNA Replication to DNA Damage and Stress Responses. *Journal of cellular physiology*. 2006;208(2):267-73. doi: 10.1002/jcp.20622. PubMed PMID: 16523492; PubMed Central PMCID: PMC3107514.
167. Niimi A, Brown S, Sabbioneda S, Kannouche PL, Scott A, Yasui A, et al. Regulation of proliferating cell nuclear antigen ubiquitination in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America*. 2008;105(42):16125-30. Epub 2008/10/11. doi: 10.1073/pnas.0802727105. PubMed PMID: 18845679; PubMed Central PMCID: PMC2571029.
168. Wyatt MD, Pittman DL. Methylating agents and DNA repair responses: Methylated bases and sources of strand breaks. *Chemical research in toxicology*. 2006;19(12):1580-94. Epub 2006/12/19. doi: 10.1021/tx060164e. PubMed PMID: 17173371; PubMed Central PMCID: PMC2542901.
169. Shah KA, Shishkin AA, Voineagu I, Pavlov YI, Shcherbakova PV, Mirkin SM. Role of DNA polymerases in repeat-mediated genome instability. *Cell reports*.

2012;2(5):1088-95. Epub 2012/11/13. doi: 10.1016/j.celrep.2012.10.006. PubMed PMID: 23142667; PubMed Central PMCID: PMCPMC3513503.

170. Huang M, D'Andrea AD. A new nuclease member of the FAN club. *Nature structural & molecular biology*. 2010;17(8):926-8. Epub 2010/08/05. doi: 10.1038/nsmb0810-926. PubMed PMID: 20683477; PubMed Central PMCID: PMCPMC2945811.

171. Liu T, Ghosal G, Yuan J, Chen J, Huang J. FAN1 acts with FANCI-FANCD2 to promote DNA interstrand cross-link repair. *Science (New York, NY)*. 2010;329(5992):693-6. Epub 2010/07/31. doi: 10.1126/science.1192656. PubMed PMID: 20671156.

172. Thongthip S, Bellani M, Gregg SQ, Sridhar S, Conti BA, Chen Y, et al. Fan1 deficiency results in DNA interstrand cross-link repair defects, enhanced tissue karyomegaly, and organ dysfunction. *Genes & Development*. 2016;30(6):645-59. doi: 10.1101/gad.276261.115. PubMed PMID: 26980189; PubMed Central PMCID: PMCPMC4803051.

173. Identification of Genetic Factors that Modify Clinical Onset of Huntington's Disease. *Cell*. 2015;162(3):516-26. Epub 2015/08/02. doi: 10.1016/j.cell.2015.07.003. PubMed PMID: 26232222; PubMed Central PMCID: PMCPMC4524551.

174. Bettencourt C, Hensman-Moss D, Flower M, Wiethoff S, Brice A, Goizet C, et al. DNA repair pathways underlie a common genetic mechanism modulating onset in polyglutamine diseases. *Annals of Neurology*. 2016;79(6):983-90. doi: 10.1002/ana.24656. PubMed PMID: 27044000; PubMed Central PMCID: PMCPMC4914895.

175. Jin H, Cho Y. Structural and functional relationships of FAN1. *DNA repair*. 2017;56:135-43. Epub 2017/06/18. doi: 10.1016/j.dnarep.2017.06.016. PubMed PMID: 28623094.

176. Gates KS. An overview of chemical processes that damage cellular DNA: spontaneous hydrolysis, alkylation, and reactions with radicals. *Chem Res Toxicol*. 2009;22(11):1747-60. Epub 2009/09/18. doi: 10.1021/tx900242k. PubMed PMID: 19757819; PubMed Central PMCID: PMCPMC2806061.

177. Lindahl T. Instability and decay of the primary structure of DNA. *Nature*. 1993;362(6422):709-15. Epub 1993/04/22. doi: 10.1038/362709a0. PubMed PMID: 8469282.

178. Krokan HE, Bjoras M. Base excision repair. *Cold Spring Harb Perspect Biol*. 2013;5(4):a012583. Epub 2013/04/03. doi: 10.1101/cshperspect.a012583. PubMed PMID: 23545420; PubMed Central PMCID: PMCPMC3683898.

