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
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 βR137Q, in leading TNR instability-mediated cancers or neurodegenerative diseases. We find that the pol βR137Q has slightly weaker DNA synthesis activity compared to that of wild-type (WT) pol β. Because of the similar abilities between pol βR137Q and WT pol β in bypassing a template loop structure, both pol βR137Q and WT pol β induces similar amount of repeat deletion. We conclude that the slightly weaker DNA synthesis activity of pol βR137Q does not alter the TNR instability compared to that of WT pol β, suggesting that the pol βR137Q 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 βY265C 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’ exonuclease 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

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>A. DNA Damage and DNA Repair</td>
<td>1</td>
</tr>
<tr>
<td>B. Mechanism of Base Excision Repair, Repair Enzymes, and Cofactors</td>
<td>4</td>
</tr>
<tr>
<td>C. Trinucleotide Repeat Instability and Human Diseases</td>
<td>12</td>
</tr>
<tr>
<td>D. Trinucleotide Repeat Instability and DNA Replication</td>
<td>17</td>
</tr>
<tr>
<td>E. Trinucleotide Repeat Instability and Base Excision Repair</td>
<td>21</td>
</tr>
<tr>
<td>OVERVIEW</td>
<td>37</td>
</tr>
<tr>
<td>CHAPTER 1: MODULATION OF TRINUCLEOTIDE REPEAT INSTABILITY BY DNA POLYMERASE (\beta) POLYMORPHIC VARIANT R137Q</td>
<td>39</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>39</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>40</td>
</tr>
<tr>
<td>MATERIALS AND METHODS</td>
<td>43</td>
</tr>
<tr>
<td>RESULTS</td>
<td>51</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>72</td>
</tr>
<tr>
<td>CHAPTER 2: WEAK DNA SYNTHESIS BY DNA POLYMERASES PROMOTES TRINUCLEOTIDE REPEAT DELETIONS DURING BASE EXCISION REPAIR</td>
<td>76</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>76</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>77</td>
</tr>
<tr>
<td>MATERIALS AND METHODS</td>
<td>80</td>
</tr>
<tr>
<td>RESULTS</td>
<td>86</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>114</td>
</tr>
<tr>
<td>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</td>
<td>119</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>119</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>120</td>
</tr>
<tr>
<td>MATERIALS AND METHODS</td>
<td>122</td>
</tr>
<tr>
<td>RESULTS</td>
<td>129</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>157</td>
</tr>
<tr>
<td>SUMMARY</td>
<td>161</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>165</td>
</tr>
<tr>
<td>VITA</td>
<td>196</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

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

1.1 Pol βR137Q 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 βR137Q 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 βR137Q during BER of a base lesion on (CAG)_{20} repeats .......................................................... 58

1.4 Pol βR137Q leads to weak FEN1 cleavage during BER in the context of (CAG)_{20} repeats .......................................................... 61

1.5 Effects of pol βR137Q 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 βR137Q 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 βR137Q 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 βY265C exhibits weak DNA synthesis activity during BER of a base lesion in the context of (CAG)_{20} repeats .......................................................... 88

2.2 Pol βY265C promotes (CAG)_{20} repeat deletions .......................................................... 91

2.3 Pol βY265C 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 βY265C promotes CTG repeat deletion during BER in the context of (CTG)_{20} repeats .......................................................... 100
2.7 Weak DNA synthesis of pol \( \nu \) promotes (CGG)\(_{20} \) repeat deletions

