9-30-2016

Trinucleotide Repeat Instability is Modulated by DNA Base Lesions and DNA Base Excision Repair

Jill M. Beaver

Florida International University, jbeav004@fiu.edu

DOI: 10.25148/etd.FIDC001178

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

Part of the Biochemistry Commons

Recommended Citation

Beaver, Jill M., "Trinucleotide Repeat Instability is Modulated by DNA Base Lesions and DNA Base Excision Repair" (2016). FIU Electronic Theses and Dissertations. 3056.

https://digitalcommons.fiu.edu/etd/3056

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.
TRINUCLEOTIDE REPEAT INSTABILITY IS MODULATED BY DNA BASE LESIONS AND DNA BASE EXCISION REPAIR

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

DOCTOR OF PHILOSOPHY in

BIOCHEMISTRY

by

Jill Beaver

2016
To: Dean Michael R. Heithaus  
College of Arts, Sciences and Education  

This dissertation, written by Jill Beaver, and entitled Trinucleotide Repeat Instability is Modulated by DNA Base Lesions and DNA Base Excision Repair, having been approved in respect to style and intellectual content, is referred to you for judgment.

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

_______________________________________  
Lidia Kos  
  
_______________________________________  
Watson Lees  
  
_______________________________________  
Yuk-Ching Tse-Dinh  
  
_______________________________________  
Xiaotang Wang  
  
_______________________________________  
Yuan Liu, Major Professor  

Date of Defense: September 30, 2016

The dissertation of Jill Beaver 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, 2016
DEDICATION

I dedicate this dissertation to my parents, Catherine and Charles Therrell, and my husband, William Beaver, for their love, support, and encouragement.
ACKNOWLEDGMENTS

I would like to express my gratefulness to all the people who have helped me and given me support throughout the completion of my Ph.D. First of all, I would like to thank my advisor, Dr. Yuan Liu, for her unending assistance, advice, and support. Under her excellent mentorship and guidance, I have gained invaluable experience in scientific research, writing, and presentation. She has encouraged and supported me every step of the way.

I would like to thank my committee members, Dr. Lidia Kos, Dr. Watson Lees, Dr. Yuk-Ching Tse-Dinh, 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 immeasurable assistance and guidance in the lab supported my research at every step. She has offered patient instruction and support throughout the entirety of my research training. I would also like to thank Zhongliang Jiang, Yaou Ren, Ruipeng Lei, and Eduardo Laverde, as well as all other past and present members of the Liu Research Lab, for their help and support.

Finally, I thank the FIU Research Initiative for Science Enhancement (RISE) fellowship program which funded my graduate studies.
ABSTRACT OF THE DISSERTATION
TRINUCLEOTIDE REPEAT INSTABILITY IS MODULATED BY DNA BASE LESIONS AND DNA BASE EXCISION REPAIR

by

Jill Beaver
Florida International University, 2016
Miami, Florida
Professor Yuan Liu, Major Professor

Trinucleotide repeat (TNR) expansions are the cause of over 40 human neurodegenerative diseases, and are linked to DNA damage and base excision repair (BER). We explored the role of DNA damage and BER in modulating TNR instability through analysis of DNA structures, BER protein activities, and reconstitution of repair using human BER proteins and synthesized DNA containing various types of damage. We show that DNA damage and BER can modulate TNR expansions by promoting removal of a TNR hairpin through coordinated activities of BER proteins and cofactors. We found that during repair in a TNR hairpin, coordination between the 5’-flap endonuclease activity of flap endonuclease 1 (FEN1), 3’-5’ exonuclease activity of AP endonuclease 1 (APE1), and activity of DNA ligase I (LIG I) can resolve the double-flap structure produced during BER in the hairpin loop. The resolution of the double-flap structure resulted in hairpin removal and prevention or attenuation of TNR expansions and provides the first evidence that coordination among BER proteins can remove a TNR hairpin. We further explored the role of BER cofactors in modulating TNR instability and found that the repair cofactor proliferating cell nuclear antigen (PCNA) facilitates genomic stability by promoting
removal of a TNR hairpin. Hairpin removal was accomplished by altering dynamic TNR structures to allow more efficient FEN1 cleavage and DNA polymerase β (pol β) synthesis and stimulating the activity of LIG I. This study provides the first evidence that a DNA repair cofactor plays an important role in modulating TNR instability. Finally, we explored the effects of sugar modifications in abasic sites on activities of BER proteins and BER efficiency during repair in a TNR tract. We found that an oxidized sugar inhibits the activities of BER enzymes, interrupting their coordination and preventing efficient repair. Inefficient repair results in accumulation of repair intermediates with DNA breaks, contributing to genomic instability. Our results indicate that DNA base lesions and BER play a crucial role in modulating TNR instability. The research presented herein provides a molecular basis for further developing BER as a target for prevention and treatment of neurodegenerative diseases caused by TNR expansion.
# TABLE OF CONTENTS

<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 BASE EXCISION REPAIR</td>
<td>1</td>
</tr>
<tr>
<td>B. TRINUCLEOTIDE REPEAT INSTABILITY AND HUMAN DISEASE</td>
<td>13</td>
</tr>
<tr>
<td>C. BASE EXCISION REPAIR AND TRINUCLEOTIDE REPEAT INSTABILITY</td>
<td>19</td>
</tr>
<tr>
<td>D. BASE EXCISION REPAIR IN A TRINUCLEOTIDE REPEAT HAIRPIN</td>
<td>30</td>
</tr>
<tr>
<td>E. CROSSTALK BETWEEN BASE EXCISION REPAIR AND OTHER DNA REPAIR PATHWAYS</td>
<td>35</td>
</tr>
<tr>
<td>OVERVIEW</td>
<td>44</td>
</tr>
</tbody>
</table>

| CHAPTER 1: AP ENDONUCLEASE 1 PREVENTS TRINUCLEOTIDE REPEAT EXPANSION VIA A NOVEL MECHANISM DURING DNA BASE EXCISION REPAIR | 44 |
| ABSTRACT | 47 |
| INTRODUCTION | 47 |
| MATERIALS AND METHODS | 51 |
| RESULTS | 58 |
| DISCUSSION | 74 |

| CHAPTER 2: PROLIFERATING CELL NUCLEAR ANTIGEN PREVENTS TRINUCLEOTIDE REPEAT EXPANSIONS BY PROMOTING REPEAT DELETION AND HAIRPIN REMOVAL | 86 |
| ABSTRACT | 86 |
| INTRODUCTION | 86 |
| MATERIALS AND METHODS | 90 |
| RESULTS | 97 |
| DISCUSSION | 120 |

| CHAPTER 3: AN OXIDIZED ABASIC LESION INHIBITS DNA BASE EXCISION REPAIR RESULTING IN DNA STRAND BREAKS | 127 |
| ABSTRACT | 128 |
| INTRODUCTION | 128 |
| MATERIALS AND METHODS | 132 |
| RESULTS | 140 |
| DISCUSSION | 153 |

| SUMMARY | 156 |
| REFERENCES | 162 |
| VITA | 185 |
LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTRODUCTION</td>
<td></td>
</tr>
<tr>
<td>I.1 DNA Damage and DNA Repair Pathways</td>
<td>3</td>
</tr>
<tr>
<td>I.2 8-oxoG base pairs with C and A</td>
<td>5</td>
</tr>
<tr>
<td>I.3 Pathways of DNA Base Excision Repair</td>
<td>7</td>
</tr>
<tr>
<td>I.4 The dRP lyase mechanism of pol β</td>
<td>9</td>
</tr>
<tr>
<td>I.5 TNR expansions in coding and non-coding regions of DNA lead to human disease</td>
<td>15</td>
</tr>
<tr>
<td>I.6 Non-B form DNA secondary structures formed by trinucleotide repeats</td>
<td>17</td>
</tr>
<tr>
<td>I.7 The location of a TNR hairpin governs instability</td>
<td>18</td>
</tr>
<tr>
<td>I.8 Loss of coordination between BER proteins leads to TNR instability</td>
<td>24</td>
</tr>
<tr>
<td>I.9 Toxic oxidation cycle of BER in a TNR tract</td>
<td>25</td>
</tr>
<tr>
<td>I.10 TNR deletion during BER via pol β hairpin bypass synthesis</td>
<td>28</td>
</tr>
<tr>
<td>I.11 Resistance of lesions in a TNR hairpin to BER exacerbates the toxic oxidation cycle</td>
<td>32</td>
</tr>
<tr>
<td>I.12 BER of an oxidized lesion in the loop of a CAG hairpin is coupled to hairpin removal</td>
<td>33</td>
</tr>
<tr>
<td>I.13 Flap equilibration facilitates efficient FEN1 cleavage to prevent TNR expansions</td>
<td>35</td>
</tr>
<tr>
<td>I.14 Crosstalk between glycosylases and other DNA repair pathways</td>
<td>38</td>
</tr>
<tr>
<td>I.15 MSH2/MSH3 stabilizes a TNR hairpin</td>
<td>40</td>
</tr>
<tr>
<td>I.16 MSH2/MSH3 stabilization of a TNR hairpin promotes TNR expansions</td>
<td>41</td>
</tr>
<tr>
<td>I.17 Crosstalk between BER and MMR promotes expansions</td>
<td>43</td>
</tr>
</tbody>
</table>
CHAPTER 1

1.1 APE1 promotes the removal of a trinucleotide repeat hairpin during BER of 8-oxoG in a TNR hairpin loop

1.2 APE1 promotes the removal of a trinucleotide repeat hairpin during BER of an abasic site in a CAG repeat hairpin loop

1.3 APE1 stimulates removal of a double-flap intermediate during BER

1.4 APE1 3’-5’ exonuclease activity shortens the 3’-flap of the double-flap intermediates

1.5 APE1 fails to stimulate FEN1 cleavage of the double-flap intermediate

1.6 The 3’-5’exonuclease activity of APE1 at low concentrations

1.7 APE1 stimulates LIG I activity on a double-flap intermediate

1.8 APE1 fails to stimulate pol β DNA synthesis on the double-flap intermediate

1.9 APE1 prevents trinucleotide repeat expansions during BER in a hairpin loop

CHAPTER 2

2.1 PCNA stimulates repeat deletion during BER in a duplex (CAG)20 tract

2.2 A bubble forms in the template of a duplex (CAG)20 tract during BER

2.3 PCNA does not stimulate FEN1 flap cleavage in a duplex (CAG)20 tract

2.4 PCNA has no effect on pol β DNA synthesis during BER of an abasic site in a duplex (CAG)20 tract

2.5 PCNA stimulates LIG I activity on a nick across a small loop on the template strand

2.6 PCNA stimulates processing of double-flap intermediates resulting from a CAG repeat hairpin

2.7 PCNA stimulates FEN1 flap cleavage on double-flap intermediates

x
2.8 Formation of a flap in the upstream strand of double-flap intermediates......116

2.9 No secondary structure is formed in the template strand of the double-flap.........118

2.10 PCNA slightly stimulates pol β DNA synthesis on double-flap intermediates during BER ..................................................................................................................................................120

2.11 PCNA prevents TNR expansions during BER ..............................................122

CHAPTER 3

3.1 Abasic DNA lesions with modifications to the sugar residue ..................131

3.2 Preparation of DOB-containing substrates via photolysis .........................136

3.3 An oxidized abasic lesion inhibits pol β DNA synthesis activity during BER in a small TNR hairpin loop......................................................................................................................142

3.4 An oxidized abasic lesion inhibits pol β DNA synthesis activity during BER in a TNR duplex.................................................................................................................................144

3.5 An oxidized abasic lesion inhibits FEN1 5’-flap cleavage activity during BER in a small TNR hairpin loop ........................................................................................147

3.6 An oxidized abasic lesion slightly inhibits FEN1 5’-flap cleavage activity during BER in a TNR duplex.................................................................................................................................148

3.7 An oxidized abasic lesion decreases repair efficiency during BER in a small TNR hairpin loop..................................................................................................................................150

3.8 An oxidized abasic lesion decreases repair efficiency during BER in a TNR duplex..................................................................................................................................................152

3.9 An oxidized abasic lesion inhibits efficient BER in a TNR tract ...............155
<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>AAG</td>
<td>3-alkyladenine DNA glycosylase</td>
</tr>
<tr>
<td>ADP</td>
<td>adenosine diphosphate</td>
</tr>
<tr>
<td>AID</td>
<td>Activation-induced cytidine deaminase</td>
</tr>
<tr>
<td>ALS</td>
<td>amyotrophic lateral sclerosis</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>APOBEC</td>
<td>apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like</td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine triphosphate</td>
</tr>
<tr>
<td>BER</td>
<td>base excision repair</td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
</tr>
<tr>
<td>C</td>
<td>Cytosine</td>
</tr>
<tr>
<td>C4-AP</td>
<td>C4-oxidized abasic site</td>
</tr>
<tr>
<td>cdPu</td>
<td>5’,8-cyclo-2’-deoxypurines</td>
</tr>
<tr>
<td>CPDs</td>
<td>cyclobutane–pyrimidine dimers</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethyl sulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>dNTP</td>
<td>deoxyribonucleotide triphosphate</td>
</tr>
<tr>
<td>DOB</td>
<td>5’-(2-phosphoryl-1,4-dioxobutane)</td>
</tr>
<tr>
<td>dRP</td>
<td>deoxyribose phosphate</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>-------------</td>
<td>-------------------------------------------------</td>
</tr>
<tr>
<td>DSBR</td>
<td>double-strand break repair</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>E. coli</td>
<td>Escherichia coli</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetic acid</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>5hmC</td>
<td>5-hydroxymethylcytosine</td>
</tr>
<tr>
<td>5hmU</td>
<td>5-hydroxymethyluracil</td>
</tr>
<tr>
<td>5mC</td>
<td>5-methylcytosine</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>FPLC</td>
<td>fast protein liquid chromatography system</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>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>Abbreviation</td>
<td>Full Name</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</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>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>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>MNU</td>
<td>methyl nitrosourea</td>
</tr>
<tr>
<td>MRN</td>
<td>Mre11-Rad50-Nbs1 complex</td>
</tr>
<tr>
<td>MSH</td>
<td>MutS homologue</td>
</tr>
<tr>
<td>MutSα</td>
<td>MSH2/MSH6</td>
</tr>
<tr>
<td>MutSβ</td>
<td>MSH2/MSH3</td>
</tr>
<tr>
<td>MYH</td>
<td>MutY homologue</td>
</tr>
<tr>
<td>N</td>
<td>nitrogen</td>
</tr>
<tr>
<td>N7-meG</td>
<td>N⁷-methylguanine</td>
</tr>
<tr>
<td>NEIL1</td>
<td>endonuclease VIII-like 1</td>
</tr>
<tr>
<td>Acronym</td>
<td>Full Form</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
</tr>
<tr>
<td>NER</td>
<td>nucleotide excision repair</td>
</tr>
<tr>
<td>NHEJ</td>
<td>nonhomologous end joining</td>
</tr>
<tr>
<td>NP-40</td>
<td>nonidet P-40</td>
</tr>
<tr>
<td>nt</td>
<td>nucleotide</td>
</tr>
<tr>
<td>O</td>
<td>oxygen</td>
</tr>
<tr>
<td>O6-meG</td>
<td>O6-methylguanine</td>
</tr>
<tr>
<td>OGG1</td>
<td>8-oxoguanine DNA Glycosylase</td>
</tr>
<tr>
<td>PAGE</td>
<td>polyacrylamide Gel Electrophoresis</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>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>polyQ</td>
<td>polyglutamine</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>RNS</td>
<td>reactive nitrogen species</td>
</tr>
<tr>
<td>ROS</td>
<td>reactive oxygen species</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>SMUG 1</td>
<td>single-strand selective monofunctional uracil DNA glycosylase</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>SP-BER</td>
<td>short patch base excision repair</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>TET</td>
<td>ten eleven translocation protein</td>
</tr>
<tr>
<td>THF</td>
<td>tetrahydrofuran</td>
</tr>
<tr>
<td>3meA</td>
<td>3-methyladenine</td>
</tr>
<tr>
<td>TNR</td>
<td>trinucleotide repeats</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>U</td>
<td>uracil</td>
</tr>
<tr>
<td>UDG</td>
<td>uracil DNA glycosylase</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>
</tbody>
</table>
INTRODUCTION