179. Nazarkina ZK, Khodyreva SN, Marsin S, Lavrik OI, Radicella JP. XRCC1 interactions with base excision repair DNA intermediates. *DNA repair*. 2007;6(2):254-64. Epub 2006/11/23. doi: 10.1016/j.dnarep.2006.10.002. PubMed PMID: 17118717.
180. Brouwer JR, Willemsen R, Oostra BA. Microsatellite repeat instability and neurological disease. *Bioessays*. 2009;31(1):71-83. Epub 2009/01/21. doi: 10.1002/bies.080122. PubMed PMID: 19154005; PubMed Central PMCID: PMCPMC4321794.
181. Loeb LA. Microsatellite instability: marker of a mutator phenotype in cancer. *Cancer Res*. 1994;54(19):5059-63. Epub 1994/10/01. PubMed PMID: 7923117.
182. Boland CR, Goel A. Microsatellite Instability in Colorectal Cancer. *Gastroenterology*. 2010;138(6):2073-87.e3. doi: 10.1053/j.gastro.2009.12.064. PubMed PMID: 20420947; PubMed Central PMCID: PMCPMC3037515.
183. Aguilera A, Gomez-Gonzalez B. Genome instability: a mechanistic view of its causes and consequences. *Nature reviews Genetics*. 2008;9(3):204-17. Epub 2008/01/30. doi: 10.1038/nrg2268. PubMed PMID: 18227811.
184. Parsons R, Li GM, Longley MJ, Fang WH, Papadopoulos N, Jen J, et al. Hypermutability and mismatch repair deficiency in RER+ tumor cells. *Cell*. 1993;75(6):1227-36. Epub 1993/12/17. PubMed PMID: 8261516.
185. Gatchel JR, Zoghbi HY. Diseases of unstable repeat expansion: mechanisms and common principles. *Nature reviews Genetics*. 2005;6(10):743-55. Epub 2005/10/06. doi: 10.1038/nrg1691. PubMed PMID: 16205714.
186. Wells RD. Molecular basis of genetic instability of triplet repeats. *The Journal of biological chemistry*. 1996;271(6):2875-8. Epub 1996/02/09. PubMed PMID: 8621672.
187. van Roon-Mom WM, Reid SJ, Faull RL, Snell RG. TATA-binding protein in neurodegenerative disease. *Neuroscience*. 2005;133(4):863-72. Epub 2005/05/27. doi: 10.1016/j.neuroscience.2005.03.024. PubMed PMID: 15916858.
188. Brooks BP, Fischbeck KH. Spinal and bulbar muscular atrophy: a trinucleotide-repeat expansion neurodegenerative disease. *Trends Neurosci*. 1995;18(10):459-61. Epub 1995/10/01. PubMed PMID: 8545913.
189. Turner C, Hilton-Jones D. The myotonic dystrophies: diagnosis and management. *Journal of Neurology, Neurosurgery & Psychiatry*. 2010;81(4):358-67.
190. Montermini L, Andermann E, Labuda M, Richter A, Pandolfo M, Cavalcanti F, et al. The Friedreich ataxia GAA triplet repeat: premutation and normal alleles. *Human molecular genetics*. 1997;6(8):1261-6. Epub 1997/08/01. PubMed PMID: 9259271.

191. Nolin SL, Brown WT, Glicksman A, Houck JGE, Gargano AD, Sullivan A, et al. Expansion of the Fragile X CGG Repeat in Females with Premutation or Intermediate Alleles. *American Journal of Human Genetics*. 2003;72(2):454-64. PubMed PMID: 12529854; PubMed Central PMCID: PMCPmc379237.
192. Jin P, Warren ST. Understanding the molecular basis of fragile X syndrome. *Hum Mol Genet*. 2000;9(6):901-8. Epub 2000/04/18. PubMed PMID: 10767313.
193. Schaefer MH, Wanker EE, Andrade-Navarro MA. Evolution and function of CAG/polyglutamine repeats in protein-protein interaction networks. *Nucleic Acids Res*. 2012;40(10):4273-87. Epub 2012/01/31. doi: 10.1093/nar/gks011. PubMed PMID: 22287626; PubMed Central PMCID: PMCPmc3378862.
194. Giovannucci E, Stampfer MJ, Krithivas K, Brown M, Dahl D, Brufsky A, et al. The CAG repeat within the androgen receptor gene and its relationship to prostate cancer. *Proceedings of the National Academy of Sciences of the United States of America*. 1997;94(7):3320-3. Epub 1997/04/01. PubMed PMID: 9096391; PubMed Central PMCID: PMCPmc20367.
195. Kazemi-Esfarjani P, Trifiro MA, Pinsky L. Evidence for a repressive function of the long polyglutamine tract in the human androgen receptor: possible pathogenetic relevance for the (CAG) n-expanded neuronopathies. *Human molecular genetics*. 1995;4(4):523-7.
196. Tirabassi G, Cignarelli A, Perrini S, Delli Muti N, Furlani G, Gallo M, et al. Influence of CAG Repeat Polymorphism on the Targets of Testosterone Action. *International journal of endocrinology*. 2015;2015:298107. Epub 2015/10/01. doi: 10.1155/2015/298107. PubMed PMID: 26421011; PubMed Central PMCID: PMCPmc4572434.
197. Heinlein CA, Chang C. Androgen receptor in prostate cancer. *Endocr Rev*. 2004;25(2):276-308. Epub 2004/04/15. doi: 10.1210/er.2002-0032. PubMed PMID: 15082523.
198. Lonergan PE, Tindall DJ. Androgen receptor signaling in prostate cancer development and progression. *J Carcinog*. 2011;10. doi: 10.4103/1477-3163.83937. PubMed PMID: 21886458; PubMed Central PMCID: PMCPmc3162670.
199. Kang S, Ohshima K, Shimizu M, Amirhaeri S, Wells RD. 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*. 1995;270(45):27014-21. Epub 1995/11/10. PubMed PMID: 7592950.
200. Oussatcheva EA, Hashem VI, Zou Y, Sinden RR, Potaman VN. Involvement of the nucleotide excision repair protein UvrA in instability of CAG*CTG repeat sequences

in *Escherichia coli*. *The Journal of biological chemistry*. 2001;276(33):30878-84. Epub 2001/06/20. doi: 10.1074/jbc.M104697200. PubMed PMID: 11413147.

201. Jakupciak JP, Wells RD. Genetic instabilities of triplet repeat sequences by recombination. *IUBMB life*. 2000;50(6):355-9. Epub 2001/05/01. doi: 10.1080/713803749. PubMed PMID: 11327307.

202. Lin Y, Hubert L, Wilson JH. TRANSCRIPTION DESTABILIZES TRIPLET REPEATS. *Mol Carcinog*. 2009;48(4):350-61. doi: 10.1002/mc.20488. PubMed PMID: 18973172; PubMed Central PMCID: PMC2671855.