2.8 Pol \( \theta \) induces (GAA)\(_{20} \) repeat deletions

2.9 DNA synthesis activity of pol \( \beta \) is inhibited by \( \text{a pol } \beta \text{ inhibitor, NSC666719} \)

2.10 Inhibition of DNA synthesis activity of pol \( \beta \) promotes (CAG)\(_{20} \) repeat deletions

2.11 Pol \( \beta Y265C \) does not significantly alter (GT)\(_{30} \) repeat instability

2.12 Pol \( \beta Y265C \) does not significantly alter (TTAGGG)\(_{10} \) repeat instability

CHAPTER 3

3.1 P11 phosphocellulose cation exchange chromatography of PCNA

3.2 DEAE ion-exchange chromatography of PCNA

3.3 Phenyl sepharose chromatography of PCNA

3.4 Q sepharose anion exchange chromatography of PCNA

3.5 The first Heparin chromatography of PCNA

3.6 The second Heparin chromatography of PCNA

3.7 Mono Q anion exchange chromatography of PCNA

3.8 Expression of UBE1 and UbcH5c S22R

3.9 The Mono Q anion exchange chromatography of UBE1

3.10 The SP sepharose cation exchange chromatography of UbcH5c S22R

3.11 Mono-ubiquitination of PCNA

3.12 Q sepharose anion exchange chromatography of ub-PCNA

3.13 Phenyl sepharose chromatography of ub-PCNA

3.14 Mono Q anion exchange chromatography of ub-PCNA

3.15 The 5’-flap and hairpin structures formed during BER in the context of TNR tracts
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
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>°C</td>
<td>degree Celsius</td>
</tr>
<tr>
<td>A</td>
<td>adenine</td>
</tr>
<tr>
<td>APE1</td>
<td>AP Endonuclease 1</td>
</tr>
<tr>
<td>AP site</td>
<td>apurinic/apyrimidinic site</td>
</tr>
<tr>
<td>AR</td>
<td>androgen receptor</td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine triphosphate</td>
</tr>
<tr>
<td>BER</td>
<td>base excision repair</td>
</tr>
<tr>
<td>Bis-Tris</td>
<td>2,2-Bis(hydroxymethyl)-2,2’-2’-nitrilotriethanol</td>
</tr>
<tr>
<td>BRCT</td>
<td>BRCA1 C terminus domain</td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
</tr>
<tr>
<td>C (Cys)</td>
<td>cysteine</td>
</tr>
<tr>
<td>CDK</td>
<td>cyclin-dependent kinase</td>
</tr>
<tr>
<td>cdPu</td>
<td>5’,8-cyclo-2’-deoxypurines</td>
</tr>
<tr>
<td>Cordycepin</td>
<td>3’-deoxyadenosine</td>
</tr>
<tr>
<td>DM</td>
<td>myotonic dystrophy</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>dNTP</td>
<td>deoxyribonucleotide triphosphate</td>
</tr>
<tr>
<td>dRP</td>
<td>deoxyribose phosphate</td>
</tr>
<tr>
<td>dsDNA</td>
<td>double-stranded DNA</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>8-oxoG</td>
<td>8-oxoguanine</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>E. coli</td>
<td>Escherichia coli</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EME1</td>
<td>essential meiotic structure-specific endonuclease 1</td>
</tr>
<tr>
<td>EXO1</td>
<td>endonuclease 1</td>
</tr>
<tr>
<td>5’,8-cdA</td>
<td>5’,8-cyclo-2’-deoxyadenosine</td>
</tr>
<tr>
<td>5’-dRP</td>
<td>5’-deoxyribose phosphate</td>
</tr>
<tr>
<td>5-hU</td>
<td>5-hydroxyuracil</td>
</tr>
<tr>
<td>FAN1</td>
<td>Fanconi associated nuclease 1</td>
</tr>
<tr>
<td>FEN1</td>
<td>flap endonuclease 1</td>
</tr>
<tr>
<td>FMR1</td>
<td>fragile X mental retardation 1</td>
</tr>
<tr>
<td>FMSF</td>
<td>phenylmethylsulfonyl fluoride</td>
</tr>
<tr>
<td>FRDA</td>
<td>Friedreich’s ataxia</td>
</tr>
<tr>
<td>FXN</td>
<td>frataxin</td>
</tr>
<tr>
<td>FXS</td>
<td>fragile X syndrome</td>
</tr>
<tr>
<td>G</td>
<td>guanine</td>
</tr>
<tr>
<td>Gln</td>
<td>glutamine (Q)</td>
</tr>
<tr>
<td>HCl</td>
<td>hydrogen chloride</td>
</tr>
<tr>
<td>HD</td>
<td>Huntington’s disease</td>
</tr>
<tr>
<td>HEPES</td>
<td>4-(2-hydroxyethyl)‐1‐piperazineethanesulfonic acid</td>
</tr>
<tr>
<td>HMGB1</td>
<td>high‐mobility group protein B1</td>
</tr>
<tr>
<td>HR</td>
<td>homologous recombination</td>
</tr>
<tr>
<td>H₂O₂</td>
<td>hydrogen peroxide</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>I (Ile)</td>
<td>Isoleucine</td>
</tr>
<tr>
<td>ICL</td>
<td>interstrand crosslink</td>
</tr>
<tr>
<td>IPTG</td>
<td>isopropyl β-D-1-thiogalactopyranoside</td>
</tr>
<tr>
<td>IR</td>
<td>ionizing radiation</td>
</tr>
<tr>
<td>K (Lys)</td>
<td>lysine</td>
</tr>
<tr>
<td>KCl</td>
<td>potassium chloride</td>
</tr>
<tr>
<td>L (Leu)</td>
<td>leucine</td>
</tr>
<tr>
<td>LB</td>
<td>lysogeny broth</td>
</tr>
<tr>
<td>LIG I</td>
<td>DNA ligase I</td>
</tr>
<tr>
<td>LIG III</td>
<td>DNA ligase III</td>
</tr>
<tr>
<td>LP-BER</td>
<td>long patch-base excision repair</td>
</tr>
<tr>
<td>MBD4</td>
<td>methyl-binding domain glycosylase 4</td>
</tr>
<tr>
<td>MD</td>
<td>myotonic dystrophy</td>
</tr>
<tr>
<td>MEF</td>
<td>mouse embryonic fibroblast</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>magnesium chloride</td>