A. DNA DAMAGE AND BASE EXCISION REPAIR

Deoxyribonucleic acid (DNA) is the essential biomolecule that carries all the genetic information of a cell and therefore determines the sequence and structure of all proteins in an organism. Thus, it is critical that the integrity and stability of the DNA sequence is maintained in order to ensure the successful passage of the genetic sequence to each new generation. However, the DNA molecule itself is unstable and prone to DNA damage and mutation (1). Damage to DNA can occur spontaneously as a result of self-decomposition or can result from exposure to damaging agents. Self-decomposition can take on one of two forms: spontaneous hydrolysis of the glycosidic bond or hydrolytic deamination. The N-glycosidic bond that connects the deoxyribose sugar of the nucleotide to the DNA nitrogenous base is vulnerable to spontaneous hydrolysis resulting in loss of the base and making the DNA more susceptible to DNA strand breaks (1, 2). Spontaneous base loss is known as depurination because the reaction predominantly occurs to remove the purines Guanine (G) and Adenine (A). Spontaneous depurination is thought to generate ~10,000 abasic (orapurinic/apyrimidinic (AP)) sites per cell per day (3). The resulting abasic sites are both cytotoxic and mutagenic (2, 4, 5) and can generate single strand breaks (6, 7), and modified sugars in an abasic site can lead to interstrand crosslinks (ICLs) that block replication and transcription (8). Another form of self-decomposition is hydrolytic deamination, particularly of Cytosine (C) to Uracil (U) and 5-methylcytosine (5mC) to Thymine (T) (9-12). This can result in C to T and C to A mutations. Thus, the nature of the DNA molecule lends itself to natural instability.
Aside from self-decomposition, DNA is also subjected to damage through exposure to both endogenous and exogenous stressors that can damage the DNA nitrogenous bases and the DNA backbone (1, 13, 14). Exogenous sources include Ultraviolet (UV) radiation from the sun, ionizing radiation (IR), chemotherapeutic drugs, and environmental toxins and pollutants. Ultraviolet radiation can induce bulky lesions such as cyclobutane-pyrimidine dimers (CPDs) (15) which distort the DNA helix and block DNA metabolic processes. Ionizing radiation can induce the production of Reactive Oxygen Species (ROS) that create strand breaks through the generation of free radicals and oxidative DNA base lesions (16) and can also induce formation of covalent bonds between bases to create thymine dimers (17) that block DNA replication and transcription. Environmental toxins such as chromate, bromate, and chemicals found in cigarette smoke can create a variety of DNA damage, including bulky lesions and both oxidative and alkylated DNA damage (18). Chemotherapeutic drugs such as bleomycin and neocarzinostatin also create a variety of lesions including bulky adducts, ICLs (16, 19), and both oxidative and alkylated base damage. Endogenous damage is caused by ROS and Reactive Nitrogen Species (RNS) that result from our cellular energy metabolism and create oxidative damage (1, 13, 14).

It is essential that DNA damage be repaired efficiently and accurately, as unrepaired damage can lead to distortion of the double helix, strand breaks, or mutations that can result in the onset of human diseases, including cancer if they alter the sequence or expression of oncogenes or tumor suppressor genes (5, 14). In addition, if damage accumulates in non-dividing cells, such as the neurons, it can lead to neurodegeneration and aging (20, 21). Furthermore, unrepaired DNA lesions and strand breaks can inhibit normal DNA metabolic processes such as DNA replication and gene transcription by blocking the advance of DNA
polymerase or RNA polymerase, respectively, and can further lead to cell death via apoptosis (22-24). Cells contain many DNA repair pathways to maintain the integrity and stability of the genome (Figure I.1). These include DNA base excision repair (BER), nucleotide excision repair (NER), mismatch repair (MMR), and double-strand break repair (DSBR)(25).

**Figure I.1 DNA Damage and DNA Repair Pathways (25)**

Different types of damage are repaired through different DNA repair pathways. Small base lesions that do not induce severe distortion in the DNA double helix are repaired by BER. These include oxidized and alkylated lesions, deaminated bases, and U, which
can result from either deamination of C or misincorporation by DNA polymerases (26-29). Base excision repair also provides a mechanism of repairing single-strand DNA (ssDNA) breaks. Larger lesions are repaired by other repair pathways, with bulky lesions and thymine dimers being repaired by NER, small loops and mismatches resulting from replication errors being repaired by MMR, and double-strand DNA (dsDNA) breaks being repaired by the DSBR pathways of homologous recombination (HR) or non-homologous end joining (NHEJ).

Base excision repair of small DNA base lesions is an efficient process that is used to remove the most common type of DNA damage, oxidation (30). There are over 100 types of oxidized base modifications that can occur in the DNA (31). Oxidative DNA damage is created by ROS, including the hydroxyl radical, hydrogen peroxide, and superoxide, which form as byproducts of cellular energy metabolism, inflammation, and exposure to IR and environmental toxicants such as bromate and chromate (30, 32). The accumulation of unrepaired oxidative damage can result in mutations, epigenetic instability, and interruption of transcription (33). The most common form of oxidized lesion in human cells is 8-oxo-7,8-dihydroguanine (8-oxoG), which is generated by radical attack at the C8 position of guanine due to its low redox potential (30) and is estimated to occur at a rate of about 180 8-oxoG lesions per day in a mammalian genome (1, 31). 8-oxoguanine has the ability to rotate about the glycosidic bond to adopt the anti- or syn-conformations. 8-oxoguanine is therefore mutagenic because of its ability to form both a Watson-Crick base pair with C when in the anti-conformation and a Hoogsteen base pair with A almost as efficiently when in the syn-conformation (32, 34) (Figure I.2). The pairing of 8-oxoG with A can result in G to T transversion mutations during DNA replication as
the lesion is bypassed by a DNA polymerase and may interrupt gene transcription depending on the location of the damage (35-37).

![Figure I.2 8-oxoG base pairs with C and A (32)](image)

In addition, some types of oxidized lesions such as the C4-oxidized AP site (C4-AP) and 5’-(2-phosphoryl-1,4-dioxobutane) (DOB) can irreversibly inhibit the BER protein DNA polymerase β (pol β) by forming a covalent crosslink with a key lysine residue in the active site of the deoxyribose phosphate (dRP) lyase domain (38), thus preventing efficient repair.

Though oxidative damage is the most common type of DNA damage, BER is also able to efficiently repair alkylated damage. Alkylating agents, such as methylating and ethylating agents, include chemotherapeutic agents like temozolomide and toxicants such as methyl nitrosourea (MNU), which is found in cigarette smoke. These agents primarily alkylate the
N and O atoms of the nitrogenous bases, with the N7 position of guanine being the most reactive, to form lesions such as N7-methylguanine (N\textsuperscript{7}-meG) and O\textsuperscript{6}-methylguanine (O\textsuperscript{6}-meG) \cite{2, 39}. Alkylation at certain positions can also accelerate the deamination process \cite{2}. Therefore, efficient BER is key in maintaining genomic integrity and stability by accurately repairing these frequently occurring lesions.

Removal of a single non-bulky base lesion is predominantly repaired in mammalian cells via BER \cite{40}. Base excision repair can efficiently repair oxidized and alkylated lesions, U, and ssDNA breaks, and has been found to occur in both dividing and non-dividing cells \cite{41}, indicating its critical role in the maintenance of genomic integrity. There are 2 subpathways of BER: short-patch BER (SP-BER, also known as single-nucleotide BER, SN-BER) and long-patch BER (LP-BER) (Figure I.3), which differ from each other in the number of nucleotides synthesized during repair and the repair enzymes and cofactors that are involved in lesion removal \cite{42-46}. Whether repair occurs via the short-patch or long-patch pathway depends on the type of DNA lesion and can also depend on the context of the DNA sequence.
A small DNA base lesion such as 8-oxoG or 3-methyladenine (3meA) is typically repaired by SP-BER. In SP-BER, only a single nucleotide is synthesized (42, 47) (Figure I.3 Pathways of DNA Base Excision Repair (46))

**Figure I.3 Pathways of DNA Base Excision Repair (46)**

A small DNA base lesion such as 8-oxoG or 3-methyladenine (3meA) is typically repaired by SP-BER. In SP-BER, only a single nucleotide is synthesized (42, 47) (Figure
I.3). The initiating steps of SP-BER and LP-BER are the same. Both pathways of BER are initiated by the recognition of a DNA lesion by a specific DNA glycosylase protein, and the subsequent removal of the lesion. The glycosylase protein removes the lesion by cleaving the N-glycosidic bond that links the base with the deoxyribose sugar, resulting in the formation of an abasic site (48). There are many DNA glycosylases in human cells which exhibit specificity for a wide range of base lesions (48-51). Examples include removal of 8-oxoG by 8-oxoguanine DNA glycosylase 1 (OGG1), removal of U by Uracil DNA glycosylase (UDG), and removal of 3meA by 3-alkyladenine DNA glycosylase (AAG), among many others. The abasic site that results from cleavage of the N-glycosidic bond is recognized and bound by AP endonuclease 1 (APE1), which makes an incision at the 5’-side of the abasic site, creating a single-strand nick in the DNA backbone with an upstream 3’-hydroxyl and a downstream 5’-dRP (52, 53). At this point in the repair process, if the 5’-dRP contains a native sugar residue, repair continues along the SP-BER pathway, however, if the sugar has experienced modification, such as oxidative stress, repair will occur through the LP-BER pathway. The next step in SP-BER is removal of the 5’-dRP by the lyase domain of DNA polymerase β (pol β) (Figure I.4 (54)), which acts as the rate-limiting step of SP-BER (55).
The dRP lyase mechanism occurs via a β-elimination reaction involving Schiff base formation and creates a temporary pol β-DNA crosslink intermediate (55). The removal of the 5’-dRP creates a 1 nt gap that is filled in by pol β DNA polymerase through its gap-filling synthesis activity creating a nick which is sealed by a DNA ligase protein, such as DNA Ligase I (LIG I) or the DNA Ligase III (LIG III)-x-ray cross-complementing group 1 (XRCC1) complex, to complete the repair process (56). Thus, in this process, only 1 nt is replaced.

When the abasic site itself has been modified, for example, the deoxyribose residue has been modified by oxidation, repair follows the LP-BER pathway (46) (Figure I.3). In the LP-BER pathway, 2 or more nucleotides are synthesized (42). While the initiating steps are the same as those for SP-BER, including removal of the base lesion by a glycosylase and incision of the DNA by APE1, the modified sugar residue is refractory to removal by pol β dRP lyase (46). The LP-BER pathway therefore requires the recruitment of the structure-specific nuclease flap endonuclease 1 (FEN1) in order to remove the DNA lesion.
Flap endonuclease 1 is a member of the RAD2 superfamily of 5’ endo/exonucleases and plays roles both in the processing of Okazaki fragments during lagging strand synthesis of DNA replication and in removing a modified 5’-lesion during LP-BER. There are 2 subpathways of LP-BER, which occur via the “hit and run” mechanism and the “strand displacement synthesis” mechanism (Figure I.3). In the “hit and run” subpathway, pol β inserts 1 nt to fill in the gap created by APE1 incision. Flap endonuclease 1 is then recruited to cleave the modified sugar along with one attached downstream nucleotide, creating a new 1 nt gap (57). DNA polymerase β gap-filling synthesis then fills in the 1 nt gap, creating a nick which is sealed by LIG I. The “hit and run” pathway requires coordinated sequential BER enzymatic reactions that are referred to as “passing the baton” (56), and the protein coordination may be critical in preventing the exposure of the BER intermediate structures, which could allow nucleolytic cleavage of the DNA or triggering of DNA recombination (56, 58, 59). In this “hit and run” BER process only 2 nt are replaced (46).

In contrast, in the “strand displacement synthesis” subpathway of LP-BER, 3 or more nucleotides are synthesized. Instead of filling in a 1nt gap, pol β and/or the replicative polymerases DNA polymerase δ (pol δ) and DNA polymerase ε (pol ε) carryout strand displacement synthesis by inserting 3 or more nucleotides to displace the downstream region of the DNA, creating a 5’-flap structure (43, 60-62). Flap endonuclease 1 is believed to load onto the free 5’-end of the flap and track down to the base where it cleaves the flap, resulting in removal of the DNA lesion (63-65), and creating a nick which is sealed by LIG I to complete the repair process (46). While the “hit and run” pathway is more efficient than the “strand displacement synthesis” pathway, certain DNA sequences, such as DNA repeats, are prone to formation of gaps larger than 1 nt during the repair process and repair
is forced into the “strand displacement synthesis” pathway. In addition, while SP-BER is certainly more efficient that LP-BER, stalling of repair, particularly at the excision step when the lesion is removed, can shift repair into the LP-BER pathway.

The proper coordination of BER protein enzymatic activities and between BER proteins and cofactors is critical for efficient repair and maintenance of genome integrity (46, 66-68). For example, the DNA glycosylase that removes the base lesion remains bound to the DNA until displaced by APE1 (51). Following lesion removal by the DNA glycosylase, APE1 has a rapid incision rate followed by a slow product-release step, which allows APE1 to directly pass the intermediate along to the next protein in the sequence of repair and prevents the accumulation of potentially cytotoxic DNA breaks (56, 69, 70). The coordination between the activities of FEN1 and pol β has also been found to be vital, as FEN1 cleavage is dependent upon pol β strand displacement synthesis, and pol β strand displacement synthesis has also been found to be stimulated by FEN1, indicating their functional cooperation (63). In fact, all of the core BER proteins have been found to contain disordered regions that have been shown to interact with other proteins (51). The coordination of BER proteins with repair cofactors is equally important. Poly (ADP-ribose) polymerase 1 (PARP1) acts as a nick surveillance protein and coordinates with FEN1 to activate LP-BER and stimulate strand displacement synthesis (71). When a stall occurs in the excision step of SP-BER, PARP1 binds the nick and may activate BER by recruiting other BER proteins to initiate the long-patch pathway (71). Thus, PARP1 plays a critical role in determining the proper BER pathway for repair. Likewise, proliferating cell nuclear antigen (PCNA) has been shown to be involved in facilitating the steps of LP-BER by stimulating the replicative polymerases and FEN1 (45, 72-74). Proliferating cell nuclear
antigen may also facilitate cooperation between repair enzymes as evidenced by the direct interaction of PCNA with all of the core BER proteins, including FEN1, pol β, and LIG I (46, 73-75). Other cofactors have also been found to facilitate more efficient repair. X-ray repair cross-complementing protein 1 physically interacts with pol β and enhances the efficiency of DNA LIG III during BER (76-78). High mobility group box 1 (HMGB1) stimulates both pol β synthesis and FEN1 cleavage activities (46, 79). If the coordination between BER proteins and cofactors is lost, there can be adverse consequences for the cells. Unrepaired AP sites are deleterious, as they lack genetic coding information and can block replication forks and lead to DNA strand breaks (80, 81). As mentioned previously, if DNA repair fails, either because of insufficiency of repair proteins or interruption of protein coordination, the accumulation of damage can have severe consequences for human health, including neurodegeneration, premature aging, and cancer (31, 51, 68, 82).