203. Dion V. Tissue specificity in DNA repair: lessons from trinucleotide repeat instability. *Trends in genetics : TIG*. 2014;30(6):220-9. Epub 2014/05/21. doi: 10.1016/j.tig.2014.04.005. PubMed PMID: 24842550.

204. Kanvah S, Joseph J, Schuster GB, Barnett RN, Cleveland CL, Landman U. Oxidation of DNA: damage to nucleobases. *Acc Chem Res*. 2010;43(2):280-7. Epub 2009/11/27. doi: 10.1021/ar900175a. PubMed PMID: 19938827.

205. Entezam A, Lokanga AR, Le W, Hoffman G, Usdin K. Potassium bromate, a potent DNA oxidizing agent, exacerbates germline repeat expansion in a fragile X premutation mouse model. *Hum Mutat*. 2010;31(5):611-6. Epub 2010/03/10. doi: 10.1002/humu.21237. PubMed PMID: 20213777; PubMed Central PMCID: PMC2951473.

206. Zhang J, Stevens MF, Bradshaw TD. Temozolomide: mechanisms of action, repair and resistance. *Curr Mol Pharmacol*. 2012;5(1):102-14. Epub 2011/11/30. PubMed PMID: 22122467.

207. Lai Y, Beaver JM, Lorente K, Melo J, Ramjagsingh S, AgoulNIK IU, et al. Base excision repair of chemotherapeutically-induced alkylated DNA damage predominantly causes contractions of expanded GAA repeats associated with Friedreich's ataxia. *PloS one*. 2014;9(4):e93464. Epub 2014/04/03. doi: 10.1371/journal.pone.0093464. PubMed PMID: 24691413; PubMed Central PMCID: PMC3972099.

208. Sobol RW, Horton JK, Kuhn R, Gu H, Singhal RK, Prasad R, et al. Requirement of mammalian DNA polymerase-beta in base-excision repair. *Nature*. 1996;379(6561):183-6. Epub 1996/01/11. doi: 10.1038/379183a0. PubMed PMID: 8538772.

209. Chan NL, Guo J, Zhang T, Mao G, Hou C, Yuan F, et al. Coordinated processing of 3' slipped (CAG)_n/(CTG)_n hairpins by DNA polymerases beta and delta preferentially induces repeat expansions. *The Journal of biological chemistry*. 2013;288(21):15015-22. Epub 2013/04/16. doi: 10.1074/jbc.M113.464370. PubMed PMID: 23585564; PubMed Central PMCID: PMC3663522.

210. Lai Y, Xu M, Zhang Z, Liu Y. Instability of CTG repeats is governed by the position of a DNA base lesion through base excision repair. *PloS one*. 2013;8(2):e56960. Epub 2013/03/08. doi: 10.1371/journal.pone.0056960. PubMed PMID: 23468897; PubMed Central PMCID: PMC3582642.
211. Xu M, Gabison J, Liu Y. Trinucleotide repeat deletion via a unique hairpin bypass by DNA polymerase beta and alternate flap cleavage by flap endonuclease 1. *Nucleic acids research*. 2013;41(3):1684-97. Epub 2012/12/22. doi: 10.1093/nar/gks1306. PubMed PMID: 23258707; PubMed Central PMCID: PMC3561997.
212. Xu M, Lai Y, Jiang Z, Terzidis MA, Masi A, Chatgililoglu C, et al. A 5', 8-cyclo-2'-deoxypurine lesion induces trinucleotide repeat deletion via a unique lesion bypass by DNA polymerase beta. *Nucleic acids research*. 2014;42(22):13749-63. Epub 2014/11/28. doi: 10.1093/nar/gku1239. PubMed PMID: 25428354; PubMed Central PMCID: PMC4267656.
213. Hill JW, Evans MK. Dimerization and opposite base-dependent catalytic impairment of polymorphic S326C OGG1 glycosylase. *Nucleic Acids Res*. 2006;34(5):1620-32. Epub 2006/03/22. doi: 10.1093/nar/gkl060. PubMed PMID: 16549874; PubMed Central PMCID: PMC3561997.
214. Terry PD, Umbach DM, Taylor JA. APE1 genotype and risk of bladder cancer: evidence for effect modification by smoking. *Int J Cancer*. 2006;118(12):3170-3. Epub 2006/01/21. doi: 10.1002/ijc.21768. PubMed PMID: 16425270.
215. Chang-Claude J, Popanda O, Tan XL, Kropp S, Helmbold I, von Fournier D, et al. Association between polymorphisms in the DNA repair genes, XRCC1, APE1, and XPD and acute side effects of radiotherapy in breast cancer patients. *Clin Cancer Res*. 2005;11(13):4802-9. Epub 2005/07/08. doi: 10.1158/1078-0432.ccr-04-2657. PubMed PMID: 16000577.
216. Dogru-Abbasoglu S, Aykac-Toker G, Hanagasi HA, Gurvit H, Emre M, Uysal M. The Arg194Trp polymorphism in DNA repair gene XRCC1 and the risk for sporadic late-onset Alzheimer's disease. *Neurol Sci*. 2007;28(1):31-4. Epub 2007/03/27. doi: 10.1007/s10072-007-0744-x. PubMed PMID: 17385092.
217. Ali MF, Meza JL, Rogan EG, Chakravarti D. Prevalence of BER gene polymorphisms in sporadic breast cancer. *Oncol Rep*. 2008;19(4):1033-8. Epub 2008/03/22. PubMed PMID: 18357393.
218. Sak SC, Barrett JH, Paul AB, Bishop DT, Kiltie AE. Comprehensive analysis of 22 XPC polymorphisms and bladder cancer risk. *Cancer Epidemiol Biomarkers Prev*. 2006;15(12):2537-41. Epub 2006/12/14. doi: 10.1158/1055-9965.epi-06-0288. PubMed PMID: 17164382.