</tr>
<tr>
<td>MLH</td>
<td>MutL homologue</td>
</tr>
<tr>
<td>MMR</td>
<td>mismatch repair</td>
</tr>
<tr>
<td>MMS</td>
<td>methyl methanesulfonate</td>
</tr>
<tr>
<td>MRN</td>
<td>Mre11-Rad50-Nbs1 complex</td>
</tr>
<tr>
<td>MSH</td>
<td>MutS homologue</td>
</tr>
<tr>
<td>Mus81</td>
<td>MUS81 structure-specific endonuclease subunit</td>
</tr>
<tr>
<td>MutSβ</td>
<td>MSH2/MSH3</td>
</tr>
</tbody>
</table>
MW  molecular weight
MYH  MutY homologue
N  nitrogen
NaCl  sodium chloride
NaH$_2$PO$_4$  sodium dihydrogen phosphate
NaOH  sodium hydroxide
N7-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
O6-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
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAH</td>
<td>polycyclic aromatic hydrocarbon</td>
</tr>
<tr>
<td>PARP1</td>
<td>poly(ADP-ribose) polymerase 1</td>
</tr>
<tr>
<td>PCNA</td>
<td>proliferating cell nuclear antigen</td>
</tr>
<tr>
<td>pH</td>
<td>potential of hydrogen</td>
</tr>
<tr>
<td>pI</td>
<td>isoelectric point</td>
</tr>
<tr>
<td>PIP</td>
<td>PCNA-interacting protein</td>
</tr>
<tr>
<td>PMSF</td>
<td>phenylmethylsulfonyl fluoride</td>
</tr>
<tr>
<td>pol β</td>
<td>DNA polymerase β</td>
</tr>
<tr>
<td>pol δ</td>
<td>DNA polymerase δ</td>
</tr>
<tr>
<td>pol ε</td>
<td>DNA polymerase ε</td>
</tr>
<tr>
<td>pol η</td>
<td>DNA polymerase η</td>
</tr>
<tr>
<td>pol κ</td>
<td>DNA polymerase κ</td>
</tr>
<tr>
<td>pol ι</td>
<td>DNA polymerase ι</td>
</tr>
<tr>
<td>pol ν</td>
<td>DNA polymerase ν</td>
</tr>
<tr>
<td>pol θ</td>
<td>DNA polymerase θ</td>
</tr>
<tr>
<td>polyA</td>
<td>polyalanine</td>
</tr>
<tr>
<td>polyQ</td>
<td>polyglutamine</td>
</tr>
<tr>
<td>PTM</td>
<td>post-translational modification</td>
</tr>
<tr>
<td>Q</td>
<td>glutamine</td>
</tr>
<tr>
<td>R (Arg)</td>
<td>arginine</td>
</tr>
<tr>
<td>REV1</td>
<td>REV1, DNA directed polymerase</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>ROS</td>
<td>reactive oxygen species</td>
</tr>
<tr>
<td>RPA</td>
<td>replication protein A</td>
</tr>
<tr>
<td>S (Ser)</td>
<td>serine</td>
</tr>
<tr>
<td>SBMA</td>
<td>spinal and bulbar muscular atrophy</td>
</tr>
<tr>
<td>SCA10</td>
<td>spinocerebellar ataxia type 10</td>
</tr>
<tr>
<td>6-FAM</td>
<td>6-carboxyfluorescein</td>
</tr>
<tr>
<td>SMUG 1</td>
<td>single-strand selective monofunctional uracil DNA glycosylase</td>
</tr>
<tr>
<td>SP-BER</td>
<td>short patch base excision repair</td>
</tr>
<tr>
<td>SPD</td>
<td>synpolydactyly type II</td>
</tr>
<tr>
<td>ssDNA</td>
<td>single-stranded DNA</td>
</tr>
<tr>
<td>T</td>
<td>thymine</td>
</tr>
<tr>
<td>TDG</td>
<td>thymine DNA glycosylase</td>
</tr>
<tr>
<td>Tg</td>
<td>thymine glycol</td>
</tr>
<tr>
<td>THF</td>
<td>tetrahydrofuran</td>
</tr>
<tr>
<td>3meA</td>
<td>3-methyladenine</td>
</tr>
<tr>
<td>TL</td>
<td>translesion synthesis</td>
</tr>
<tr>
<td>TNR</td>
<td>trinucleotide repeat</td>
</tr>
<tr>
<td>Tris</td>
<td>tris(Hydroxymethyl)aminomethane</td>
</tr>
<tr>
<td>2-hA</td>
<td>2-hydryuracil</td>
</tr>
<tr>
<td>XPG</td>
<td>xeroderma pigmentosum complementation group G</td>
</tr>
<tr>
<td>XRCC1</td>
<td>X-ray repair cross-complementing protein 1</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>U</td>
<td>uracil</td>
</tr>
<tr>
<td>UbcH5c</td>
<td>ubiquitin-conjugating enzyme H5c</td>
</tr>
<tr>
<td>UBE1</td>
<td>ubiquitin activating enzyme E1</td>
</tr>
<tr>
<td>UBM</td>
<td>ubiquitin-binding module</td>
</tr>
<tr>
<td>UBZ</td>
<td>ubiquitin-binding domain</td>
</tr>
<tr>
<td>ub-PCNA</td>
<td>ubiquitinated PCNA</td>
</tr>
<tr>
<td>UDG</td>
<td>uracil DNA glycosylase</td>
</tr>
<tr>
<td>µM</td>
<td>micromolar</td>
</tr>
<tr>
<td>UNG</td>
<td>uracil DNA N-glycosylase</td>
</tr>
<tr>
<td>UTR</td>
<td>untranslated region</td>
</tr>
<tr>
<td>UV</td>
<td>ultraviolet</td>
</tr>
<tr>
<td>WT</td>
<td>wild-type</td>
</tr>
<tr>
<td>XPG</td>
<td>xeroderma pigmentosum complementation group G</td>
</tr>
<tr>
<td>Zic2</td>
<td>zinc-finger protein of the cerebellum 2</td>
</tr>
<tr>
<td>ZnCl₂</td>
<td>zinc chloride</td>
</tr>
</tbody>
</table>
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 exiguous 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$_{2}^-$), hydroxyl radicals (OH$^-$), singlet oxygen (O$_2^+$), and hydrogen peroxide (H$_2$O$_2$) 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 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]](image)
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 subject 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 IIIα (LIG IIIα) 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].
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 damage.