Furthermore, BER can also affect human health through its role in active DNA demethylation. Gene expression is in part controlled by the methylation pattern of CpG islands in the promoter region of genes which govern the ability of transcription factors to bind. By actively demethylating CpG islands in the promoter region of genes, BER can affect gene expression. The modified base 5-hydroxymethylcytosine (5hmC) is formed from oxidation of 5mC by the Ten-Eleven Translocation (TET) oxidases, and can be converted to 5-hydroxymethyluracil (5hmU) by the Activation-Induced Deaminase (AID)/apolipoprotein B editing complex (APOBEC) proteins, making it a substrate for the glycosylase Single-Strand-Selective Monofunctional Uracil-DNA Glycosylase 1 (SMUG1). In addition, the AID/APOBEC proteins can also deaminate 5mC to T, which is a substrate for Thymine DNA glycosylase (TDG). The removal of these lesions by their
respective glycosylases initiates BER, resulting in repair of the lesions and replacement of
the 5mC with an unmodified C, thus resulting in demethylation \((51)\) that can affect gene
transcription and expression.

Base excision repair takes on many roles in influencing genome and epigenome
integrity, stability, and gene expression, and efficient BER is critical for the maintenance
of human health. In the remaining sections, the role of BER in mediating the instability of
DNA trinucleotide repeats will be discussed, as it relates to the onset of human
neurodegenerative disease and cancer.

**B. TRINUCLEOTIDE REPEAT INSTABILITY AND HUMAN DISEASE**

Trinucleotide repeats (TNRs) are a class of microsatellite DNA made up of tandem
repeats of 3 nucleotides. The number of repeats in a TNR tract is highly polymorphic in
the human genome. The TNRs are prone to instability in the length of the repeat tract in
the form of both expansions and deletions \((83, 84)\). Expansions of TNRs were first
recognized as a cause of human disease in 1991 when CGG/CCG expansions in the 5’-
untranslated region (UTR) of the Fragile X mental retardation 1 (FMR1) gene were shown
to be responsible for Fragile X syndrome (FXS) \((85, 86)\) and expansions of CAG/CTG
repeats in the coding region of the gene for the androgen receptor were found to be
responsible for spinal-bulbar muscular atrophy (SBMA) \((87)\). Since then, more than 40
human neurodegenerative diseases have been shown to be associated with TNR
expansions, including Huntington’s disease (HD) (CAG/CTG), Friedreich’s ataxia
(FRDA) (GAA/TTC), and myotonic dystrophy (MD) (CTG/CAG) \((88-92)\). While all
individuals contain some repeats in the affected regions, each disease has a threshold
number of repeats for TNR expansion, which is the number of repeats (typically 30-35 repeats) at which the rate and probability of expansions greatly increases, leading to disease (93, 94). A longer repeat tract is associated with earlier onset and more severe disease symptoms, a phenomenon known as anticipation (89, 90, 95), which means that as a TNR tract expands, the likelihood of further expansion increases. Expansions of TNRs can occur in the coding or non-coding regions of the DNA, as TNR instability in the coding region, 5’-UTR, 3’-UTR, intron, and promoter region of genes have all been associated with human disease (89) (Figure 1.5). For example, Huntington’s disease is caused by CAG expansions in the encoding region of the gene for huntingtin protein, which leads to expression of a mutant protein with an expanded polyglutamine (polyQ) tract. The longer polyQ tract results in aggregation and precipitation of protein, leading to neuronal death and neurodegeneration (93, 94). Expansions that occur in the coding region of genes tend to be more modest, with an increase of less than 10 repeat units per generation of cells (94-96). On the other hand, expansions in the non-coding regions can be just as deleterious. For example, Friedreich’s ataxia is caused by GAA expansions in the first intron of the frataxin (FXN) gene, which results in disruption of gene transcription and suppression of FXN protein expression. Because the FXN protein plays a crucial role in cellular energy metabolism, the reduction of FXN levels creates a metabolic deficiency that results in neuronal death and neurodegeneration (93, 97, 98). Expansions in the non-coding regions are typically much larger and can increase by 100-100,000 repeat units (94-96). It is also important to note that TNRs also experience deletions, which are associated with certain types of cancer. For example, deletions in the CAG repeat tract in the coding region of the gene for androgen receptor are associated with prostate and ovarian cancers (99, 100).
Thus, proposed mechanisms for generating TNR instability in cells must account for both their expansion and deletion.

The mechanisms of generating TNR instability are not well understood, however, changes to TNR length have been found to be mediated by the formation of non-B form DNA secondary structures that are made possible by the self-base pairing ability of the C and G rich TNRs through both traditional Watson-Crick base pairing and Hoogsteen base pairing \((21, 26, 27)\). The TNRs form small loops, hairpins, triplexes, tetraplexes such as G-quadruplexes, and sticky DNA \((101, 102)\) (Figure I.6). The longer the repeat tract, the more stable the secondary structures that can form, contributing to further instability of the TNR tract \((103, 104)\). Both CAG and the complementary CTG repeats are able to form hairpins, with the CTG hairpin being the more stable of the two. In addition, both CGG and the complementary CCG repeats can also form hairpins, with the CGG hairpin being more stable and capable of additionally forming G-quadruplexes. GAA/TTC repeats are able to form triplexes \((105)\). In addition to secondary structures composed purely of DNA, the

Figure I.5 TNR expansions in coding and non-coding regions of DNA lead to human disease \((89)\)
formation of R-loops also plays a role in TNR-instability related diseases. G-rich regions of the DNA such as CAG/CTG and CGG/CCG repeats are able to form these DNA:RNA hybrids during transcription while the DNA is unwound (105). Formation of secondary structures, such as hairpins, in vivo has been experimentally confirmed by detecting inhibition of repair by hairpin structures during meiotic recombination and by using synthetic zinc-finger nucleases specific for CAG/CTG hairpin structures in human cells (106-108).
These non-B form DNA secondary structures are able to form during DNA metabolic processes including DNA replication, repair, and recombination as well as gene transcription (89, 95). The secondary structures occur when a nick is formed in the backbone of the DNA that allows DNA slippage and realignment, allowing the formation of self-base pairs within strands of the DNA (105). Both expansions and deletions have been found to result from secondary structure formation (89, 90, 109). When a secondary
structure, such as a hairpin, forms in the newly synthesized strand, this can result in TNR expansions when the extra repeats become incorporated into the genome. When secondary structures form in the template strand, deletions can occur as a result of polymerases skipping over the structure during synthesis (110, 111) (Figure I.7). The deletions have also been found in studies of bacteria and yeast (112, 113). DNA secondary structures can also be hotspots for the formation of further DNA damage as they expose ssDNA regions to DNA damaging agents. They can further cause remodeling the chromatin structure, which also makes the DNA more accessible to damaging agents and leads to further damage (114).

![Figure I.7 The location of a TNR hairpin governs instability (89)](image)

Longer repeats have also recently been found to be associated with human disease, including ATTCT expansions, which cause spinocerebellar ataxia type 10 (SCA10) (115) and G\textsubscript{4}C\textsubscript{2} expansions which are the most common cause of familial amyotrophic lateral sclerosis (ALS) (116). However, the majority of repeat expansion diseases are caused by TNRs (105). Though TNR-expansion diseases are very debilitating and eventually fatal, there are currently no cures and only limited treatments for this family of diseases. Therefore, identifying the mechanisms that govern TNR instability will be critical for
finding ways to prevent and reverse TNR instability for the treatment and prevention of TNR instability-associated diseases.

C. BASE EXCISION REPAIR AND TRINUCLEOTIDE REPEAT INSTABILITY

The TNR instability via DNA replication has been the most studied pathway of TNR instability, however, TNR instability has more recently been found to occur in non-dividing somatic cells, indicating a role for other DNA metabolic processes (114, 117, 118). The TNR instability has since been found to result from any DNA metabolic process that creates regions of ssDNA, including DNA replication (119, 120), repair (46, 95), recombination (102), and gene transcription (121), due to the ability of ssDNA to form non-B form secondary structures. Secondary structures can block DNA and RNA polymerases during replication and repair or gene transcription, respectively (122, 123), allowing DNA slippage and misalignment. DNA repair has specifically been implicated in facilitating TNR instability because of the age-dependency seen in TNR expansions of non-dividing neurons (117, 118). The role of DNA repair in TNR instability is supported by the fact that TNR expansions seen in the neurons are relatively small (5-20 repeat units) (117, 118), whereas other processes that lead to TNR instability, such as DNA replication and recombination, usually result in much larger expansions (124). Many recent studies have indicated BER as the main pathway of mediating this age-dependent TNR instability (95, 117). The neurodegenerative symptoms associated with TNR expansion diseases can be accounted for by the fact that neurons are postmitotic and have high metabolic rates, resulting in increased amounts of oxidative damage and ssDNA breaks with age, which are subjected to BER (31, 125).
Base excision repair is the major pathway of repairing oxidative damage and ssDNA breaks. TNRs are hotspots for the formation of these lesions because they contain G-rich sequences, and the low redox potential of G makes it an ideal target for formation of oxidative damage (126). The first steps of BER include removal of a lesion by a specific glycosylase and APE1 5’-incision of the abasic site, which generates a ssDNA break (127) necessary for the formation of secondary structures (95). These secondary structures can cause pausing of the repair DNA polymerase and allow further DNA slippage (122). There are now many studies in a variety of organisms linking oxidative stress, BER, and TNR instability (16, 19, 117, 128-130). Increased levels of oxidative damage have been associated with TNR expansions (117, 130), indicating a role of BER in facilitating TNR instability in response to oxidative damage. Furthermore, CAG expansions in HD transgenic mouse models have been found to be dependent on OGG1, with knock out of OGG1 completely abolishing expansions, indicating that mammalian somatic expansions are initiated by an OGG1-mediated BER mechanism (117). Because OGG1 removal of a lesion initiates the BER process, this indicates a role for BER in mediating TNR expansions. In addition, loss of the endonuclease VIII-like 1 (NEIL1) glycosylase has also been shown to decrease somatic expansions in HD mouse models, further supporting the role of BER in TNR expansion (131). Other BER proteins and cofactors have also been found to modulate TNR instability in vitro (67) and in mouse neurons (132), including pol β and FEN1.

Decreased efficiency in any of the core BER protein activities results in decreased overall BER efficiency. A decrease in BER efficiency leads to accumulation of DNA damage and ssDNA breaks, which allow the further formation of DNA secondary
structures that lead to TNR instability. Pol β plays a crucial role in both SP-BER and LP-BER through its dRP lyase and gap-filling synthesis activities. Efficient FEN1 activity is also crucial for maintenance of TNR stability by removing TNR-containing flaps. A 5’-TNR flap can fold into a hairpin through self-base pairing, which inhibits FEN1 flap cleavage activity (133-135), preventing hairpin removal and allowing incorporation of extra repeats as expansions. Longer repeat tracts are able to form more stable hairpins that are resistant to FEN1 cleavage (133, 134), while short repeats form more transient secondary structures that can be resolved by FEN1 (136) as it captures and loads onto the 5’-end of a transiently formed single strand. Thus, the ability of FEN1 to cleave a TNR-flap is determined by the length of the repeats, the formation of secondary structures within the repeats, and coordination of FEN1 with other repair proteins (46). The formation of hairpins that prevent FEN1 cleavage at the flap base force FEN1 to carry out “alternate flap cleavage” in which FEN1 cleaves a short 5’-TNR flap attached to the hairpin (46, 67), leaving extra repeats that can be incorporated as an expansion. The interactions between BER proteins and their effect on TNR instability have been explored in a series of in vitro and cell-based experiments. Reconstituted BER with purified human proteins was used to explore the effect of BER on TNR stability in vitro (67). Trinucleotide repeat expansions were found to be dependent on pol β multi-nucleotide gap-filling synthesis to add extra repeats during LP-BER. The TNR expansions required FEN1 alternate flap cleavage to remove the lesion from the 5’-end of the downstream hairpin to create a ligatable nick for LIG I to seal, incorporating the hairpin as an expansion (67). These studies indicate that not only is BER involved in mediating TNR expansions, but that the coordination of BER protein activities is critical in the maintenance of TNR stability.
The proper coordination between BER proteins and cofactors has been found to be critical for the maintenance of TNR stability. Secondary structures, such as hairpins, can interrupt the coordination between proteins, disrupting repair and leading to TNR instability. Specifically, disruption of the coordination between pol β and FEN1 activities can promote LP-BER and multi-nucleotide gap-gilling synthesis by pol β. The promotion of LP-BER can lead to addition of extra repeats that results in expansions. In this expansion process, the weak processivity and strand displacement activity of pol β allows for strand slippage, misalignment, and hairpin formation (114) and leaves a gap that is filled in by pol β multi-nucleotide gap-filling synthesis. The gap-filling synthesis of pol β not only produces extra repeats, but also fills in the gap, preventing downstream TNRs from reannealing to the template and instead stabilizing the hairpin (46, 67). The stable hairpin further prevents FEN1 typical flap cleavage activity and forces alternate cleavage, removing fewer repeats and creating a ligatable nick for LIG I to incorporate the hairpin as an expansion (Figure I.8) (46, 67). Coordination of LIG I activity is also crucial, as LIG I competes with FEN1 to ligate the DNA, which then incorporates the hairpin as an expansion (137, 138). Figure I.8 shows the pathway by which loss of coordination between the BER proteins leads to TNR expansions (46). In this LP-BER pathway, damage occurs in a TNR tract which is subsequently removed by a DNA glycosylase. APE1 5’-incision results in formation of a ssDNA break and resulting 1 nt gap. If no strand slippage occurs, pol β fills in the gap and FEN1 cleaves the 5’-dRP and adjacent nucleotide. The “hit and run” mechanism is used to repair the DNA and no extra repeats are synthesized so that no TNR instability results. However, following APE1 5’-incision, if DNA slippage results in the formation of a hairpin, the coordination between pol β and FEN1 activities is disrupted.
Pol β multi-nucleotide gap-filling synthesis produces extra repeats, while FEN1 alternate flap cleavage removes fewer repeats and LIG I activity seals in the expansion (46). The imbalance of pol β and FEN1 activities leads to TNR instability. When pol β activity is greater than FEN1, more repeats are inserted by pol β than are removed by FEN1, leading to expansion. Conversely, when FEN1 activity is greater than pol β activity, more repeats are removed by FEN1 than are synthesized by pol β, leading to deletion. This model has been supported experimentally in HD mouse models, in which low expression of FEN1 was found to be associated with an increase in somatic CAG expansions (132). In addition, pol β has been found to be enriched at expanded and unstable CAG repeats, but not on more stable repeat regions (132), while deficiency of pol β and HMGB1 has been found to suppress CAG expansions (67). This evidence supports a model in which proper protein levels and coordination of repair protein activities is critical in maintaining TNR stability.
Each round of LP-BER in a TNR tract results in only a 1-5 repeat expansion as a result of pol β multi-nucleotide gap-filling synthesis (139, 140). However, approximately 50,000 DNA base lesions are estimated to be generated per cell per day (141). Thus, BER of these abundant lesions could lead to many repair processes that result in an accumulation of small expansions. Because the hairpins formed by CAG and CTG repeats are themselves hotspots for the formation of oxidative damage, hairpin formation could initiate a “toxic oxidation cycle” in which each round of lesion repair increases the opportunity for further hairpin formation and generation of additional lesions for further repair (Figure I.9) (142).