219. Jiang X, Zhou LT, Zhang SC, Chen K. XPC Polymorphism Increases Risk of Digestive System Cancers: Current Evidence from A Meta-Analysis. *Chin J Cancer Res*. 2012;24(3):181-9. Epub 2013/01/30. doi: 10.1007/s11670-012-0181-0. PubMed PMID: 23359774; PubMed Central PMCID: PMC3555276.
220. Tome S, Manley K, Simard JP, Clark GW, Slean MM, Swami M, et al. MSH3 polymorphisms and protein levels affect CAG repeat instability in Huntington's disease mice. *PLoS genetics*. 2013;9(2):e1003280. Epub 2013/03/08. doi: 10.1371/journal.pgen.1003280. PubMed PMID: 23468640; PubMed Central PMCID: PMC3585117.
221. Miao HK, Chen LP, Cai DP, Kong WJ, Xiao L, Lin J. MSH3 rs26279 polymorphism increases cancer risk: a meta-analysis. *Int J Clin Exp Pathol*. 2015;8(9):11060-7. Epub 2015/12/01. PubMed PMID: 26617824; PubMed Central PMCID: PMC4637639.
222. Xu XL, Yao YL, Xu WZ, Feng JG, Mao WM. Correlation of MSH3 polymorphisms with response and survival in advanced non-small cell lung cancer patients treated with first-line platinum-based chemotherapy. *Genet Mol Res*. 2015;14(2):3525-33. Epub 2015/05/13. doi: 10.4238/2015.April.15.16. PubMed PMID: 25966119.
223. Morales F, Vasquez M, Santamaria C, Cuenca P, Corrales E, Monckton DG. A polymorphism in the MSH3 mismatch repair gene is associated with the levels of somatic instability of the expanded CTG repeat in the blood DNA of myotonic dystrophy type 1 patients. *DNA repair*. 2016;40:57-66. Epub 2016/03/20. doi: 10.1016/j.dnarep.2016.01.001. PubMed PMID: 26994442.
224. Martins S, Pearson CE, Coutinho P, Provost S, Amorim A, Dube MP, et al. Modifiers of (CAG)(n) instability in Machado-Joseph disease (MJD/SCA3) transmissions: an association study with DNA replication, repair and recombination genes. *Hum Genet*. 2014;133(10):1311-8. Epub 2014/07/17. doi: 10.1007/s00439-014-1467-8. PubMed PMID: 25026993.
225. Dogru-Abbasoglu S, Tanrikulu S, Ademoglu E, Erbil Y, Ozderya A, Karadag B, et al. Polymorphisms of DNA base-excision repair genes APE/Ref-1 and XRCC1 are not associated with the risk for Graves' disease. *Cell Biochem Funct*. 2009;27(7):462-7. Epub 2009/08/28. doi: 10.1002/cbf.1595. PubMed PMID: 19711438.
226. Parildar-Karpuzoglu H, Dogru-Abbasoglu S, Hanagasi HA, Karadag B, Gurvit H, Emre M, et al. Single nucleotide polymorphisms in base-excision repair genes hOGG1, APE1 and XRCC1 do not alter risk of Alzheimer's disease. *Neurosci Lett*. 2008;442(3):287-91. Epub 2008/08/02. doi: 10.1016/j.neulet.2008.07.047. PubMed PMID: 18672023.

227. Zhang Y, Newcomb PA, Egan KM, Titus-Ernstoff L, Chanock S, Welch R, et al. Genetic polymorphisms in base-excision repair pathway genes and risk of breast cancer. *Cancer Epidemiol Biomarkers Prev.* 2006;15(2):353-8. Epub 2006/02/24. doi: 10.1158/1055-9965.epi-05-0653. PubMed PMID: 16492928.
228. Yamtich J, Nemec AA, Keh A, Sweasy JB. A germline polymorphism of DNA polymerase beta induces genomic instability and cellular transformation. *PLoS genetics.* 2012;8(11):e1003052. Epub 2012/11/13. doi: 10.1371/journal.pgen.1003052. PubMed PMID: 23144635; PubMed Central PMCID: PMC3493456.
229. Sobol RW. Genome instability caused by a germline mutation in the human DNA repair gene POLB. *PLoS genetics.* 2012;8(11):e1003086. Epub 2012/11/13. doi: 10.1371/journal.pgen.1003086. PubMed PMID: 23144636; PubMed Central PMCID: PMC3493457.
230. Mohrenweiser HW, Xi T, Vazquez-Matias J, Jones IM. Identification of 127 amino acid substitution variants in screening 37 DNA repair genes in humans. *Cancer Epidemiol Biomarkers Prev.* 2002;11(10 Pt 1):1054-64. Epub 2002/10/12. PubMed PMID: 12376507.
231. Yamtich J, Speed WC, Straka E, Kidd JR, Sweasy JB, Kidd KK. Population-specific variation in haplotype composition and heterozygosity at the POLB locus. *DNA repair.* 2009;8(5):579-84. Epub 2009/01/27. doi: 10.1016/j.dnarep.2008.12.005. PubMed PMID: 19167932; PubMed Central PMCID: PMC2791466.
232. Lin GC, Jaeger J, Eckert KA, Sweasy JB. Loop II of DNA polymerase beta is important for discrimination during substrate binding. *DNA repair.* 2009;8(2):182-9. Epub 2008/11/18. doi: 10.1016/j.dnarep.2008.10.006. PubMed PMID: 19013261; PubMed Central PMCID: PMC2653711.
233. El-Andaloussi N, Valovka T, Toueille M, Hassa PO, Gehrig P, Covic M, et al. Methylation of DNA polymerase beta by protein arginine methyltransferase 1 regulates its binding to proliferating cell nuclear antigen. *FASEB J.* 2007;21(1):26-34. Epub 2006/11/23. doi: 10.1096/fj.06-6194com. PubMed PMID: 17116746.
234. Gary R, Kim K, Cornelius HL, Park MS, Matsumoto Y. Proliferating cell nuclear antigen facilitates excision in long-patch base excision repair. *J Biol Chem.* 1999;274(7):4354-63. Epub 1999/02/06. PubMed PMID: 9933638.
235. Tom S, Henricksen LA, Park MS, Bambara RA. DNA ligase I and proliferating cell nuclear antigen form a functional complex. *J Biol Chem.* 2001;276(27):24817-25. Epub 2001/05/02. doi: 10.1074/jbc.M101673200. PubMed PMID: 11331287.
236. Kedar PS, Kim SJ, Robertson A, Hou E, Prasad R, Horton JK, et al. Direct interaction between mammalian DNA polymerase beta and proliferating cell nuclear