**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.
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 IIIα [68]. Disruption of the interaction between pol β and XRCC1 leads to a reduced efficiency of ligation by LIG IIIα [73], which
compromises the efficiency of BER. Meanwhile, an alternation in the BRCT domain of XRCC1 that interact with LIG IIIα directly affects expression of LIG IIIα [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 IIIα.

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 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.
Although several DNA metabolic pathways including DNA replication, NER, BER, and DNA recombination may induce the formation of secondary structures, many \textit{in vivo} and \textit{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, \text{pol } \varepsilon \text{ and } \text{pol } \delta, \text{ 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]. \textit{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 realignment 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 occur 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 DM1 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].
On the other hand, BER can also induce TNR deletions. It has been reported that some chemotherapeutical reagents such as temozolomide, ironizing 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

---

**Figure I.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.
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 base 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].
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

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.
deletions. In fact, pol βY265C with impaired DNA synthesis activity has been reported to reduce repeat expansion frequency in both sperm and brain cells of the heterozygous pol βY265C 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).
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’ exonucleaseolytic 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 IIIα [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 Lsy164 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.
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 1.17 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 βR137Q 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, polβR137Q 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 βR137Q variant on TNR instability. We showed that pol βR137Q 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 βR137Q 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 βR137Q 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,
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 βR137Q, pol βP242R, and pol βQ8R variant that are associated with cancer [61, 228-232]. Among them, pol βR137Q 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 βR137Q 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 βR137Q variant and its effects on CAG and CTG repeat instability during BER. We found that pol βR137Q 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 βR137Q variant modulates TNR instability in a similar manner as wild-type pol β. Our results further suggest that the individuals who carry pol βR137Q 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-32P ATP (6000 mCi/mmol) and Cordycepin 5’-triphosphate 3’-α-32P (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 Technologie (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)\textsubscript{20} repeats or (CTG)\textsubscript{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)\textsubscript{20} or (CTG)\textsubscript{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)\textsubscript{20} or (CTG)\textsubscript{20} repeats without damage was used as a size marker. The sequences of substrates are listed in Table 1.

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>nt</th>
<th>Sequence (5’-3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undamaged strand</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAG-0 (undamaged)</td>
<td>100</td>
<td>CGA GTC ATC TAG CAT CCG TA CAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AGA CTT ACT CAT TGC</td>
</tr>
<tr>
<td>CTG-0 (undamaged)</td>
<td>100</td>
<td>GCA ATG AGT AAG TCT ACG TA CTG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG</td>
</tr>
</tbody>
</table>

Table 1. Oligonucleotide Sequences [13]
<table>
<thead>
<tr>
<th>THF containing substrate</th>
<th>TA CGG ATG CTA GAT GAC TCG</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAG-1 (THF 5'-end)</td>
<td>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 CAG TA CGT AGA CTT ACT CAT TGC</td>
</tr>
<tr>
<td>CAG-10 (THF middle)</td>
<td>99</td>
</tr>
<tr>
<td>CTG-1 (THF 5'-end)</td>
<td>CGA GTC ATC TAG CAT CCG TA CATHF 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</td>
</tr>
<tr>
<td>CTG-10 (THF middle)</td>
<td>99</td>
</tr>
<tr>
<td>Template strand</td>
<td>GCA ATG AGT AAG TCT ACG TA CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG TA CGG ATG CTA GAT GAC TCG</td>
</tr>
<tr>
<td>CAG-T (CAG-1 and CAG-10 template)</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>GCA ATG AGT AAG TCT ACG TA CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG TA CGG ATG CTA GAT GAC TCG</td>
</tr>
<tr>
<td>CTG-1-T (CTG-1 template)</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>GCA ATG AGT AAG TCT ACG TA CAG CAG CAG CAG CAG CAG CAG CAG CAG 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</td>
</tr>
<tr>
<td>CTG-10-T (CTG-10 template)</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>CAG GTC ATC TAG CAT CCG TA CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG 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</td>
</tr>
<tr>
<td>1-nt gap substrate</td>
<td>50</td>
</tr>
<tr>
<td>CAG-gap-upstream</td>
<td>CAG GTC ATC TAG CAT CCG TA CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG TA CGT AGA CTT ACT</td>
</tr>
<tr>
<td>CAG-gap-downstream</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>AG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG TA CGT AGA CTT ACT</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>-------</td>
</tr>
<tr>
<td><strong>CAG-gap-template</strong></td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Random-gap-upstream</strong></td>
<td>50</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Random-gap-downstream</strong></td>
<td>49</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Random-gap-template</strong></td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Nick substrate</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Nick-CAG-1-up</strong></td>
<td>23</td>
</tr>
<tr>
<td><strong>Nick-CAG-1-down</strong></td>
<td>77</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Nick-CAG-1-Template</strong></td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Nick-CAG-10-up</strong></td>
<td>50</td>
</tr>
<tr>
<td><strong>Nick-CAG-10-down</strong></td>
<td>50</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Nick-CAG-10-Template</strong></td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Primers for DNA fragment analysis</strong></td>
<td></td>
</tr>
<tr>
<td><strong>CAG reverse primer</strong></td>
<td>23</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CAG forward primer</strong></td>
<td>21</td>
</tr>
<tr>
<td><strong>CTG-1 reverse primer</strong></td>
<td>23</td>
</tr>
</tbody>
</table>
Construction of pol βR137Q variant expression vector and expression and purification of pol βR137Q variant protein