Figure I.8 Loss of coordination between BER proteins leads to TNR instability (46)
The hotspot occurs as the result of the high proportion of G’s found in the TNRs and the single stranded loop regions formed in the hairpins that make the DNA bases more accessible to DNA damaging agents. The cycle of damage, excision, and expansion escalates with age as a result of the age-dependent accumulation of oxidative lesions (117). Thus, the relatively modest expansions associated with each round of BER can account for the accumulation of expansions with age that lead to the onset of TNR-associated neurodegenerative diseases.

Figure I.9 Toxic oxidation cycle of BER in a TNR tract (117)
While many studies now indicate BER as a key process in mediating TNR expansions, BER has more recently been found to also play a role in inducing TNR deletions. Several DNA damaging agents such as chemotherapeutic drugs and environmental toxicants, including ethyl methanesulfonate, ethylnitrosourea, mitomycin C, cisplatin, and UV and ionizing radiation, that induce damage subjected to the BER pathway have been shown to induce deletions in human, bacterial, and mouse cells (16, 19, 143-145). The TNR deletions were found to result from agents that cause both oxidized and alkylated DNA damage and ssDNA breaks, indicating that BER is responsible for inducing these deletions, as it is the major repair pathway for removing these types of damage. The generation of damage by the alkylating chemotherapeutic agent temozolomide in human cells containing plasmids with CTG repeats and in FRDA patient lymphoblasts was found to result in large deletions of the repeats (146). These deletions are proposed to result from pol β bypass of a large template hairpin or loop in coordination with FEN1 alternate cleavage of a downstream TNR-flap (Figure I.10) (111, 147). Temozolomide-induced alkylated DNA damage was found to induce massive intronic GAA deletions, as well as limited expansions, in the FRDA patient lymphoblasts. The deletions occur via ssDNA break formation during BER of the alkylated lesions that allows DNA slippage and the formation of a TTC loop in the template strand that is bypassed by pol β during synthesis, along with formation of a long 5’-GAA downstream flap that is removed by FEN1, resulting in removal of many more repeats by FEN1 than were synthesized by pol β (146). The formation of TNR deletions via BER further indicates the possibility that chemotherapeutically-induced DNA damage could be used to induce large deletions of expanded repeats as a therapeutic treatment for TNR expansion-associated
diseases. Oxidative damage was also found to be able to induce deletions. 5’,8-cyclo-2’-deoxypurines (cdPus) are oxidized lesions which are repaired inefficiently and thus accumulate in the genome. Cyclodeoxyguanine (cdG) and cyclodeoxyadenine (cdA) contain an extra covalent bond between the C5 of the 2’-deoxyribose and the C8 of the purine creating a structural abnormality and preventing their removal by DNA glycosylases. The high proportion of Gs and As in TNRs make them hotspots for the accumulation of these lesions. The formation of cdA in the template strand was found to induce formation of a CAG loop which is skipped over by pol β during the synthesis step of BER in coordination with FEN1 removal of a downstream TNR-flap, resulting in deletions during BER (148). However, these deletions occurred only when the BER-initiated gap formed upstream or opposite to the cdPu and not when the gap occurred downstream of the lesion, indicating that lesion location also plays a role in determining TNR instability.
Further studies have also indicated that lesion location governs TNR instability. In general, structures forming in the newly synthesized strand have been found to induce expansions, while structures in the template strand induce deletions (89). *In vitro* studies exploring the effect of location of an abasic site in a 20 CAG repeat tract using reconstitution of BER with purified human proteins have shown that lesion location dictates the size and location of secondary structures that further govern whether expansions or deletions occur (147). Abasic lesions located at the 5’-end of a TNR tract were shown to be repaired with much lower efficiency than lesions located in the middle.
or at the 3’-end of a TNR tract (149). The difference in the location of secondary structures formed and the resulting imbalance in pol β and FEN1 activities determines whether expansions or deletions result. A base lesion at the 5’-end of the TNR tract was found to induce expansions as a result of more efficient pol β synthesis and less efficient FEN1 cleavage, resulting in more nucleotides synthesized than are removed. When a base lesion was located in the middle of the repeat tract, FEN1 activity was more efficient than pol β synthesis so that more nucleotides were removed than were added by pol β, resulting in deletion. Damage at the 3’-end of the TNR tract had no effect on TNR stability (147). The size of the hairpins can also govern TNR instability, as pol β was shown to be able to synthesize through almost the entirety of a small template hairpin causing only small deletions, whereas it skipped over almost the entirety of a large template hairpin (Figure I.10) (111). This indicates that, like expansions, deletions exhibit length dependency with longer repeats inducing larger deletions.

The molecular mechanisms underlying TNR instability via BER remain to be elucidated. Determining these mechanisms will be crucial in identification of novel targets for the therapeutic treatment of TNR instability-associated diseases. As lesion location has been found to play a role in governing TNR instability, determining how BER of lesions in different contexts of TNRs could aid in identification of mechanisms for removing expanded repeats for disease treatment. In the next section, BER of lesions located in the context of a TNR hairpin will be reviewed.
D. BASE EXCISION REPAIR IN A TRINUCLEOTIDE REPEAT HAIRPIN

Paradoxically, though BER has been shown to facilitate TNR instability, BER has also been found to modulate TNR instability by removing TNR secondary structures such as hairpins (95). Hairpins serve as hotspots of oxidative DNA damage and guanines located in the hairpin loop of a CAG hairpin have been shown to be particularly susceptible to formation of oxidized lesions (142, 150). Using peroxynitrite to induce oxidative damage, a (CAG)\textsubscript{10} hairpin was found to be much more susceptible to formation of 8-oxoG compared to a (CAG)\textsubscript{10} duplex (150). The susceptibility of guanines in the loop of the hairpin to oxidative damage was found to be structure specific, not sequence specific, as guanines in a loop of random DNA sequence were also hypersensitive to oxidation (150). Studies using chemical probes to determine DNA base accessibility to damaging agents have shown that bases located in a loop or bulge are preferentially modified because of their exposure to solution as compared to paired bases (150-153). Thus, oxidized lesions are likely to accumulate in hairpin loops as a result of the high proportion of guanines and easier exposure to damaging agents that make them hotspots for oxidative damage.

However, recent evidence indicates that damage in the hairpin loop may be more resistant to BER. The first step of BER is removal of the lesion by a DNA glycosylase. OGG1, the glycosylase responsible for removing 8-oxoG, was found to act as efficiently on a TNR duplex as on a duplex of random sequence (142), however, OGG1 acted on 8-oxoG in a hairpin loop at a rate ~700-fold slower than in a duplex (150). The resistance of damage in the hairpin loop to BER is proposed to exacerbate the toxic oxidation cycle. During the toxic oxidation cycle, duplex TNR tracts are oxidized and BER is initiated by OGG1 and APE1. The nicked DNA strand allows strand slippage and hairpin formation
which becomes ligated into the DNA as an expansion. The expanded region is then again subjected to oxidation and the cycle continues. Because the incorporated hairpins act as hotspots for further damage and yet are more resistant to BER, generation of TNR hairpins may exacerbate the toxic oxidation cycle as a result of accumulation of TNR hairpins containing oxidative damage (142) (Figure I.11). Furthermore, 8-oxoG has been found to be repositioned into the loop of a hairpin via dynamic intrastrand rearrangements of the DNA when the lesion occurs near the hairpin. The intrastrand rearrangements occur to minimize the thermodynamic impact of 8-oxoG and avoid the destabilization that occurs when 8-oxoG pairs with C (154). Differential scanning calorimetry was used to show that there is no significant change in thermodynamic parameters when 8-oxoG is located in the loop as compared to a non-damaged G (154), further supporting the claim that TNR hairpins containing oxidized lesions in the loop of the hairpin would accumulate in TNR tracts.
As these studies have suggested that hairpins containing 8-oxoG in the hairpin loop are likely to accumulate in TNR tracts, a series of in vitro experiments were carried out to test the ability of BER to remove 8-oxoG from the hairpin loop which is in the context of a duplex TNR tract. It was found that BER of 8-oxoG in a CAG hairpin loop led to hairpin removal and the prevention of TNR expansion (155). Using reconstitution of BER with purified human BER proteins and using the single strand specific Mung Bean Nuclease to probe for ssDNA regions, it was shown that OGG1 was able to remove 8-oxoG from the hairpin loop followed by APE1 5’-incision of the abasic site, which converted the hairpin

Figure I.11 Resistance of lesions in a TNR hairpin to BER exacerbates the toxic oxidation cycle (142)
into a “double-flap intermediate” of BER containing an upstream 3-flap and a downstream 5’-flap (Figure I.12) (155).

Figure I.12 BER of an oxidized lesion in the loop of a CAG hairpin is coupled to hairpin removal (155)

The double-flap repair intermediate was resolved by FEN1 cleavage of the downstream 5’-flap in coordination with cleavage of the upstream 3’-flap by the structure specific 3’-5’ endonuclease Mus81/Eme1 complex which is typically involved in the resolution of DNA Holliday Junctions. The resolution of the repair intermediate resulted in complete or partial removal of the hairpin and prevention or attenuation of TNR expansions (Figure I.12) (155). The findings indicate that BER of a lesion in the hairpin loop can be coupled to hairpin removal. The removal of TNR hairpins via BER could also implicate a method of
removing expanded repeats by inducing DNA damage as a mechanism of treating TNR expansion-associated disease. The results further support a model in which lesion location plays a major role in governing TNR stability by determining whether an expansion or deletion occurs.

Other 5’-3’ nucleases such as the NER protein xeroderma pigmentosum complementation group G (XPG) (156) and 3’-5’ nucleases such as the Mre11-Rad50-Nbs1 (MRN) complex (124), which is involved in DSBR, have also been shown to facilitate the removal of TNR hairpins both in vitro and in vivo, supporting a model of hairpin removal by cleavage of a 3’-flap and 5’-flap, and further indicates that other nucleases may play a role in facilitating resolution of the double-flap intermediate and removal of a TNR hairpin. Resolution of the double-flaps involves flap equilibration. Flap equilibration is the ability of a flap to partially or completely reanneal to the template strand, which can allow the formation of a variety of double-flap intermediates (Figure I.13) (157). Flap equilibration plays an important role in preventing TNR expansions. FEN1 was shown to be able to capture a 5’-flap among equilibrating intermediates in order to make efficient cleavage, and is actually more efficient at cleaving equilibrating flaps than non-equilibrating flaps, indicating that this is a normal part of FEN1’s cleavage mechanism (157).
Small hairpins can be resolved by FEN1 alone via flap equilibration that allows the upstream flap to reanneal and forms a downstream 5'-flap that FEN1 can remove, however, FEN1 is much less efficient at removing larger hairpins and longer flaps, so coordination with other repair proteins may be much more critical for removal of a large hairpin (155). The involvement of Mus81/Eme1, XPG, and the MRN complex indicates that crosstalk between BER and other repair pathways occurs in order to maintain TNR stability in vivo. Crosstalk between repair pathways will be discussed in the following section.

**E. CROSSTALK BETWEEN BASE EXCISION REPAIR AND OTHER DNA REPAIR PATHWAYS**

Several lines of evidence suggest that other DNA repair pathways, such as MMR, may crosstalk with the BER machinery to repair difficult lesions. While DNA glycosylases are well established as the proteins that carry out the first step of BER to remove the base lesion, glycosylases have been shown to crosstalk with both MMR and NER proteins. The MMR system utilizes two types of heterodimeric MutS homologue (MSH) complexes to
initiate the repair process. MutSα is composed of MSH2/MSH6 and MutSβ is composed of MSH2/MSH3, with the two complexes having differences in substrate preference (158-161). MSH2/MSH6 prefers single base mismatches and single or double base insertions, while MSH2/MSH3 prefers single base insertions and small extrahelical loops (158, 159, 162-165). The MSH2/MSH6 and MSH2/MSH3 complexes carry out the initial detection step of MMR, and then recruit a MutL complex, such as the MutL homology (MLH1)/PMS2 complex, to form a ternary complex that removes the mismatch and initiates repair (105, 166). Several glycosylases involved in BER have been found to also crosstalk with this MMR machinery (Figure I.14). The MutY homologue (MYH) glycosylase removes adenine from a newly synthesized strand when paired with 8-oxoG (167), resulting in an 8-oxoG:C base pair that can be repaired by OGG1 (168). In yeast lacking OGG1, the MSH2/MSH6 complex has been shown to suppress G:C to T:A transversion mutations, suggesting that MMR may be able to compensate for OGG1 in the removal of 8-oxoG (169). Further evidence also suggests that OGG1, MYH, and MMR may cooperate to repair oxidative damage (168). In mammalian cells, the absence of OGG1 or MSH2 results in the accumulation of 8-oxoG (168, 170, 171). Deficiencies in MLH1 or MSH2 in tumor cells and mouse embryonic fibroblasts (MEFs) were also shown to result in increased 8-oxoG levels following DNA damage events (172, 173). Supporting the idea that MYH and the MMR machinery cooperate to repair 8-oxoG, MSH2/MSH6 was found to interact with MYH through the C-terminal domain of MSH6 (174). Crosstalk between DNA glycosylases and MMR has also been implicated to play a role in the repair of alkylated DNA damage (175). Cell lines that are deficient in the MMR proteins MSH2/MSH6 are hypersensitive to methylating agents (176). In addition, both Uracil-N-
glycosylase (UNG), which removes uracil, and Thymine DNA glycosylase (TDG) have been implicated to crosstalk with MMR proteins (168). TDG repairs O6-meG, yet MMR proteins have been found to also bind O6-meG when it is paired with T, which could suggest that TDG and MMR are able to direct distinct pathways depending on the context of the lesion, with TDG and BER repairing O6-meG paired with C and MMR repairing O6-meG paired with T (150). Thus, BER and MMR may cooperate to repair DNA base lesions, and crosstalk between BER and MMR proteins may help resolve lesions that are resistant to repair.
There is also accumulating evidence that MMR proteins play a role in mediating TNR instability, indicating a role for the crosstalk between BER and MMR for preventing TNR instability. Knockout of MSH2 or MSH3 was found to greatly decrease the frequency of TNR expansions in several TNR expansion disease mouse models (177-180), indicating...
that the MMR proteins play a role in promoting TNR expansions. In a transgenic HD mouse model, CAG expansion was found to be dependent on MSH2/MSH3, and expansion was completely abolished in the MSH2 or MSH3 knockout mice in both the somatic and germ cells (177, 178, 181-184). MSH2/MSH3 has also been found to be required for expansion of CTG repeats in a mouse model for MD (179, 180, 185-187), the GAA repeats in the FXN gene associated with FRDA (105, 188-192), and the CGG repeats associated with FXS (193). The MSH3 protein was found to be particularly important for the progressive, incremental (as small as 1 repeat unit) expansions that promote the transition into the pathogenic range of TNR length, while knockout of MSH3 resulted in more stable repeats in a yeast system (194). The loss of other MMR proteins such as PMS2 and MLH-related endonucleases have also been found to suppress TNR expansions (195). MLH1, which is involved in all 3 mammalian MutL complexes, and MLH3 were found to be involved in somatic TNR expansions in HD mouse models (195, 196). The protein PCNA acts as a BER cofactor to stimulate FEN1 cleavage and further interacts with all core BER proteins. PCNA also plays role in MMR and has further been shown to be able to load onto small CAG and CTG loop outs of 2-3 repeats, thus, PCNA loading may be facilitated in the presence of small hairpins (105) and may be able to facilitate crosstalk between BER and MMR for more efficient repair. This evidence further supports a model in which BER and MMR are jointly involved in mediating TNR instability.