antigen. *J Biol Chem.* 2002;277(34):31115-23. Epub 2002/06/14. doi: 10.1074/jbc.M201497200. PubMed PMID: 12063248.

237. Desai NA, Shankar V. Single-strand-specific nucleases. *FEMS Microbiol Rev.* 2003;26(5):457-91. Epub 2003/02/15. PubMed PMID: 12586391.

238. Coppede F, Migheli F, Ceravolo R, Bregant E, Rocchi A, Petrozzi L, et al. The hOGG1 Ser326Cys polymorphism and Huntington's disease. *Toxicology.* 2010;278(2):199-203. Epub 2009/10/28. doi: 10.1016/j.tox.2009.10.019. PubMed PMID: 19857538.

239. Berger F, Vaslin L, Belin L, Asselain B, Forlani S, Humbert S, et al. The impact of single-nucleotide polymorphisms (SNPs) in OGG1 and XPC on the age at onset of Huntington disease. *Mutat Res.* 2013;755(2):115-9. Epub 2013/07/09. doi: 10.1016/j.mrgentox.2013.04.020. PubMed PMID: 23830927.

240. Freudenthal BD, Gakhar L, Ramaswamy S, Washington MT. Structure of monoubiquitinated PCNA and implications for translesion synthesis and DNA polymerase exchange. *Nature structural & molecular biology.* 2010;17(4):479-84. Epub 2010/03/23. doi: 10.1038/nsmb.1776. PubMed PMID: 20305653; PubMed Central PMCID: PMC2920209.

241. Kannouche PL, Wing J, Lehmann AR. Interaction of human DNA polymerase ϵ with monoubiquitinated PCNA: a possible mechanism for the polymerase switch in response to DNA damage. *Mol Cell.* 2004;14(4):491-500. Epub 2004/05/20. PubMed PMID: 15149598.

242. Schulte J, Littleton JT. The biological function of the Huntingtin protein and its relevance to Huntington's Disease pathology. *Current trends in neurology.* 2011;5:65-78. PubMed PMID: 22180703; PubMed Central PMCID: PMC2937673.

243. Myers RH, MacDonald ME, Koroshetz WJ, Duyao MP, Ambrose CM, Taylor SA, et al. De novo expansion of a (CAG) n repeat in sporadic Huntington's disease. *Nature genetics.* 1993;5(2):168-73. Epub 1993/10/01. doi: 10.1038/ng1093-168. PubMed PMID: 8252042.

244. McColgan P, Tabrizi SJ. Huntington's disease: a clinical review. *European journal of neurology.* 2018;25(1):24-34. Epub 2017/08/18. doi: 10.1111/ene.13413. PubMed PMID: 28817209.

245. Ross CA, Tabrizi SJ. Huntington's disease: from molecular pathogenesis to clinical treatment. *The Lancet Neurology.* 2011;10(1):83-98. Epub 2010/12/18. doi: 10.1016/s1474-4422(10)70245-3. PubMed PMID: 21163446.