The expression vector of pol βR137Q 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 βR137Q 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 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 by elution buffer containing 30 mM 4-(2-hydroxyethyl)-1-piperazinethane-sulfonic acid (HEPES), pH 7.5, 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 βR137Q variant, along with FEN1, LIG I and (CAG)₂₀ or (CTG)₂₀ substrates (25 nM) containing 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 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 βR137Q for 30 min in the BER buffer as described previously. Subsequently, the 10 µl reaction mixtures were subjected to S1 nuclease digestion at 37°C for 1, 3, 5, 10, 15 min, respectively with 5 µl 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 ZnCl$_2$, 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 µg 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 βR137Q 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 µM dNTPs and 5 mM Mg$^{2+}$ 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 βR137Q variant at 37°C for 15 min in reaction buffer under the conditions described previously. The 20 µl of reaction mixture was assembled on ice, and reaction was quenched by addition of 20 µl 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 µl of 100 mM EDTA, and reaction mixtures were incubated with 50 µl 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 βR137Q 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 βR137Q 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 βR137Q 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 βR137Q 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 βR137Q 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 βR137Q 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 βR137Q 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 βR137Q 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 βR137Q 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 βR137Q 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 βR137Q exhibited weaker DNA synthesis activity than wild-type pol β in the context of TNRs.
To compare the gap-filling synthesis activity of pol βR137Q variant and wild-type pol β at low concentrations, the gap-filling synthesis of pol βR137Q 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 βR137Q 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 βR137Q 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 βR137Q 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 βR137Q 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 βR137Q 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 βR137Q 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 βR137Q 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 βR137Q 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 βR137Q 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).
PCNA does not affect DNA synthesis activity of pol βR137Q 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 βR137Q variant has weak interaction with PCNA [61], it is possible that this can affect pol βR137Q 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 βR137Q variant. We found that in the absence of PCNA or presence of 50 nM and 100 nM PCNA, at 0.2 nM, both pol βR137Q 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 βR137Q variant (Figure 1.3), indicating that PCNA did not affect the processivity of pol β WT and pol βR137Q variant either. This further demonstrates that the impaired interaction between PCNA and pol βR137Q variant did not affect the polymerase gap-filling synthesis activity of the pol β variant.
The weak DNA synthesis of pol βR137Q 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 βR137Q on FEN1 cleavage activity by measuring FEN1 activity in the presence of 1 nM or 5 nM polβR137Q 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 βR137Q 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 βR137Q variant or wild-type pol β (Figure 1.4A, compare lanes 5-8 with lanes 3-4). In the presence of 5 nM pol βR137Q, 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 of 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 βR137Q and wild-type pol β by comparing to in the absence of pol βR137Q and wild-type pol β, FEN 1 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 βR137Q 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 βR137Q, 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 βR137Q resulted in weaker FEN1 cleavage of TNRs than that from wild-type pol β by creating a shorter repeat flap.
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 βR137Q 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 βR137Q) 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 βR137Q). Further analysis on the repeat size of the repaired products showed that BER reconstituted with 1 nM pol βR137Q 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 βR137Q 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 βR137Q 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 βR137Q 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 βR137Q exhibited the same ability as wild-type pol β enzyme to modulate TNR instability by causing one or two repeat deletions.
Pol βR137Q 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 βR137Q 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 βR137Q 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 βR137Q 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 βR137Q 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 βR137Q 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 βR137Q 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 βR137Q variant and wild-type pol β did not alter the formation of the downstream flap.
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 \(\beta\) (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 \(\beta\) or pol \(\beta R137Q\) 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 \(\beta\) and pol \(\beta R137Q\) with the substrate (Figure 1.5C and 1.5D, the left panel). This further suggested that pol \(\beta\) 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 \(\beta\), 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 βR137Q, 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 βR137Q 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 βR137Q 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 βR137Q 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 βR137Q 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.
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 1.7. 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 βR137Q 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 βR137Q. 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 βR137Q, 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.
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 βR137Q exhibited weaker DNA synthesis activity than wild-type pol β in the context.