Mismatch repair has been found to modulate TNR instability by removing small 1-3 repeat unit CAG loops (177, 178, 197). However, as discussed above, the MMR proteins also play a key role in mediating TNR expansions. To explore this phenomenon, in vitro studies were used to test the binding and activity of the MSH2/MSH3 complex on TNR
tracts and structures. MSH2/MSH3 was found to be trapped by hairpin structures \((181, 198)\), and to then stabilize the hairpin and promote incorporation of the hairpin as an expansion (Figure I.15). So, while small loops are easily repaired by the MMR machinery, hairpins are stabilized by the trapped complex \((106)\).

Figure I.15 MSH2/MSH3 stabilizes a TNR hairpin \((89)\)
The MSH proteins bind and hydrolyze adenosine triphosphate (ATP). Hydrolysis of ATP is critical for recognition and repair of DNA, and loss of hydrolysis activity leads to mutation (199). Purified proteins and synthetic DNA were used to show that the A:A mismatch that occurs at every third repeat in the hairpin stem (200) recruits the MSH2/MSH3 complex and inhibits the ATPase activity of the protein complex, while also altering the ATP/adenosine diphosphate (ADP) binding affinity as well as the binding affinity of the protein complex to the DNA (181), which prevents the complex from being able to slide off the hairpin and prevents repair of the hairpin by MMR. The inability of the MSH2/MSH3 complex to repair the hairpin therefore facilitates expansions, in contrast to a small loop or small unstable hairpin that is efficiently removed to prevent instability (197). The model explains how MSH2/MSH3 can promote TNR expansions through repeated cycles of damage and BER, by stabilizing the hairpin and preventing removal of extra repeats (Figure I.16).

Figure I.16 MSH2/MSH3 stabilization of a TNR hairpin promotes TNR expansions (95)
A molecular mechanism was recently proposed to explain how crosstalk between BER and MMR leads to TNR expansions during DNA repair (Lai et al., in press, *Nature Comm* DOI: 10.1038/ncomms12465). The Pol β protein is inefficient at synthesizing repeats and tends to skip over a template hairpin or loop, leading to deletions (111, 146-148). Using purified human proteins and synthetic DNA, MSH2/MSH3 was found to stimulate pol β to synthesize through a template hairpin or loop and enhance the formation of a downstream 5’-flap precursor for expansion (Lai et al., in press, *Nature Comm* DOI: 10.1038/ncomms12465). The promotion of pol β activity by MSH2/MSH3 was found to alter the outcome of TNR instability, shifting the outcome from deletions to expansions during the repair of oxidized lesions. MSH2/MSH3 was found to form a physical complex with pol β to stimulate pol β synthesis through the repeats, and DNA damage was found to promote recruitment of this pol β-MSH2/MSH3 complex. Thus BER and MMR form a novel hybrid pathway of repair during the removal of oxidized bases from TNR tracts (Figure I.17). These findings suggest that crosstalk occurs between repair pathways to share machinery when a canonical repair pathway cannot remove a lesion. In this case, the crosstalk between BER and MMR contributes to expansions during repair in a TNR tract.
Thus, BER alone or in crosstalk with other repair pathways can promote both TNR expansions and deletions.

The molecular mechanisms by which BER mediates TNR instability, as well as mechanisms by which cells prevent TNR instability, remain to be elucidated. Evidence shows that the position of the lesion within a TNR tract governs the type of instability that results. While both oxidative and alkylated DNA damage have been shown to be involved in mediating TNR instability, whether the type of lesion has any effect on TNR instability is yet to be explored. Understanding how lesion location and lesion type govern TNR

**Figure I.17 Crosstalk between BER and MMR promotes expansions** (Lai et al., in press, *Nature Comm* DOI: 10.1038/ncomms12465)
instability will help identify new targets and pathways for the prevention and treatment of TNR instability-associated diseases, for which there is currently no treatment. In my Ph.D. dissertation research project, I explored the molecular mechanisms by which DNA damage and BER in different contexts of TNRs modulate TNR instability. In the three chapters of this dissertation, I have explored how BER proteins and cofactors cooperate to prevent TNR expansions and how the type of DNA lesion affects TNR instability during the repair process. In CHAPTER 1, I studied the cooperation between key BER enzymes that promote removal of a TNR hairpin during BER in the hairpin loop to prevent or attenuate TNR expansions. In CHAPTER 2, I explored the role of the BER cofactor PCNA in facilitating TNR deletion, TNR hairpin removal, and prevention of TNR expansions during BER in a TNR tract. Finally, in CHAPTER 3, I investigated the effect of lesion type on the efficiency and outcome of BER during repair in a TNR hairpin and TNR duplex.

OVERVIEW

TNR expansions are associated with over 40 human neurodegenerative diseases, including Huntington’s disease and myotonic dystrophy, while TNR deletions are linked to certain types of cancers. TNR instability has been found to result from DNA metabolism within the repeat tracts, including DNA replication and repair, recombination, and gene transcription. Base excision repair has been implicated as a major pathway of TNR instability in postmitotic neural cells. It is important to identify the molecular mechanisms of TNR instability during DNA repair in order to identify novel targets for the therapeutic treatment and prevention of TNR instability-associated diseases for which there are
currently no cures and only limited treatments. While lesion location has been shown to govern TNR instability, how the type of lesion affects instability of the repeats and the molecular mechanisms by which TNR hairpins are removed in cells to prevent expansions are yet to be determined. Identifying pathways used by cells to prevent expansions could help identify a new strategy for the treatment of TNR expansion-associated neurodegenerative diseases. Previous studies have shown that BER of DNA damage in the loop region of a TNR hairpin can lead to hairpin removal. The interactions between BER proteins and cofactors that facilitate hairpin removal need to be determined to identify novel therapeutic targets. This dissertation project explored the molecular mechanisms underlying TNR hairpin removal during BER through the cooperation of BER proteins and cofactors to prevent TNR expansions, and further explores how the type of DNA lesion affects BER in a TNR tract. In CHAPTER 1, exploration of the cooperation between the activities of the key BER enzymes allowed us to establish a model for the prevention and attenuation of TNR expansions in a CAG tract by resolution of a BER intermediate via the coordinated actions of APE1, FEN1, pol β, and LIG I. We were able to identify a novel and significant role for the APE1 3’-5’ exonuclease that had previously remained enigmatic. In CHAPTER 2, we examined the role of the BER cofactor PCNA in preventing TNR expansions during BER in a CAG repeat tract. We showed that PCNA was able to facilitate removal of a TNR hairpin during BER by stimulating FEN1 flap cleavage activity and facilitating pol β synthesis activity by altering the structure of dynamic DNA, and to facilitate CAG deletions during repair in a CAG duplex tract by stimulating LIG I across a small template loop. Our study provided the first evidence of a BER cofactor facilitating TNR deletions and TNR hairpin removal to prevent TNR expansions. In CHAPTER 3, we
investigated the effect of modifications of the sugar residue in an abasic site on BER efficiency in a TNR tract and found that an oxidized abasic lesion inhibits efficient BER as result of crosslink formation between the oxidized sugar and pol β that prevents efficient FEN1 cleavage and pol β strand displacement synthesis activities, thereby resulting in accumulation of DNA strand breaks and contributing to genomic instability.
CHAPTER 1: AP ENDONUCLEASE 1 PREVENTS TRINUCLEOTIDE REPEAT EXPANSION VIA A NOVEL MECHANISM DURING BASE EXCISION REPAIR

ABSTRACT

Base excision repair (BER) of an oxidized base within a trinucleotide repeat (TNR) tract can lead to TNR expansions that are associated with over 40 human neurodegenerative diseases. This occurs as a result of DNA secondary structures such as hairpins formed during repair. We have previously shown that BER in a TNR hairpin loop can lead to removal of the hairpin, attenuating or preventing TNR expansions. Here, we further provide the first evidence that AP endonuclease 1 (APE1) prevented TNR expansions via its 3’-5’ exonuclease activity and stimulatory effect on DNA ligation during BER in a hairpin loop. Coordinating with flap endonuclease 1, the APE1 3’-5’ exonuclease activity cleaves the annealed upstream 3’-flap of a double-flap intermediate resulting from 5’-incision of an abasic site in the hairpin loop. Furthermore, APE1 stimulated DNA ligase I to resolve a long double-flap intermediate, thereby promoting hairpin removal and preventing TNR expansions.

INTRODUCTION

The expansion of trinucleotide repeats (TNR) is associated with over 40 human neurodegenerative diseases including Huntington’s disease (CAG/CTG), Friedreich’s ataxia (GAA/TTC), and myotonic dystrophy (CTG/CAG) (84, 95), among others. The repeat instability results from non-B form DNA secondary structures including hairpins, triplexes, tetraplexes, and sticky DNA (102) formed during DNA replication, repair,
recombination, and gene transcription (89, 101, 108). Recent studies have shown that the repair of oxidative DNA damage within trinucleotide repeats is associated with somatic TNR instability (46, 67, 111, 117, 147). During the repair of oxidized base lesions such as 8-oxoguanine (8-oxoG), an abasic site is incised at its 5’-side generating a nick in the DNA backbone. If this occurs in trinucleotide repeats, it allows the formation of hairpins in the DNA, which leads to repeat expansions and deletions if the hairpins are incorporated into the DNA and sustained upon completion of repair (111, 147). It has been proposed that repeated cycles of oxidative DNA damage and BER lead to cumulative repeat expansions through a ‘toxic oxidation cycle’, which can result in the onset of disease caused by TNR expansion (95, 117, 142).

We previously reported that the location of a DNA base lesion within a TNR repeat tract determines whether an expansion or deletion occurs, with a lesion near the 5’-end leading to expansions, and a lesion in the middle leading to deletions (147). An 8-oxoG located in the hairpin loop has been found to be resistant to DNA repair (142), resulting in the accumulation of the lesion and sustainment of the hairpin loop that leads to repeat expansion. However, we have discovered that during BER in a TNR hairpin loop, the hairpin can be converted to a double-flap intermediate, containing an upstream 3’-flap and a downstream 5’-flap, which can be subsequently resolved by a 3’-5’ endo/exonuclease, Mus81/Eme1, and flap endonuclease 1 (FEN1), respectively. This subsequently results in hairpin removal and prevention or attenuation of TNR expansions (155).

AP endonuclease 1 (APE1) is a multi-functional protein abundant in human cells (82, 201-203) that is essential in maintaining multiple cellular functions. This is demonstrated by the lethality of APE1 knockout in mice (204). The enzyme can
endonucleolytically incise the 5’-end of an abasic site in duplex DNA and in single-stranded DNA (205), including a TNR hairpin loop (155, 205), by hydrolyzing the phosphodiester bond during BER. In addition, APE1 can promote the recycling of DNA glycosylases by dislodging the glycosylase from an abasic site, resulting in stimulation of the activity of the enzymes such as 8-oxoguanine DNA glycosylase 1 (OGG1) (206, 207) which efficiently initiates BER (59, 61, 208). APE1 can also stimulate the activity of FEN1 via a physical protein-protein interaction and the activity of DNA ligase I (LIG I) (209, 210), as well as stimulate the activities of both the polymerase and deoxyribose phosphate (dRP) lyase activities of polymerase β (pol β) (66) by facilitating pol β binding to a gapped DNA and dRP residue (211), thereby enhancing the efficiency of BER (66, 209, 211, 212). Furthermore, APE1 can physically interact with proliferating cell nuclear antigen (PCNA) (210), suggesting its involvement in BER coordination (56, 213). In addition, APE1 has 3’-5’ exonuclease activity (214) that is 100-fold less efficient than its AP endonuclease activity (82). The exonuclease activity of APE1 can be regulated by poly(ADP-ribose) polymerase-1 (PARP-1), as these proteins have been shown to compete for binding to the 5’-end of a gapped intermediate during BER (215). However, the competition between PARP-1 and APE1 in binding to BER intermediates stimulates APE1 3’-5’ exonuclease activity, suggesting that APE1 is dislodged by PARP-1 from the 5’-end of a BER intermediate, allowing its binding to the 3’-end of a BER intermediate on which the exonuclease activity can act (215). It has been shown that APE1 3’-5’ exonuclease activity removes a 3’-mismatch and may provide proof-reading activity for pol β during BER (216-218), indicating a crucial role of APE1 3’-5’ exonuclease in maintaining genome integrity.