246. Roos RAC. Huntington's disease: a clinical review. *Orphanet Journal of Rare Diseases*. 2010;5:40. doi: 10.1186/1750-1172-5-40. PubMed PMID: 21171977; PubMed Central PMCID: PMCPMC3022767.
247. Kar K, Hoop CL, Drombosky KW, Baker MA, Kodali R, Arduini I, et al. beta-hairpin-mediated nucleation of polyglutamine amyloid formation. *Journal of molecular biology*. 2013;425(7):1183-97. Epub 2013/01/29. doi: 10.1016/j.jmb.2013.01.016. PubMed PMID: 23353826; PubMed Central PMCID: PMCPMC3602386.
248. Chen S, Berthelie V, Hamilton JB, O'Nuallain B, Wetzel R. Amyloid-like features of polyglutamine aggregates and their assembly kinetics. *Biochemistry*. 2002;41(23):7391-9. Epub 2002/06/05. PubMed PMID: 12044172.
249. Sukhanova MV, Khodyreva SN, Lebedeva NA, Prasad R, Wilson SH, Lavrik OI. Human base excision repair enzymes apurinic/apyrimidinic endonuclease1 (APE1), DNA polymerase beta and poly(ADP-ribose) polymerase 1: interplay between strand-displacement DNA synthesis and proofreading exonuclease activity. *Nucleic acids research*. 2005;33(4):1222-9. Epub 2005/02/26. doi: 10.1093/nar/gki266. PubMed PMID: 15731342; PubMed Central PMCID: PMCPMC549570.
250. Hubscher U, Maga G, Spadari S. Eukaryotic DNA polymerases. *Annual review of biochemistry*. 2002;71:133-63. Epub 2002/06/05. doi: 10.1146/annurev.biochem.71.090501.150041. PubMed PMID: 12045093.
251. Takata K, Shimizu T, Iwai S, Wood RD. Human DNA polymerase N (POLN) is a low fidelity enzyme capable of error-free bypass of 5S-thymine glycol. *The Journal of biological chemistry*. 2006;281(33):23445-55. Epub 2006/06/22. doi: 10.1074/jbc.M604317200. PubMed PMID: 16787914.
252. Beard WA, Wilson SH. Structures of human DNA polymerases nu and theta expose their end game. *Nature structural & molecular biology*. 2015;22(4):273-5. Epub 2015/04/04. doi: 10.1038/nsmb.3006. PubMed PMID: 25837870.
253. Zahn KE, Averill AM, Aller P, Wood RD, Double S. Human DNA polymerase theta grasps the primer terminus to mediate DNA repair. *Nature structural & molecular biology*. 2015;22(4):304-11. Epub 2015/03/17. doi: 10.1038/nsmb.2993. PubMed PMID: 25775267; PubMed Central PMCID: PMCPMC4385486.
254. Lee YS, Gao Y, Yang W. How a homolog of high-fidelity replicases conducts mutagenic DNA synthesis. *Nature structural & molecular biology*. 2015;22(4):298-303. Epub 2015/03/17. doi: 10.1038/nsmb.2985. PubMed PMID: 25775266; PubMed Central PMCID: PMCPMC4469489.
255. Jaiswal AS, Panda H, Law BK, Sharma J, Jani J, Hromas R, et al. NSC666715 and Its Analogs Inhibit Strand-Displacement Activity of DNA Polymerase β and Potentiate Temozolomide-Induced DNA Damage, Senescence and Apoptosis in

Colorectal Cancer Cells. PLoS ONE. 2015;10(5). doi: 10.1371/journal.pone.0123808. PubMed PMID: 25933036; PubMed Central PMCID: PMC4416822.

256. Rizvi S, Raza ST, Mahdi F. Telomere length variations in aging and age-related diseases. *Current aging science*. 2014;7(3):161-7. Epub 2015/01/24. PubMed PMID: 25612739.

257. Batra R, Nelles DA, Pirie E, Blue SM, Marina RJ, Wang H, et al. Elimination of Toxic Microsatellite Repeat Expansion RNA by RNA-Targeting Cas9. *Cell*. 2017;170(5):899-912.e10. Epub 2017/08/15. doi: 10.1016/j.cell.2017.07.010. PubMed PMID: 28803727.

258. Manning KS, Rao AN, Castro M, Cooper TA. BNA(NC) Gapmers Revert Splicing and Reduce RNA Foci with Low Toxicity in Myotonic Dystrophy Cells. *ACS chemical biology*. 2017;12(10):2503-9. Epub 2017/08/31. doi: 10.1021/acscchembio.7b00416. PubMed PMID: 28853853; PubMed Central PMCID: PMC5694563.

259. Garcia-Lopez A, Llamusi B, Orzaez M, Perez-Paya E, Artero RD. In vivo discovery of a peptide that prevents CUG-RNA hairpin formation and reverses RNA toxicity in myotonic dystrophy models. *Proceedings of the National Academy of Sciences of the United States of America*. 2011;108(29):11866-71. Epub 2011/07/07. doi: 10.1073/pnas.1018213108. PubMed PMID: 21730182; PubMed Central PMCID: PMC3141925.

260. Bieschke J, Herbst M, Wiglenda T, Friedrich RP, Boeddrich A, Schiele F, et al. Small-molecule conversion of toxic oligomers to nontoxic beta-sheet-rich amyloid fibrils. *Nature chemical biology*. 2011;8(1):93-101. Epub 2011/11/22. doi: 10.1038/nchembio.719. PubMed PMID: 22101602.

261. Ehrnhoefer DE, Bieschke J, Boeddrich A, Herbst M, Masino L, Lurz R, et al. EGCG redirects amyloidogenic polypeptides into unstructured, off-pathway oligomers. *Nature structural & molecular biology*. 2008;15(6):558-66. Epub 2008/05/31. doi: 10.1038/nsmb.1437. PubMed PMID: 18511942.

262. Mossuto MF, Bolognesi B, Guixer B, Dhulesia A, Agostini F, Kumita JR, et al. Disulfide Bonds Reduce the Toxicity of the Amyloid Fibrils Formed by an Extracellular Protein. *Angewandte Chemie (International Ed in English)*. 2011;50(31):7048-51. doi: 10.1002/anie.201100986. PubMed PMID: 21671315; PubMed Central PMCID: PMC3166663.

263. Malinina L. Possible involvement of the RNAi pathway in trinucleotide repeat expansion diseases. *Journal of biomolecular structure & dynamics*. 2005;23(3):233-5. Epub 2005/10/13. doi: 10.1080/07391102.2005.10531230. PubMed PMID: 16218750.