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.
of TNRs (Figure 1.1). However, the DNA synthesis activity of pol βR137Q 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 βR137Q 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 βR137Q 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 βR137Q 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 βR137Q 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 βR137Q 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 βR137Q 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 βR137Q 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 βR137Q variant in the context of a TNR tract (Figure 1.3). This indicates that the impaired interaction between pol βR137Q 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 βR137Q 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 βR137Q 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 βR137Q 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 βR137Q 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 βY265C, 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 βY265C and a translesion DNA polymerase (pol ν) with weak DNA synthesis predominantly promoted TNR deletions through BER. The inefficient DNA synthesis of pol βY265C 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 prevents repeat expansion, which is supported by the fact that a pol β mutant, pol βY265C variant, with impaired DNA synthesis activity has been reported to reduce the repeat expansion frequency in both sperm and brain cells from heterozygous pol βY265C 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 βY265C 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 internationally 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-32P ATP (6000 mCi/mmol) and Cordycepin 5’-triphosphate 3’-α-32P (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 βY265C 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 Heath. 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>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>nt</th>
<th>Sequence (5’-3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(CAG)$_{20}$ substrate</td>
<td></td>
<td>CAG-0 (Undamaged strand)</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>CGA GTC ATC TAG CAT CCG TA CAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAG CAG CAG CAG CAG CAG CAG CAG CAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAG CAG CAG CAG CAG CAG CAG CAG CAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAG CAG CAG CAG CAG CAG CAG CAG CAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAG CAG CAG CAG CAG CAG TA CGT AGA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CTT ACT CAT TGC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAG-1 (THF 1$^{st}$ repeat)</td>
</tr>
<tr>
<td></td>
<td>99</td>
<td>CGA GTC ATC TAG CAT CCG TA CAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAT$^\text{THF}$ CAG CAG CAG CAG CAG CAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAG CAG CAG CAG CAG CAG CAG CAG CAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAG CAG CAG CAG CAG CAG CAG CAG CAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAG CAG CAG CAG CAG CAG CAG CAG CAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAG CAG CAG CAG CAG CAG CAG CAG TA CGT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AGA CTT ACT CAT TGC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAG-10 (THF 10$^{th}$ repeat)</td>
</tr>
<tr>
<td></td>
<td>99</td>
<td>CGA GTC ATC TAG CAT CCG TA CAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAG CAG CAG CAG CAG CAG CAG CAG CAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAG CAG CAG CAG CAG CAG CAG CAG CAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAG CAG CAG CAG CAG TA CGT AGA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CTT ACT CAT TGC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAG-T (Template strand)</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>GCA ATG AGT AAG TCT ACG TA CTG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CTG CTG CTG CTG CTG TA CGG ATG</td>
</tr>
<tr>
<td>Substrate Type</td>
<td>Damaged Strand</td>
<td>Sequence</td>
</tr>
<tr>
<td>----------------</td>
<td>----------------</td>
<td>----------</td>
</tr>
<tr>
<td><strong>(CTG)₂₀ substrate</strong></td>
<td>CTG-0 (Undamaged strand)</td>
<td>100 CTA GAT GAC TCG CTA GAT GAC TCG CTA GAT GAC TCG CTA GAT GAC TCG CTA GAT GAC TCG CTA GAT GAC TCG CTA GAT GAC TCG</td>
</tr>
<tr>
<td></td>
<td>CTG-1 (THF 1st repeat)</td>
<td>99 CTA GAT GAC TCG CTA GAT GAC TCG CTA GAT GAC TCG CTA GAT GAC TCG CTA GAT GAC TCG CTA GAT GAC TCG</td>
</tr>
<tr>
<td></td>
<td>CTG-10 (THF 10th repeat)</td>
<td>99 CTA GAT GAC TCG CTA GAT GAC TCG CTA GAT GAC TCG</td>
</tr>
<tr>
<td></td>
<td>CTG-T (Template strand)</td>
<td>100 CTA GAT GAC TCG CTA GAT GAC TCG CTA GAT GAC TCG</td>
</tr>
<tr>
<td><strong>(CGG)₂₀ substrate</strong></td>
<td>CGG-0 (Undamaged strand)</td>
<td>100 CTA GAT GAC TCG CTA GAT GAC TCG CTA GAT GAC TCG CTA GAT GAC TCG CTA GAT GAC TCG CTA GAT GAC TCG</td>
</tr>
<tr>
<td></td>
<td>CGG-10 (THF 10th repeat)</td>
<td>99 CTA GAT GAC TCG CTA GAT GAC TCG CTA GAT GAC TCG</td>
</tr>
<tr>
<td></td>
<td>CGG-T (Template strand)</td>
<td>100 CTA GAT GAC TCG CTA GAT GAC TCG CTA GAT GAC TCG</td>
</tr>
<tr>
<td><strong>(GAA)₂₀ substrate</strong></td>
<td>GAA-0 (Undamaged strand)</td>
<td>100 CTA GAT GAC TCG CTA GAT GAC TCG CTA GAT GAC TCG</td>
</tr>
</tbody>
</table>