The APE1 exonuclease activity has also been found to be able to remove 3’-blocking
groups such as L-configuration nucleoside analogs (219), suggesting its importance in trimming the “dirty 3’-end” of a BER intermediate to facilitate the completion of BER. However, the role of the exonuclease activity of APE1 in sustaining genome stability remains unknown. We previously showed that the 3'-5' flap endonuclease activity of the Mus81/Eme1 complex promotes the removal of a TNR hairpin by cleaving the 3'-flap of a double-flap intermediate generated by BER in a TNR hairpin loop (155). Yet, no BER protein that can process a 3'-5' flap during BER in a TNR hairpin has been identified. Because APE1 exhibits 3'-5' exonuclease activity that can process mismatches at the 3’-end of an upstream DNA strand, we hypothesized that the 3'-5' exonuclease activity may shorten the upstream 3'-flap of a double-flap intermediate resulting from the incision of the loop region of a TNR hairpin during BER. This could subsequently promote the removal of the hairpin, thereby facilitating prevention or attenuation of TNR expansions. We tested this hypothesis by reconstituting BER on both TNR hairpins containing a base lesion in the loop region and on the double-flap intermediates that are generated during BER. We found that APE1 promoted the removal of a TNR hairpin during BER of a base lesion in the hairpin loop region. This was accomplished by the 3’-5’ exonuclease activity of the enzyme that cleaved the upstream 3'-region exonucleolytically, resolving the double-flap intermediate and preventing TNR expansions. Surprisingly, we also found that APE1 significantly stimulated the ligation activity of LIG I to specifically facilitate the completion of hairpin removal. This is the first evidence of APE1 preventing TNR expansions by facilitating hairpin removal.
MATERIALS AND METHODS

Materials

The DNA oligonucleotides which contain an 8-oxoG were synthesized by Eurofins MWG Operon (Huntsville, AL, USA), and all others were synthesized by Integrated DNA Technologies (IDT, Coralville, IA, USA). Deoxynucleoside 5’-triphosphates (dNTPs) were from Fermentas (Glen Burnie, MD, USA). Terminal deoxynucleotidyl transferase and T4 polynucleotide kinase were purchased from Thermo Fisher Scientific (Waltham, MA, USA). Radionucleotides \[\gamma^{-32}\text{P}]\ ATP (6000 mCi/mmol) and Cordycepin 5’-triphosphate 3’-[\alpha^{-32}\text{P}] (5000 mCi/mmol) were purchased from Perkin Elmer Inc. (Boston, MA, USA). Micro Bio-Spin 6 chromatography columns were purchased from Bio-Rad (Hercules, CA, USA). All standard chemical reagents were purchased from Sigma-Aldrich (St. Louis, MO, USA) and from Thermo Fisher Scientific (Waltham, MA, USA). Purified OGG1, pol β, and LIG I were generous gifts from Dr. Samuel H. Wilson at the National Institute of Environmental Health Sciences/National Institutes of Health. APE1 and FEN1 were expressed in E. coli and purified as described below. Oligonucleotide substrates were prepared as described previously (155). Briefly, substrates containing a (CAG)₇ or (CAG)₁₄ hairpin were constructed by annealing a damaged strand containing an 8-oxoG or tetrahydrofuran (THF), an analogue of an abasic site, in the loop-forming region of a (CAG)₁₃ or (CAG)₂₀ tract with a template strand containing (CTG)₇ repeats at a molar ratio of 1:2. Substrates mimicking the double-flap intermediates with a 5’-sugar phosphate residue, the THF residue, were constructed by annealing an upstream primer containing a 3’-(CAG)₄ or (CAG)₇ flap and a downstream primer containing a 5’-(CAG)₃ or (CAG)₇
flap with the template strand at a molar ratio of 1:2:2. For each substrate, three CAG repeats located at both the 5’- and 3’-side of the hairpins/flaps, base paired with three CTG repeats in the template strand. Oligonucleotide sequences were described previously and are included below (155) (Table 1). Substrates were labeled with $^{32}$P at the 5’- or 3’-end of the damaged strand, or at the upstream primer or downstream primer, as indicated.
### Table 1 Oligonucleotide sequences (155)

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

F. tetrahydrofuran, THF 8-oxoG: 8-oxoguanine
Protein expression and purification

FEN1 was expressed in *E. Coli* BL21(DE3). Two liters of lysogeny broth (LB) medium were inoculated with one colony each of the transformed BL21(DE3) cells and were incubated overnight without shaking. The cells were then incubated at 37°C at 225 rpm until OD$_{600}$ reached 0.6. The protein expression was induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 3.5 hours and harvested by centrifugation at 2600 rpm for 45 minutes at 4°C. Cells were lysed with a French press cell disruptor (Glen Mills, Clifton, NJ, USA) in lysis buffer which contained 30 mM 4-(2-hydroxyethyl)-1-piperazineethane-sulfonic acid (HEPES), pH 7.5, 30 mM KCl, 1 mM dithiothreitol (DTT), 1 mM ethylenediaminetetraacetic acid (EDTA), 1 mM phenylmethysulfonyl fluoride (PMSF), and 0.5% inositol. The cell lysates were centrifuged at 12,000 rpm for 30 minutes at 4°C. The supernatant was loaded onto a 10-ml Sepharose Q column operated by an AKTA Fast Protein Liquid Chromatography system (FPLC) (GE Healthcare, Piscataway, NJ, USA). The flow-through was collected and dialyzed into buffer containing 30 mM HEPES, pH 7.5, 30 mM KCl, 0.5% inositol, and 1 mM PMSF, and subsequently loaded onto a 5-ml CM sepharose column (Bio-Rad, Hercules, CA, USA), with fractions eluted using a linear gradient of KCl from 30 mM to 2 M. Peak fractions were combined and dialyzed into buffer containing 30 mM HEPES, pH 7.5, 0.5% inositol, 1.7 M (NH$_4$)$_2$SO$_4$, and 1 mM PMSF. Samples were then loaded onto a 2-ml phenyl sepharose column, with fractions eluted using a linear gradient of (NH$_4$)$_2$SO$_4$ from 1.7 M to 0 M. The peak fractions were combined and dialyzed into buffer containing 30 mM HEPES, pH 7.5, 30 mM KCl, 0.5% inositol, and 1 mM PMSF. Samples were then loaded onto a 1-ml Mono-S column.
GE Healthcare, Piscataway, NJ, USA), and eluted using a linear gradient of KCl (30 mM to 2 M). Purified FEN1 was aliquoted and frozen at -80 °C until further use.

APE1 was expressed in *E. Coli* BL21(DE3). Two liters of LB medium were inoculated with one colony each of the transformed BL21(DE3) cells and were incubated overnight without shaking. The cells were then incubated at 37°C at 225 rpm until OD$_{600}$ reached 0.6. The APE1 expression was induced by 0.5 mM IPTG for 3.5 hours. Cells were harvested by centrifugation at 2500 rpm for 30 minutes at 4 °C. The supernatant was discarded and cell pellets were lysed in lysis buffer, which contained 50 mM HEPES, pH 7.5, 30 mM NaCl, 1 mM DTT, 1 mM EDTA, and 1 mM PMSF. The cell lysates were centrifuged at 12,000 rpm for 30 minutes at 4 °C. The supernatant was loaded onto a 10-ml sepharose Q column operated by an AKTA FPLC. The flow-through was collected and dialyzed into buffer containing 50 mM HEPES, pH 7.5, 30 mM NaCl, and 1 mM PMSF, and then loaded onto a 5-ml CM sepharose column (Bio-Rad, Hercules, CA, USA), with fractions eluted using a linear gradient of NaCl (30 mM to 2 M). The peak fractions were combined and dialyzed into buffer containing 50 mM HEPES, pH 7.5, 30 mM NaCl, and 1 mM PMSF. Samples were then loaded onto a 1-ml Mono-S column, and eluted using a linear gradient of NaCl from 30 mM to 2 M. Peak fractions were combined and dialyzed into buffer containing 50 mM HEPES, pH 7.5, 1.7 M (NH$_4$)$_2$SO$_4$, and 1 mM PMSF. Samples were then loaded onto a 2-ml phenyl sepharose column, with fractions eluted using a linear gradient of (NH$_4$)$_2$SO$_4$ ranging from 1.7 M to 0 M. Purified APE1 was aliquoted and frozen at -80 °C until further use.
**Reconstituted BER assay**

*In vitro* BER of an 8-oxoG or abasic site analog, THF, in the loop region of a (CAG)$_7$ or (CAG)$_{14}$ hairpin was carried out by incubating 50 nM substrate with the indicated concentrations of OGG1, APE1, pol β, FEN1, and LIG I. Substrates containing an 8-oxoG were initially incubated with OGG1 (100 nM) and APE1 (50 nM) at 37 °C for 30 minutes. Substrates were then subject to a phenol:chloroform extraction for removing the OGG1 and APE1. The cleaved substrates were then precipitated with ethanol and resuspended for the BER assay. Substrates containing a THF in the hairpin loop were pre-cleaved with APE1 alone and subject to the BER assay. BER of an abasic site, the THF residue, within the loop of the (CAG)$_7$ and (CAG)$_{14}$ hairpins was measured by incubating 25 nM substrates with the indicated concentrations of APE1, pol β, FEN1, and LIG I. All reactions were carried out 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), 0.2 mM DTT, and 0.01% Nonidet P-40, with 5 mM MgCl$_2$, 2 mM ATP, and 50 μM dNTPs. The reaction mixtures were incubated at 37 °C for 15 minutes, and terminated by addition of stopping buffer containing 95% formamide and 10 mM EDTA. Reaction mixtures were denatured at 95 °C for 10 minutes and separated by 15% or 18% urea-denaturing polyacrylamide gel electrophoresis. Substrates and products were detected and analyzed using a Pharos FX Plus Phosphorimager from Bio-Rad (Hercules, CA, USA).

**BER enzymatic activity assay**

The activities of APE1 AP endonuclease and APE1 3’-5’ exonuclease, as well as FEN1 flap cleavage activity on the (CAG)$_7$ and (CAG)$_{14}$ hairpin substrates and double-flap
substrates were measured by incubating 50 nM or 100 nM substrates with the indicated concentrations of FEN1, APE1, and pol β in reaction buffer containing 50 mM Tris-HCl, pH 7.5, 50 mM KCl, 0.1 mM EDTA, 0.1 mg/ml BSA, 0.2 mM DTT, and 0.01% Nonidet P-40, with 5 mM MgCl₂ and 50 μM dNTPs. The reaction mixtures were incubated at 37 °C for 15 minutes, and terminated by addition of stopping buffer containing 95% formamide and 10 mM EDTA. Reaction mixtures were denatured at 95 °C for 10 minutes and separated by 15% urea-denaturing polyacrylamide gel electrophoresis. Substrates and products were detected using a Pharos FX Plus Phosphorimager (Bio-Rad Laboratory, Hercules, CA, USA).

The effects of APE1 on LIG I activity

The effects of APE1 on LIG I activity were determined via in vitro reconstitution of BER on the double-flap substrates resulting from APE1 5’-incision of an abasic site within the loop of the (CAG)₇ and (CAG)₁₄ hairpins at low concentrations of APE1 (0.5 and 1 nM) that exhibit little 3’-5’ exonuclease activity, and LIG I (1 nM). 25 nM substrates were incubated with the indicated concentrations of FEN1, APE1, pol β, and LIG I in reaction buffer containing 50 mM Tris-HCl, pH 7.5, 50 mM KCl, 0.1 mM EDTA, 0.1 mg/ml BSA, 0.2 mM DTT, and 0.01% Nonidet P-40, with 5 mM MgCl₂, 2 mM ATP, and 50 μM dNTPs. The reaction mixtures were incubated at 37 °C for 15 minutes, and terminated by stopping buffer containing 95% formamide and 10 mM EDTA. Reaction mixtures were denatured at 95 °C for 10 minutes and separated by 15% or 18% urea-denaturing polyacrylamide gel electrophoresis. Results were analyzed using a Pharos FX Plus Phosphorimager (Bio-Rad Laboratory, Hercules, CA, USA).
RESULTS

APE1 stimulates CAG repeat hairpin removal during BER of a base lesion in a hairpin loop

To determine whether APE1 can facilitate the removal of a CAG repeat hairpin during BER within the loop of the hairpin, we reconstituted BER using synthesized oligonucleotide substrates containing either a (CAG)$_7$ or (CAG)$_{14}$ hairpin with an 8-oxoG (Figure 1.1) or abasic site analog, tetrahydrofuran (THF) (Figure 1.2), in the loop region of the hairpin. We found that during BER of an 8-oxoG in the loop of the small (CAG)$_7$ hairpin, the production of the “repaired unexpanded product”, the product with the same length as the template strand, was significantly increased in the presence of 50 nM and 100 nM APE1 (Figure 1.1, lanes 7-8 and lanes 10-11) compared to the amount of product resulting from BER in the absence of APE1 (Figure 1.1, lanes 6 and 9). During BER of an 8-oxoG in the loop of the large (CAG)$_{14}$ hairpin, 50 nM and 100 nM APE1 resulted in the production of the repaired unexpanded product (Figure 1.1, lanes 18-19 and lanes 21-22), whereas little unexpanded product was generated in the absence of APE1 (Figure 1.1, lanes 17 and 20). APE1 alone failed to produce any repair products (Figure 1.1, lanes 4-5 and lanes 15-16). Consistent with the findings from BER of the 8-oxoG in the hairpin loop, 50 nM and 100 nM APE1 also resulted in the production or an increase in the amount of repaired unexpanded product during BER of a THF in the loop region of the (CAG)$_7$ and (CAG)$_{14}$ hairpins (Figure 1.2, compare lanes 7-8 with lane 6 and lanes 10-11 with lane 9, lanes 18-19 with lane 17, and lanes 21-22 with lane 20). The results indicated that APE1 promoted the removal of the (CAG)$_7$ and (CAG)$_{14}$ repeat hairpins during BER of both an
Figure 1.1 APE1 promotes the removal of a trinucleotide repeat hairpin during BER of 8-oxoG in a TNR hairpin loop (222)
The effect of APE1 on hairpin removal during BER of an 8-oxoG in the loop of a small (CAG)$_7$ hairpin (left panel) and a large (CAG)$_{14}$ hairpin (right panel) was examined by reconstituting BER with the hairpin containing substrates. Lanes 1 and 12 are markers that indicate the length of the template strand or unexpanded product. Lanes 2 and 13 are markers that indicate the length of the damage-containing strand or expanded product.
Figure 1.2 APE1 promotes the removal of a trinucleotide repeat hairpin during BER of an abasic site in a CAG repeat hairpin loop (222)

The effect of APE1 on CAG repeat hairpin removal during BER of a THF residue in the loop of a small (CAG)\textsubscript{7} hairpin (left panel) and a large (CAG)\textsubscript{14} hairpin (right panel) was examined by reconstituting BER with the substrates containing a THF in the (CAG)\textsubscript{7} and (CAG)\textsubscript{14} loop. Lanes 1 and 12 indicate markers that illustrate the length of the template strand or unexpanded product, corresponding to the removal of the entire hairpin. Lanes 2 and 13 correspond to markers that represent the length of the substrate only.
8-oxoG and abasic lesion in the loop region of the hairpin, thereby preventing TNR expansions.