264. Nalavade R, Griesche N, Ryan DP, Hildebrand S, Krauß S. Mechanisms of RNA-induced toxicity in CAG repeat disorders. *Cell Death & Disease*. 2013;4(8):e752-. doi: 10.1038/cddis.2013.276. PubMed PMID: 23907466; PubMed Central PMCID: PMC3763438.
265. Eddy S, Maddukuri L, Ketkar A, Zafar MK, Henninger EE, Pursell ZF, et al. Evidence for the kinetic partitioning of polymerase activity on G-quadruplex DNA. *Biochemistry*. 2015;54(20):3218-30. doi: 10.1021/acs.biochem.5b00060. PubMed PMID: 25903680; PubMed Central PMCID: PMC3763438.
266. Washington MT, Johnson RE, Prakash L, Prakash S. Human DNA polymerase ϵ utilizes different nucleotide incorporation mechanisms dependent upon the template base. *Molecular and cellular biology*. 2004;24(2):936-43. Epub 2004/01/01. PubMed PMID: 14701763; PubMed Central PMCID: PMC3763438.
267. Arana ME, Takata K, Garcia-Diaz M, Wood RD, Kunkel TA. A unique error signature for human DNA polymerase η . *DNA repair*. 2007;6(2):213-23. Epub 2006/11/23. doi: 10.1016/j.dnarep.2006.09.012. PubMed PMID: 17118716; PubMed Central PMCID: PMC1950682.
268. Yang W. Nucleases: diversity of structure, function and mechanism. *Quarterly reviews of biophysics*. 2011;44(1):1-93. Epub 2010/09/22. doi: 10.1017/s0033583510000181. PubMed PMID: 20854710.
269. Pizzolato J, Mukherjee S, Scharer OD, Jiricny J. FANCD2-associated nuclease 1, but not exonuclease 1 or flap endonuclease 1, is able to unhook DNA interstrand cross-links in vitro. *The Journal of biological chemistry*. 2015;290(37):22602-11. Epub 2015/07/30. doi: 10.1074/jbc.M115.663666. PubMed PMID: 26221031; PubMed Central PMCID: PMC3763438.
270. Zhao Q, Xue X, Longerich S, Sung P, Xiong Y. Structural insights into 5' flap DNA unwinding and incision by the human FANL1 dimer. *Nature communications*. 2014;5:5726. Epub 2014/12/17. doi: 10.1038/ncomms6726. PubMed PMID: 25500724; PubMed Central PMCID: PMC3763438.
271. Kratz K, Schopf B, Kaden S, Sandoel A, Eberhard R, Lademann C, et al. Deficiency of FANCD2-associated nuclease KIAA1018/FANL1 sensitizes cells to interstrand crosslinking agents. *Cell*. 2010;142(1):77-88. Epub 2010/07/07. doi: 10.1016/j.cell.2010.06.022. PubMed PMID: 20603016.
272. MacKay C, Declais AC, Lundin C, Agostinho A, Deans AJ, MacArtney TJ, et al. Identification of KIAA1018/FANL1, a DNA repair nuclease recruited to DNA damage by monoubiquitinated FANCD2. *Cell*. 2010;142(1):65-76. Epub 2010/07/07. doi: 10.1016/j.cell.2010.06.021. PubMed PMID: 20603015; PubMed Central PMCID: PMC3763438.

273. Smogorzewska A, Desetty R, Saito TT, Schlabach M, Lach FP, Sowa ME, et al. A genetic screen identifies FANL1, a Fanconi anemia-associated nuclease necessary for DNA interstrand crosslink repair. *Molecular Cell*. 2010;39(1):36-47. Epub 2010/07/07. doi: 10.1016/j.molcel.2010.06.023. PubMed PMID: 20603073; PubMed Central PMCID: PMCPMC2919743.
274. Bruning JB, Shamoo Y. Structural and thermodynamic analysis of human PCNA with peptides derived from DNA polymerase-delta p66 subunit and flap endonuclease-1. *Structure (London, England : 1993)*. 2004;12(12):2209-19. Epub 2004/12/04. doi: 10.1016/j.str.2004.09.018. PubMed PMID: 15576034.
275. Hibbert RG, Sixma TK. Intrinsic flexibility of ubiquitin on proliferating cell nuclear antigen (PCNA) in translesion synthesis. *The Journal of biological chemistry*. 2012;287(46):39216-23. Epub 2012/09/20. doi: 10.1074/jbc.M112.389890. PubMed PMID: 22989887; PubMed Central PMCID: PMCPMC3493961.
276. Lopez MA, Mackler RM, Yoder KE. Removal of nuclease contamination during purification of recombinant Prototype Foamy Virus integrase. *Journal of virological methods*. 2016;235:134-8. doi: 10.1016/j.jviromet.2016.06.002. PubMed PMID: 27269588; PubMed Central PMCID: PMCPMC4992616.
277. Callis J. The Ubiquitination Machinery of the Ubiquitin System. *The Arabidopsis Book / American Society of Plant Biologists*. 2014;12. doi: 10.1199/tab.0174. PubMed PMID: 25320573; PubMed Central PMCID: PMCPMC4196676.
278. Scheffner M, Nuber U, Huibregtse JM. Protein ubiquitination involving an E1-E2-E3 enzyme ubiquitin thioester cascade. *Nature*. 1995;373(6509):81-3. Epub 1995/01/05. doi: 10.1038/373081a0. PubMed PMID: 7800044.
279. Stewart MD, Ritterhoff T, Klevit RE, Brzovic PS. E2 enzymes: more than just middle men. *Cell Research*. 2016;26(4):423-40. Epub 2016/03/24. doi: 10.1038/cr.2016.35. PubMed PMID: 27002219; PubMed Central PMCID: PMCPMC4822130.
280. Ardley HC, Robinson PA. E3 ubiquitin ligases. *Essays in biochemistry*. 2005;41:15-30. Epub 2005/10/28. doi: 10.1042/eb0410015. PubMed PMID: 16250895.
281. Strzalka W, Ziemienowicz A. Proliferating cell nuclear antigen (PCNA): a key factor in DNA replication and cell cycle regulation. *Annals of Botany*. 2011;107(7):1127-40. doi: 10.1093/aob/mcq243. PubMed PMID: 21169293; PubMed Central PMCID: PMCPMC3091797.
282. Good MC, Zalatan JG, Lim WA. Scaffold proteins: hubs for controlling the flow of cellular information. *Science (New York, NY)*. 2011;332(6030):680-6. Epub 2011/05/10. doi: 10.1126/science.1198701. PubMed PMID: 21551057; PubMed Central PMCID: PMCPMC3117218.