82
<table>
<thead>
<tr>
<th>Sequence Description</th>
<th>Percentage</th>
<th>Sequence</th>
</tr>
</thead>
</table>
| GAA-10 (THF 10th repeat)            | 99%        | CTT ACT CAT TGC  
CGA GTC ATC TAG CAT CCG TA GAA  
GAA GAA GAA GAA GAA GAA GAA GAA  
GAA THF GAA 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  
TTC TTC TTC TTC TTC TTC TTC TTC  
TTC TTC TTC TTC TTC TTC TTC TTC  
TTC TTC TTC TA CGT AGA CTA GAT |
| (GT)$_{30}$ substrate               |            | GT-0 (Undamaged strand) 100% | CGA GTC ATC TAG CAT CCG TA GT GT  
GT GT GT GT GT GT GT GT GT GT GT  
GT GT GT GT GT GT GT GT GT GT GT  
GT GT GT GT GT GT GT GT GT GT GT  
GT GT GT GT GT GT GT GT GT GT GT  
GT GT GT GT GT GT GT TA CGT AGA  
CTT ACT CAT TGC |
|                                      |            | GT-15 (THF 15th repeat) 99% | CGA GTC ATC TAG CAT CCG TA GT GT  
GT GT GT GT GT GT GT GT GT GT GT  
GT GT GT GT GT GT GT GT GT GT GT  
GT GT GT GT 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  
AC AC AC AC AC AC AC AC AC AC AC  
AC AC AC AC AC AC AC AC AC AC AC  
AC AC AC AC AC AC AC TA CGG ATG  
CTA GAT GAC TCG |
| (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 TTAGGG TA CGT AGA  
CTT ACT CAT TGC |
|                                      |            | TTAGGG-5 (THF 5th repeat) 99% | CGA GTC ATC TAG CAT CCG TA  
TTAGGG TTAGGG TTAGGG TTAGGG  
TTAGGTHF 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 |
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 βY265C, 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 µM dNTPs and 5 mM Mg$^{2+}$ 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 βY265C 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 Mg$^{2+}$ at 37°C for 15 min. The 20 µl-reaction mixture was assembled on ice, and reactions were quenched by addition of 20 µl 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-denaturating 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 βR137Q 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 MgCl$_2$, 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 MgCl$_2$ 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 βY265C promotes CAG repeat deletions independent of damage location

A polymorphic pol βR137Q 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 βY265C, 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 βY265C 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 βY265C 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 βY265C 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 βY265C indicates that pol βY265C 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.
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 βY265C 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 βY265C 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 βY265C 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 βY265C mutant protein, indicating that the DNA synthesis activity of pol βY265C 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 βY265C. 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 βY265C (Figure 2.3A, lane 7), and it removed up to 2 repeats when the pol βY265C 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 βY265C (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 βY265C did not significantly affect FEN1 cleavage activity. Comparing the major DNA synthesis products by pol βY265C 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 βY265C resulted in the loss of CAG repeats, leading to repeat deletions.
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 ν, may also be involved in modulating TNR instability during BER. Since pol ν is specialized in bypassing the thymine glycol lesion [251], it is

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 ^{32}P-labeled at the 3'-end of the damaged strand.
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 $\theta$ and pol $\nu$ 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 $\nu$ and pol $\theta$ with pol $\beta$) during BER
of an abasic lesion in the middle of a (CAG)$_{20}$ repeat tract. We found that Pol $\nu$ 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 $\beta$ 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 $\nu$ increased (Figure 2.4B, compare the
peak heights among panels b, c, and d), indicating that the increased amount of pol $\nu$
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 $\nu$, and 3 repeats
in the presence of 5 nM of pol $\nu$ (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
$\nu$ caused more repeats removed by FEN1, which led to TNR deletions.
Figure 2.4. Weak DNA synthesis of pol ν 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 ν 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).
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 0 does not induce (CAG)$_{20}$ repeat deletions.

(A) DNA synthesis activity of pol 0 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 0 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 0. 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 0.

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 βY265C 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 βY265C, and pol ν in modulating the instability of (CTG)$_{20}$ and (CGG)$_{20}$ tracts.

We initially tested the DNA synthesis activity of pol βY265C 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 βY265C 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 βY265C 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 βY265C 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 βY265C led to repeat deletion on the (CTG)$_{20}$ repeat tract. Both WT pol β and pol βY265C at the concentration of 1 nM induced 1 repeat deletion (Figure 2.6C and 2.5D, panels c and d). However, pol βY265C-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 βY265C 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)\textsubscript{20} repeats. 

(B) Measurement of DNA synthesis activity and BER reconstitution was conducted with the (CTG)\textsubscript{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.
On the other hand, pol ν 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 ν 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 ν (Figure 2.7C, lanes 7-9). The

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

Figure 2.7. Weak DNA synthesis of pol ν promotes (CGG)\textsubscript{20} repeat deletions.
(A) Measurements of DNA synthesis activity and BER reconstitution were conducted with (CGG)\textsubscript{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 ν 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 ν. Substrates were \textsuperscript{32}P-labeled at the 5'-end of the damaged strand.
Figure 2.7. Weak DNA synthesis of pol ν promotes (CGG)_{20} repeat deletions. 

(B) Sizing analysis of the repaired products from pol ν 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 ν.

C

(A) Sizing analysis of the repaired products from pol ν 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 ν.
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 0 induces (GAA)\textsubscript{20} repeat deletions.

(A) DNA synthesis activity of 1 nM, 2 nM, and 5 nM of pol 0 was measured with the (GAA)\textsubscript{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 0 at 1 nM, 2 nM, and 5 nM.