**APE1 stimulates the processing of a double-flap intermediate during BER**

Our previous studies have shown that during BER of a base lesion within the loop of a CAG repeat hairpin, the hairpin is converted to a double-flap intermediate as a result of APE1 5’-incision of an abasic site (155). We found that the double-flap intermediate was then processed by a 3'-5' flap endonuclease, Mus81/Eme1, and FEN1 5'-flap cleavage (155). This led to cleavage of the flaps, thereby resulting in the removal of the hairpin. We further hypothesized that APE1 is also involved in the processing of a double-flap intermediate by employing its 3'-5' exonuclease activity to process the upstream 3'-flap, thereby promoting removal of a TNR hairpin structure. To test this, we reconstituted BER with a substrate containing an upstream 3’-(CAG)₄ flap and a downstream 5’-(CAG)₃ flap with a THF residue, as well as with a substrate containing a (CAG)₇ flap on both the upstream and downstream strands with a 5’-THF residue, in the absence and presence of APE1. These substrates simulate the double-flap intermediates produced by APE1 5’-incision of the (CAG)₇ and (CAG)₁₄ hairpins, respectively. We found that BER reconstituted with FEN1 (0.5 nM and 1 nM) and LIG I (5 nM) without and with pol β (5 nM) on the (CAG)₃/(CAG)₄ double-flap substrate, resulted in the production of a significant amount of the repaired unexpanded product (Figure 1.3A, lanes 6, 9, and 12). However, 50 nM and 100 nM APE1 did not significantly alter the production of the unexpanded repaired product with the substrate (Figure 1.3A, compare lanes 7-8 with lane 6, lanes 10-11 with lane 9, and lanes 13-14 with lane 12). Further characterization of the
3’-5’ exonuclease activity of APE1 showed that 50 nM and 100 nM APE1 exhibited efficient 3’-5’ exonuclease activity that generated a significant amount of exonucleolytic cleavage products (Figure 1.3A, lanes 4-5). This indicated that APE1 cleaved the upstream CAG repeats exonucleolytically. The results showed that APE1 exonuclease did not play a significant role in the removal of the short double-flaps, suggesting that FEN1 flap cleavage plays a predominant role in processing the short double-flaps. Interestingly, we found that BER with the substrate containing long double-flaps in the presence of FEN1 (1 nM) and LIG I (5 nM) only resulted in a series of repaired products that are longer than the repaired unexpanded product (Figure 1.3A, lanes 20 and 23), but shorter than the original hairpin-containing substrate. They were termed “repaired expanded product” (Figure 1.3A, lanes 20-25). This occurred in the absence and presence of pol β (10 nM) (Figure 1.3A, lanes 20-25) indicating that the production of the expanded products was independent of pol β. However, we found that the presence of 50 nM and 100 nM APE1 resulted in the formation of repaired unexpanded product (Figure 1.3A, lanes 21-22 and lanes 24-25). APE1 alone failed to generate the product (Figure 1.3A, lanes 18-19), indicating that the product was specifically generated through BER. However, APE1 failed to affect the production of repaired expansion products during BER (Figure 1.3A, compare the amount of repaired expanded products in lanes 21-22 and lanes 24-25 with that in lanes 20 and 23). This suggests that APE1 promoted the complete resolution of the long double-flap intermediate resulting from a (CAG)_{14} hairpin, thereby specifically stimulating the production of repaired unexpanded product. To further determine the effect that is specifically from APE1 on the production of repaired unexpanded products during BER, we reconstituted BER with 50 nM and 100 nM APE1 without or with a low concentration
of FEN1 (1 nM) with the double-flap substrates without a THF residue. We found that on the short \((\text{CAG})_3/(\text{CAG})_4\) double-flap substrate, APE1 (50 nM and 100 nM) along with LIG I (5 nM) was sufficient to generate repaired unexpanded product in the absence and presence of 10 nM pol \(\beta\) (Figure 1.3B, lanes 7-8 and lanes 10-11). This indicated that APE1 facilitated the processing of the short \((\text{CAG})_3/(\text{CAG})_4\) double-flap by cleaving the upstream strand via its 3'-5' exonuclease activity independent of FEN1. The production of the repaired unexpanded product was significantly stimulated by the presence of 1 nM FEN1 (Figure 1.3B, left panel, lanes 13-14 and lanes 16-17). To further determine the specific effect of APE1 on the removal of the long \((\text{CAG})_7/(\text{CAG})_7\) double-flaps, we reconstituted BER with the substrate containing \((\text{CAG})_7/(\text{CAG})_7\) double-flaps without the 5'-THF residue in the absence and presence of FEN1 (1 nM) (Figure 1.3B, right panel). We found that 50 nM and 100 nM APE1 alone or APE1 along with LIG I (5 nM) in the absence and presence of pol \(\beta\) (10 nM) failed to produce any repair products (Figure 1.3B, lanes 21-22, lanes 24-25 and lanes 27-28). BER reconstituted with FEN1 and LIG I produced only repaired expanded products without and with pol \(\beta\) (10 nM) (Figure 1.3B, lanes 29 and 32). However, BER reconstituted with APE1, FEN1 and LIG I in the absence and presence of pol \(\beta\) resulted in the production of a significant amount of the repaired unexpanded product (Figure 1.3B, lanes 30-31 and lanes 33-34). The results indicated that APE1 significantly promoted the removal of both the short \((\text{CAG})_3/(\text{CAG})_4\) and long \((\text{CAG})_7/(\text{CAG})_7\) double-flap intermediates. Thus, we conclude that APE1 can facilitate the resolution of the double-flap intermediate formed during BER in a CAG repeat hairpin loop and stimulate the formation of the unexpanded repaired product. APE1 alone can lead to complete removal of a short double-flap intermediate through its 3'-5' exonuclease
Figure 1.3 APE1 stimulates removal of a double-flap intermediate during BER (222)

APE1 stimulation of resolution of the double-flap intermediate resulting from (CAG)₇ and (CAG)₁₄ hairpins during BER was examined by reconstituting BER with the substrates containing short (CAG)₃/(CAG)₄ double-flaps (left panel) or long (CAG)₇/(CAG)₇ double-flaps (right panel) with a THF residue (A) and without a THF residue (B) in the absence and presence of APE1. (A) Lanes 1 and 15 indicate the size markers of the unexpanded product. Lanes 2 and 16 represent the size markers of the expanded product. (B) Lanes 1 and 18 indicate the size markers that represent the template strand or unexpanded product. Lanes 2 and 19 correspond to the size markers that illustrate the damaged strand.
APE1 stimulation of resolution of the double-flap intermediate resulting from (CAG)₇ and (CAG)₁₄ hairpins during BER was examined by reconstituting BER with the substrates containing short (CAG)₃/(CAG)₄ double-flaps (left panel) or long (CAG)₇/(CAG)₇ double-flaps (right panel) with a THF residue (A) and without a THF residue (B) in the absence and presence of APE1. (A) Lanes 1 and 15 indicate the size markers of the unexpanded product. Lanes 2 and 16 represent the size markers of the expanded product. (B) Lanes 1 and 18 indicate the size markers that represent the template strand or unexpanded product. Lanes 2 and 19 correspond to the size markers that illustrate the damaged strand.

Figure 1.3 APE1 stimulates removal of a double-flap intermediate during BER (222)
activity (Figure 1.3B, lanes 7-8). However, it can only promote the processing of an intermediate with long double-flaps by cooperating with FEN1 flap cleavage (Figure 1.3B, lanes 30-31 and lanes 33-34).

**APE1 3’-5’ exonuclease activity shortens the upstream 3’-region of a double-flap intermediate**

Because our previous studies have shown that a 3’-5’ flap endonuclease, Mus81/Eme1, can facilitate the removal of a TNR hairpin by endonucleolytically cleaving a 3’-flap from the double-flap intermediate during BER (155), we further hypothesized that APE1 may promote the removal of a CAG repeat hairpin by shortening the upstream 3’-flap of a double-flap intermediate through its 3’-5’ exonuclease activity. To test this, we examined the 3’-5’ exonuclease activity of APE1 on both the (CAG)$_3$/(CAG)$_4$ and (CAG)$_7$/(CAG)$_7$ double-flap substrates without or with the THF residue, that were labeled at the 5’-end of the upstream flaps. This allowed the detection of shortened upstream flaps (Figure 1.4). We found that 10 nM-100 nM APE1 exonucleolytically cleaved the upstream 3’-flap of the (CAG)$_3$/(CAG)$_4$ double-flap substrate, leading to the production of cleavage products that are shorter than the substrate (Figure 1.4, lanes 2-5 and lanes 7-10). APE1 exonucleolytic cleavage was slightly greater on the substrate lacking the THF residue (Figure 1.4, compare lanes 7-10 with lanes 2-5). The results indicated that the APE1 3’-5’ exonuclease activity efficiently processed the short 3’-upstream strand of the double-flap intermediate resulting from a small CAG repeat hairpin. To further determine if APE1 can process a long 3’-CAG repeat flap, we examined the APE1 activity on the (CAG)$_7$/(CAG)$_7$ double-flap substrate, without or with the THF residue. The results showed that 10 nM-
100 nM APE1 resulted in the production of a small amount of exonuclease cleavage product from the (CAG)$_7$(CAG)$_7$ double-flap substrate with a THF (Figure 1.4 lanes 12-15). For the double-flap substrate without a THF, a small amount of APE1 exonucleolytic cleavage products was also generated (Figure 1.4, lanes 17-20). This indicated that APE1 3’-5’ exonuclease cleaved a long upstream 3’-flap with a low efficiency. This further suggests that the resolution of the long double-flap intermediate formed from the larger (CAG)$_{14}$ hairpin during BER requires a coordinated activity between APE1 and other nucleases such as FEN1 that can resolve the double-flap intermediate through its endonucleolytic cleavage of a 5’-flap.
**Figure 1.4 APE1 3’-5’ exonuclease activity shortens the 3’-flap of the double-flap intermediates (222)**

APE1 3’-5’ exonuclease activity of shortening the 3’-flap of the double-flap intermediates was examined by incubating the substrates with the short (CAG)$_3$/(CAG)$_4$ and long (CAG)$_7$/(CAG)$_7$ double-flaps without or with a THF residue with increasing concentrations of APE1 ranging from 10 nM to 100 nM. Lanes 1, 6, 11, and 16 correspond to the substrate only. Lanes 2-5, lanes 7-10, lanes 12-15, and lanes 17-20 correspond to increasing concentrations of APE1 ranging from 10 nM to 100 nM. Substrates were $^{32}$P-labeled at the 5’-end of the upstream strands and are illustrated above the gels.
APE1 fails to directly stimulate FEN1 cleavage on a double-flap intermediate

APE1 has been shown to stimulate FEN1 flap cleavage during BER of an abasic lesion located in a random DNA sequence (209). To further determine if APE1 can promote the removal of a CAG repeat hairpin by facilitating FEN1 flap cleavage of the 5’-flap of a double-flap intermediate, we tested whether APE1 could stimulate FEN1 flap cleavage on the (CAG)$_3$/ (CAG)$_4$ and (CAG)$_7$/ (CAG)$_7$ double-flap intermediates (Figure 1.5). We found that in the absence and presence of 50 nM and 100 nM APE1, similar amounts of FEN1 cleavage products were generated by 0.5 nM and 1 nM FEN1 during BER of the (CAG)$_3$/ (CAG)$_4$ and (CAG)$_7$/ (CAG)$_7$ double-flap substrate (Figure 1.5, compare lanes 4-5 with lane 3, lanes 10-11 with lane 9, lanes 18-19 with lane 17, and lanes 24-25 with lane 23). We further examined if APE1 could stimulate FEN1 cleavage in the presence of pol β (10 nM). We did not observe any increase of FEN1 cleavage products in the presence of 50 nM and 100 nM APE1 with 10 nM pol β (Figure 1.5, compare lanes 7-8 with lane 6, lanes 13-14 with lane 12, lanes 21-22 with lane 20, and lanes 27-28 with lane 26). Thus, the results indicated that APE1 failed to stimulate FEN1 cleavage activity in processing the short and long double-flap intermediates during BER. We therefore concluded that the stimulation of the removal of a CAG repeat hairpin by APE1 was not the result of APE1 stimulation of FEN1 cleavage activity.
Figure 1.5 APE1 fails to stimulate FEN1 cleavage of the double-flap intermediate
(222)
The stimulatory effects of APE1 on FEN1 cleavage activity on the short (CAG)$_3$/((CAG)$_4$ and long (CAG)$_7$/((CAG)$_1$) double-flap intermediates during BER were examined by incubating 0.5 nM or 1 nM FEN1 with the double-flap substrates in the absence and presence of APE1. Lanes 1 and 15 are size markers. Lanes 2 and 16 correspond to the substrate only. Lanes 3-5 and lanes 17-19 correspond to reaction mixture with FEN1 (0.5 nM) in the absence and presence of APE1. Lanes 6-8 and lanes 20-22 correspond to reaction mixture with FEN1 (0.5 nM) and pol β (10 nM) in the absence and presence of APE1. Lanes 9-11 and lanes 23-25 correspond to reactions with FEN1 (1 nM) in the absence and presence of APE1. Lanes 12-14 and lanes 26-28 correspond to reactions with FEN1 (1 nM) and pol β (10 nM) in the absence and presence of APE1. Substrates were $^{32}$P-labeled at the 3’-end of the downstream strands and are illustrated above each gel.
**APE1 promotes the production of unexpanded repair product by stimulating LIG I**

A study from the Bambara group showed that APE1 can stimulate LIG I activity to promote ligation of a nick and facilitate the formation of repaired products during BER (209). Thus, it is possible that APE1 may also stimulate LIG I activity to promote the removal of a CAG repeat hairpin and facilitate the production of the repaired products during BER of a base lesion in a hairpin loop. To test this possibility, we determined the effects of APE1 on LIG I activity independent of its 3’-5’ exonuclease activity by measuring the production of the repaired products in the presence of low concentrations of APE1 (0.5 nM and 1 nM) that exhibited a low 3’-5’ exonuclease activity on the (CAG)$_3$/(CAG)$_4$ double-flap substrate (Figure 1.6, lanes 3-4), and little exonucleolytic activity on the (CAG)$_7$/(CAG)$_7$ double-flap substrate (Figure 1.6, lanes 9-10). We found that low concentrations of APE1 failed to alter the production of repaired products resulting from a short double-flap substrate (Figure 1.7, compare lanes 5-6 with lane 4 and lanes 8-9 with lane 7) indicating that APE1 did not significantly alter LIG I activity during removal of a small hairpin. However, the same concentrations of APE1 resulted in the production of the unexpanded product on the long (CAG)$_7$/(CAG)$_7$ double-flap substrate (Figure 1.7, compare lanes 14-15 with lane 13 and lanes 17-18 with lane 16). The results indicated that APE1 stimulated the ligation activity of LIG I, thereby promoting the production of the unexpanded repair product. Interestingly, we observed that the same concentrations of APE1 failed to stimulate the production of shortened repaired expanded products (Figure 1.7, lanes 14-15, lanes 17-18) suggesting that APE1 failed to stimulate the ligation of a nick by LIG I that was adjacent a hairpin structure.
The 3'-5' exonuclease activity of APE1 at low concentrations (222)
The 3'-5' exonuclease activity of APE1 on the short (CAG)₃/(CAG)₄ double-flap and long (CAG)₇/(CAG)₇ double-flap substrates, was determined by incubating the double-flap substrates with APE1 at concentrations of 0.25 nM to 10 nM. Lanes 1 and 7 correspond to the substrate only. Lanes 2-6 and lanes 8-12 correspond to reaction mixture with the substrates and increasing concentrations of APE1. Substrates were ³²P-labeled at the 5'-end of the upstream strands and are illustrated above each gel.
Figure 1.7 APE1 stimulates LIG I activity on a double-flap intermediate (222)

The effects of APE1 to stimulate the ligation activity of LIG I on the short (CAG)$_3$/(CAG)$_4$ and long (CAG)$_7$/(CAG)$_7$ double-flap intermediates during BER were determined by reconstituting BER in the absence and presence of low concentrations of APE1 at which a little or no 3'–5’ exonuclease activity was observed on the short and long double-flap substrates, respectively. Lanes 1 and 10 indicate the size markers of the template strand, corresponding to complete removal of the double-flaps and full repair. Lanes 2 and 11 correspond to the markers that illustrate the length of the damaged strand containing a CAG repeat hairpin and the size of repaired expanded product.
DISCUSSION