283. Huang M, D'Andrea AD. A new nuclease member of the FAN club. *Nature structural & molecular biology*. 2010;17(8):926-8. doi: 10.1038/nsmb0810-926. PubMed PMID: 20683477; PubMed Central PMCID: PMC2945811.

VITA

YAOU REN

2012	B.S., Sanitary Science Sichuan University Sichuan, Chengdu, China
2013-2014	Teaching Assistant Florida International University Miami, FL, USA
2014-2018	Research Assistant Florida International University Miami, FL, USA
2017-2018	Doctoral Candidate Biochemistry Ph.D. program Florida International University Miami, FL, USA

PUBLICATIONS AND PRESENTATIONS

Chen, C., Jiang X., Ren, Y., and Zhang, Z. (2013) Arsenic trioxide co-exposure potentiates benzo(a)pyrene genotoxicity by enhancing the oxidative stress in vitro. *Biol Trace Elem Res.* 156(1-3): 338-349.

Cui, J., Zhao, W., Xu, X., Yang, M., Ren, Y., and Zhang, Z. (2012) DNA polymerase beta is involved in the protection against the cytotoxicity and genotoxicity of cigarette smoke. *Environ Toxicol Pharmacol.* 34(2): 370-380.

Ioannou, D., Jiang, Z., Cao, N., Ren, Y., Becerra, V., Vaccarella, A., Kandukuri, L., and Tempest, G.H. Gene repositioning in human lymphocytes following exposure to genotoxic agents. Conference paper in Chromosome Research, 2014.

Ren, Y., Garcia, C.B., Lai, Y., Wen, Z., Yang, W., Wnuk, S.F., and Liu, Y. Potential role of 5-azidomethyl-2'-deoxyuridine triphosphate in cancer therapy. EMGS 48th Annual Meeting, Sept. 9-13, 2017, Raleigh, NC.

Ren, Y., Lai, Y., Laverde, E.E., Lei, R., Rein, H.L., and Liu Y. (2017) Modulation of trinucleotide repeat instability by DNA polymerase β polymorphic variant R137Q. *PLoS ONE.* 12(5): e0177299.

Ren, Y., Lai, Y., Laverde, E.E., and Liu, Y. The DNA polymerase β polymorphic variant (R137Q) fail to alter the trinucleotide repeat instability. 17th Annual Biomedical and Comparative Immunology Symposium, Mar. 26-27, 2015, Miami, FL.

Ren, Y., Lai, Y., Laverde, E.E., and Liu, Y. Trinucleotide repeat instability is not affected by a DNA polymerase β polymorphic variant, arginine 137 glutamine. FIU Scholarly Forum, Apr. 3-4, 2015, Miami, FL.

Ren, Y., Lai, Y., Laverde, E.E., and Liu, Y. DNA Polymerase β R137Q polymorphism does not significantly alter trinucleotide repeat instability. EMGS 46th Annual Meeting, September 26-30, 2015, New Orleans, LA.

Ren, Y., Lai, Y., Laverde, E.E., and Liu, Y. Trinucleotide repeat instability is not significantly affected by the DNA polymerase β R137Q variant. 18th Annual Biomedical and Comparative Immunology Symposium, Mar. 3-4, 2016, Miami, FL.

Ren, Y., Lai, Y., Laverde, E.E., and Liu, Y. A polymorphic DNA polymerase β R137Q variant does not affect (CTG)_n trinucleotide repeat instability. FIU Scholarly Forum, Mar. 28-29, 2016, Miami, FL.

Ren, Y., Lai, Y., Laverde, E.E., Rein, H., and Liu, Y. Deficiency of DNA synthesis of a polymorphic variant of DNA polymerase β promotes trinucleotide repeat deletion. EMGS 45th Annual Meeting, September 13-17, 2014, Orlando, FL.

Ren, Y., Lai, Y., Laverde, E.E., Sweasy, J.B., and Liu, Y. Repeat instability mediated by DNA polymerase β variant Y265C during base excision repair. 2016 FASEB Conference in Dynamic DNA Structures in Biology, Jul. 10-15, 2016, Saxtons Rivers, VT.

Ren, Y., Lai, Y., Laverde, E.E., Sweasy, J.B., and Liu, Y. DNA polymerase β variant Y265C with impaired DNA synthesis activity induces repeat deletion during base excision repair. FIU Scholarly Forum, Mar. 27-31, 2017, Miami, FL.

Ren, Y., Lai, Y., Laverde, E.E., Sweasy, J.B., and Liu, Y. DNA polymerase β variant Y265C predominantly promotes repeat deletions during base excision repair. 19th Annual Biomedical and Comparative Immunology Symposium, Mar. 30-31, 2017, Miami, FL.

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