(B) Sizing analysis of the repaired products in the presence of 1 nM, 2 nM, and 5 nM of pol 0. Panels a and b represent the length of the (GAA)\textsubscript{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 v. 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 µM, 3 µM, and 5 µM of the inhibitor (Figure 2.9, lanes 4-6). When the concentrations of the inhibitor increased to 8 µM, 10 µM, and 25 µM, 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 µM, 3 µM, and 5 µM 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 µM, 10 µM, and 25 µM 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.](image)

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 µM, 3 µM, 5 µM, 8 µM, 10 µM, and 25 µM) 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 µM 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)\textsubscript{20} repeat deletions

(A) BER reconstituted in the presence of NSC666719 with the (CAG)\textsubscript{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 βY265C would promote deletions of GT repeats and telomere repeats.

First, we determined the DNA synthesis and repeat instability in the presence of pol βY265C with the substrate containing (GT)$_{30}$ with a lesion located in the middle. The results showed that both WT pol β and pol βY265C 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 βY265C 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 βY265C induced 1 repeat deletion products (Figure 2.11B, panels b and c). However, the weaker DNA synthesis activity of pol βY265C on GT repeats did not significantly alter the repeat instability during BER (compare panel c to panel b in Figure 2.11B).
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 βY265C 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 βY265C 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 βY265C 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 βY265C.

**DISCUSSION**

Trinucleotide repeat expansion diseases including Huntington’s disease, Friedreich’s ataxia, and mytonic 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-
decreasement 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, providing 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 βY265C 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 βY265C 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 βY265C performed similarly 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 \((\text{CGG/GCC})_n\) than that of \((\text{CAG/CTG})_n\). On the other hand, pol θ exhibited a marked difference in DNA synthesis on the \((\text{GAA/CTT})_n\) from that on the \((\text{CAG/CTG})_n\), causing different outcomes of repeat instabilities. On the \((\text{CAG/CTG})_n\), pol θ exhibited similar DNA synthesis activity as pol β and did not alter the instability of \((\text{CAG/CTG})_n\) (Figure 2.5). However, it performed weaker DNA synthesis activities on \((\text{GAA/CTT})_{20}\) 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 perform 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 βY265C 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 βY265C indicates that the slightly weaker DNA synthesis activity of pol βY265C on GT repeats than that of WT pol βY265C 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 βY265C 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 Exonucelase 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 characterize 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 kinases 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)\(_{15}\) 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**

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>nt</th>
<th>Sequence (5’-3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5’-flap substrate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upstream strand-(CAG)(_{15})</td>
<td>65</td>
<td>CGA GTC ATC TAG CAT CCG TA CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG</td>
</tr>
</tbody>
</table>
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
NaH$_2$PO$_4$ (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 NaH$_2$PO$_4$
(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-μl reaction mixture containing 160 nM Ube1, 32
uM UbcH5c S22R, 32 uM ubiquitin, and 16 uM 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-μl reaction mixture
was assembled on ice, and the reaction-mixtures were quenched with addition of 10 µl 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 µl) 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 strained 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 strained PAGE analysis of selective fractions 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.
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 3.4. 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.
Figure 3.5. The first Heparin chromatography of PCNA.
Coomassie strained 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 strained 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 strained 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.
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 UbcH5c S22R.](image)

Coomassie stained PAGE analysis of expression of UBE1 (118 kDa) and UbcH5c 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.](image)

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

On the other hand, the UbcH5c 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 strained PAGE analysis of selective elution fractions after the SP sepharose cation exchange chromatography of UBE1 (17.5 kDa).](image)

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

Ubiquitination of PCNA *in vivo* requires the E1 ubiquination 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 \textit{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) \textit{in vitro}. However, only 60\% of PCNA was ubiquitinationationed (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).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3.11.png}
\caption{Figure 3.11. Mono-ubiquitination of PCNA. Coomassie stained PAGE analysis of the yield of the monoubiquited PCNA (37.5 kDa). \textbf{A.} Mono-ubiquitination with low yield. Lanes 1-4 depict the ubiquitination reactions in a 10-ul reaction. PCNA, UB, UBE1 and UbcH5c illustrate the control of each protein. \textbf{B.} Monoubiquitination of PCNA under an optimal mono-ubiquitination.}
\end{figure}
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 stained PAGE analysis of selective fractions 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.
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).
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.

![Large flap formed on the TNR](image1.png) ![Hairpin structure formed on the TNR](image2.png)

**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.
To further characterize the cleavage of FEN1 or FAN1 on a short flap that was attached to a CAG repeat hairpin, we created an ahrpin 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 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.
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

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.
affect FAN1 cleavage of the (CAG)₁₅ flap or the short (CAG)₁ flap attached to a transien 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.

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

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 of 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.
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].
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’ exonuclease 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),
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’ exonucleaseolytic activity on the hairpins, we further tested whether the stimulation of FAN1 5’-3’ exonucleaseolytic 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.
DISCUSSION

Trinucleotide repeat instability is caused by the formation of secondary structures, such as hiarpins 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

**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.
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, wheras 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-ubquitination 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 mytonic 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, nuclease, 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 βR137Q 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 βR137Q 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 βR137Q 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 βY265C 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 βY265C 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
βY265C 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 to repeat deletions. We further demonstrate the weak DNA synthesis of pol βY265C 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 increases 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 monoubiquituated 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’ exonucleaseolytic activity, and this lead to disabilizatiin 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


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


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


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


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. 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.