In this study, we made the first discovery that APE1 promoted the removal of a TNR hairpin (Figures 1.1 and 1.2) by shortening the upstream 3’-flap via its 3’-5’ exonuclease activity (the left panels of Figure 1.3 and Figure 1.4). We found that APE1 specifically facilitated the production of the repaired unexpanded product by stimulating the ligation activity of LIG I (the right panel of Figure 1.7). We further demonstrated that APE1 also stimulated the removal of a large (CAG)14 repeat hairpin by exonucleolytically cleaving the 3’-flap of a long (CAG)7/(CAG)7 repeat double-flap intermediate (Figure 1.4) as well as stimulating LIG I activity (Figure 6). This specifically facilitated the formation of the repaired unexpanded product during BER (the right panels of Figure 1.3A and 1.3B and Figure 1.7). In addition, we showed that the stimulatory effects of APE1 on TNR hairpin removal were not the result of APE1 stimulation of FEN1, because APE1 did not show any stimulatory effects on FEN1 cleavage activity or pol β DNA synthesis (Figure 1.5 and Figure 1.8). Our results support a model by which APE1 stimulates TNR hairpin removal to prevent TNR expansions (Figure 1.9), showing that upon exposure of TNR tracts to oxidative DNA damaging agents and the formation of an oxidized base lesion in the loop region of a TNR hairpin, or relocation of the lesion to the loop, a DNA glycosylase removes the damaged base leaving an abasic site. APE1 incises the 5’-side of the abasic site in the hairpin loop, converting the hairpin into a double-flap intermediate. A short double-flap intermediate converted from a small hairpin can be directly processed by APE1 via its exonucleolytic cleavage of the upstream 3’-region, shortening the flap. In coordination with FEN1 cleavage of a 5’- sugar phosphate residue, this leads to the removal of a small hairpin and prevention of repeat expansion (Figure 1.9, subpathway 1). A long
double-flap intermediate resulting from a large hairpin is resolved by the coordination among APE1, FEN1, pol β, and LIG I that results in the formation of different intermediates leading to either repeat expansion or no expansion. For the long flaps that form two adjacent small hairpins with a short 5’-flap containing a sugar phosphate, FEN1 cleaves the flap with its alternate flap cleavage activity leaving a gapped DNA. Pol β then fills in the gap leaving a nick for ligation by LIG I. The nicked intermediate with small hairpins traps a large amount of LIG I, resulting in ligation of hairpins by LIG I and the production of repaired expanded product. For this intermediate, APE1 cannot access the nick to interact with LIG I and stimulate the activity of the enzyme (Figure 1.9, the left side of the subpathway 2). For the long double-flap intermediate that forms a downstream long flap, FEN1 removes the flap leaving a short upstream 3’-flap. This allows the upstream flap to reanneal into an intermediate with an upstream small repeat bubble with a 3’-annealed region that is subsequently cleaved by APE1 3’-5’ exonuclease activity leaving a gapped DNA. Pol β then fills in the gap resulting in a nick, which is sealed by the limited amount of LIG I that is not bound to the nick located in between the two small hairpins. Subsequently, APE1 facilitates the binding of the limited amount of LIG I to the nick. This stimulates the ligation activity of LIG I, thereby promoting the generation of the repaired unexpanded product (Figure 1.9, right, subpathway 2).
Figure 1.8 APE1 fails to stimulate pol β DNA synthesis on the double-flap intermediate (222)
APE1 stimulatory effects on the DNA synthesis activity of pol β on the short (CAG)$_3$/(CAG)$_4$ and long (CAG)$_7$/(CAG)$_7$ double-flap intermediates during BER were examined in the absence and presence of 0.5 nM or 1 nM APE1. Lanes 1 and 5 correspond to the substrate only. Lanes 2 and 6 indicate reaction mixture with the substrates and 1 nM pol β. Lanes 3-4 and 7-8 correspond to
Figure 1.9 APE1 prevents trinucleotide repeat expansions during BER in a hairpin loop (222)

During BER in the loop of a TNR hairpin, APE1 incision converts the hairpin into a double-flap intermediate. A short double-flap intermediate resulting from a small hairpin undergoes flap equilibration to allow annealing of the upstream flap. The annealed upstream 3’-region can be processed by APE1 3’-5’ exonuclease cleavage, shortening the upstream strand to allow annealing of the downstream 5’-flap, leading to removal of a small hairpin (subpathway 1). A long double-flap intermediate formed by APE1 5’-incision in a large hairpin can result in different consequences upon the formation of hairpin structures by the flaps. If the double-flaps form two small hairpins adjacent to each other, FEN1 removes a sugar phosphate with a short flap through alternate flap cleavage activity, leaving a gapped DNA. Pol β then fills in the gap, leaving a nick for ligation by LIG I. This intermediate traps a large quantity of LIG I, resulting in ligation of hairpins by LIG I and repeat expansion. For this intermediate, APE1 fails to access the nick to interact with LIG I to stimulate the activity (left, subpathway 2). For the long double-flap intermediates that form a downstream long flap, FEN1 removes the flap leaving a short upstream 5’-flap. This allows the upstream flap to reanneal into an intermediate with an upstream small repeat bubble with a 3’-annealed region that is subsequently cleaved by APE1 3’-5’ exonuclease activity leaving gapped DNA. Pol β then fills in the gap resulting in a nick that is sealed by the limited amount of LIG I. Subsequently, APE1 facilitates the limited amount of LIG I to bind to the nick, stimulating LIG I activity, facilitating the removal of the double-flaps and generation of the unexpanded product (right, subpathway 2).
Here, for the first time, we identified a new role of APE1 in the resolution of the upstream 3’-flap formed during removal of a TNR hairpin through BER. Our findings also provide new insight into how the BER enzymes APE1, pol β, FEN1 and LIG I can cooperate to remove TNR hairpins and maintain TNR stability. Although the APE1 exonuclease activity has been shown to be able to remove 3’-mismatched bases and 3’-blocking groups (82, 216, 218-220), the biological function of the 3’-5’ exonuclease of APE1 remains to be elucidated. Our study demonstrates that during BER in a hairpin loop, APE1 promoted the formation of repaired unexpanded product specifically by exonucleolytically cleaving the upstream strand of a double-flap intermediate and stimulating the LIG I activity in sealing a nick. We further verified the exonucleolytic cleavage of APE1 on both the short and long double-flap substrates specifically to examine if the 3’-5’ exonuclease activity of APE1 is indeed involved in the resolution of both of the short and long double-flaps. Our results showed that APE1 exhibited efficient 3’-5’ exonuclease activity on the short double-flap substrate by predominantly cleaving two repeat units from the short upstream strand (Figure 1.4, lanes 2-5 and lanes 7-10), whereas it mainly cleaved one repeat unit from the long upstream flap (Figure 1.4, lanes 13-15 and lanes 19-20). The results indicate that APE1 3’-5’ exonucleolytic cleavage of the upstream flaps plays an important role in preventing and attenuating TNR expansion during the repair of both small and large hairpin structures.

Our previous study has shown that the 3’-5’ flap endonuclease Mus81/Eme1 protein complex prevents TNR expansions by cleaving the upstream 5’-flap (155). This indicates that any nuclease that can cleave the upstream 5’-flap may lead to the prevention of TNR repeat expansions during BER in a hairpin loop. Although the 3’-5’ exonuclease
activity of APE1 has been well established and proposed to serve as a proof-reading enzyme for pol β during BER, its biological function remains to be elucidated. In this study, for the first time, we reveal a novel function of APE1 3’-5’ exonuclease activity and APE1 stimulation on the activity of LIG I in preventing TNR expansions via its unique coordination with other BER enzymes. Although the APE1 exonuclease activity is weaker than its AP endonuclease activity and accomplished by a relatively high concentration of the enzyme, the abundance of the enzyme in mammalian cells (10^5-10^6 molecules/cell) (203) appears to allow it to act as an efficient 3’-5’ exonuclease for sustaining TNR stability in cells.

Our results also suggested that for a double-flap intermediate, the upstream flap annealed back to the template strand creating a downstream 5’-flap. Subsequently, APE1 3’-5’ exonuclease cleaved the upstream strand creating an upstream gap into which the downstream 5’-flap reannealed to generate a nick. This is supported by the observation that APE1 exonuclease failed to make cleavage on a substrate containing an upstream 3’-flap that fails to anneal with the template strand, indicating that APE1 3’-5’ exonuclease cannot endonucleolytically cleave a single-stranded 3’-flap (218). This further indicates that the APE1 3’-5’ exonuclease activity was accomplished through the reannealing of the upstream 3’-flap to the template strand that creates a nick or gap in a DNA duplex through flap equilibration (221).

Our results showed that APE1 specifically stimulated the production of the repaired unexpanded product, but not the production of repaired expanded products via its 3’-5’ exonuclease activity (Figure 1.4) as well as its stimulation of the activity of LIG I (Figure 1.7, lanes 14-15 and lanes 17-18). This was indicated by the production of the repaired
unexpanded product during BER with the long double-flap substrate in the presence of low concentrations of APE1 (Figure 1.7, compare lanes 14-15 with lane 13 and lanes 17-18 with lane 16) that exhibit little 3'-5' exonuclease activity (Figure 1.6). This indicates that APE1 can specifically promote the formation of repaired unexpanded product from BER of a long double-flap substrate by stimulating LIG I activity independent of its 3'-5' exonuclease activity. The production of a large quantity of repaired expanded products suggests that the products were generated by ligation of a nick that was flanked by a small upstream and downstream CAG repeat hairpin structure that provided a limited space for LIG I alone to bind to the nick. This further prevented the binding of both LIG I and APE1 to the nick, thereby preventing any effects from APE1 on LIG I activity. Moreover, it is possible that the nick flanked by two CAG repeat hairpins may also trap LIG I, thereby resulting in depletion of LIG I that reduced the availability of LIG I for generating the unexpanded product. In this scenario, APE1 became critical for recruiting the limited amount of LIG I to bind to the nick at a duplex DNA resulting from the complete removal of the double flaps. This led to stimulation of the formation of the unexpanded product. This notion was supported by our results showing that APE1 failed to stimulate LIG I activity during BER with the short double-flap substrate that led to the production of a nick in a duplex DNA. This further suggests that there was a sufficient amount of LIG I to carry out efficient ligation on a nick in duplex DNA during BER with a short double-flap intermediate because the short double-flaps did not form hairpin structures. This was also supported by the fact that LIG I activity at a nick on duplex DNA can be stimulated by APE1 at a very low concentration of LIG I (0.06-0.1 nM) (209). Our results indicate that APE1 promoted the removal of a small TNR hairpin by employing its 3'-5' exonuclease
activity. It facilitated the removal of a large TNR hairpin by both its 3'-5' exonuclease activity and its stimulation of LIG I activity.

FEN1 was previously shown by the Bambara group to be stimulated by APE1 through a direct interaction between the two enzymes (209). However, we failed to observe the stimulatory effect on the double-flap substrates. It is possible that with the double flap intermediate, the upstream flap may interrupt the direct interaction between APE1 and FEN1, thereby eliminating APE1 stimulatory effect on FEN1 flap cleavage. It is conceivable that APE1 can be bound to the upstream region of the double-flap intermediate, which is too far away from the base of the downstream 5' flap of the intermediate. This would prevent the direct interaction between the enzymes and the stimulatory effect.

It should be noted that we found that the formation of unexpanded repaired product from both the short and long double-flap substrates was promoted in the presence of pol β (Figure 1.1 lanes 10-11 and lanes 21-22, Figure 1.2 lanes 10-11 and lanes 21-22, Figure 1.3a lanes 24-25, Figure 1.3b lanes 33-34, and Figure 1.7 lanes 17-18). However, the formation of the unexpanded product from both the short and long double-flap substrates was also observed in the absence of pol β (Figure 1.1 lanes 6-8 and lanes 18-19, Figure 1.2 lanes 7-8 and lanes 18-19, Figure 1.3a lanes 6-11 and lanes 21-22, Figure 1.3b lanes 7-8, lanes 10-11, and lanes 30-31, and Figure 1.7 lanes 4-6 and lanes 14-15). This indicates that pol β is not required for the production of the unexpanded product during BER. This is because for both a short and long double-flap intermediate, FEN1 cleavage of the downstream flap along with APE1 3'-5' exonucleolytic cleavage of the upstream strand can result in the production of nicked DNA for ligation as well as gapped intermediates that
need to be filled in by pol β gap-filling synthesis for generating nicked DNA for ligation and completion of repair. In addition, we found that the production of the unexpanded product was slightly stimulated by the presence of pol β with both the short and long double-flap substrates, indicating that there is no preference for pol β in the repair of a short or long double-flap intermediate during BER.

Our results showed that repair of the long double-flap substrates led to the formation of expanded products. This is because the long flaps can reanneal with the template to form an intermediate with a small hairpin in the upstream or/and the downstream strand attached with a short flap as described in our previous study (155). This further indicates that a hairpin intermediate formed in the downstream flap can be further processed by FEN1 alternate flap cleavage resulting in ligation of a small hairpin and production of shortened expanded products, thereby attenuating repeat expansion (155). However, our results also demonstrated that a significant amount of unexpanded product was generated during repair of the long double-flap intermediate, suggesting that the upstream and downstream flaps were removed completely by the cooperative cleavage activities of APE1 3′-5′ exonuclease activity and FEN1 flap cleavage through their direct cleavage of the flaps as well as their cleavage of a series of short flaps formed by a long flap in a stepwise manner.

Previous studies from the Delaney group have shown that a hairpin loop is more susceptible to the formation of oxidative DNA damage such as 8-oxoG than duplex DNA (150). Moreover, an 8-oxoG formed in the stem region of a hairpin can relocate to the loop region (154), indicating that the oxidized base lesion preferentially accumulates in a hairpin loop. Although this can further reduce OGG1 binding ability to the oxidized base lesion
we have demonstrated that OGG1 can efficiently remove an 8-oxoG located in the loop of a TNR hairpin within a TNR duplex tract, leaving an abasic site in the loop region. Subsequently, APE1 incises the 5' side of the abasic site cleaving the loop and converting the hairpin into a double-flap intermediate. This results in removal of the hairpin through BER, demonstrating that BER in a TNR hairpin loop can lead to prevention and attenuation of TNR expansion. Employing synthetic zinc finger endonucleases, which specifically target hairpin structures of a specific sequence (CAG/CTG), the Leffak group found a CAG and CTG repeat hairpin structure containing about 46 repeats can be generated in cells. Thus, the hairpin substrates with 7 more repeats in the upper strand than the lower strand used in our study can be readily generated as a natural DNA replication and repair intermediate in cells. Here, we used the substrates containing an 8-oxoG and abasic lesion in the loop region of a TNR hairpin to mimic a damaged TNR hairpin and its repair intermediate that can be generated during BER to study the mechanism of BER in a hairpin loop. We found that BER was not significantly affected by a TNR hairpin.

In this study, we have identified a mechanism by which APE1 prevents TNR expansion with an in vitro biochemical approach. Since there is no in vivo system that has been developed to study a mechanism that involves BER protein coordination in modulating TNR stability as yet, development of such a type of in vivo system would be helpful for further elucidating this mechanism in mammalian cells.

Our previous studies have shown that alkylated DNA damage can be induced in expanded GAA repeats associated with Friedreich’s ataxia by the chemotherapeutic drug, temozolomide, and BER of the alkylated DNA damage subsequently results in large GAA
repeat deletions \((146)\). This occurs as a result of the formation of a TTC loop in the template strand and a downstream GAA repeat flap. Pol β then skips over the loop, and FEN1 cleaves the GAA repeat flap allowing the removal of more GAA repeats than pol β synthesizes. This subsequently leads to repeat deletions \((146)\). The study indicates that chemotherapeutically induced DNA damage can shorten expanded TNR tracts via BER, and thus potentially be used as novel treatment of TNR expansion-induced neurodegeneration. In this study, we further demonstrated that BER of a DNA base lesion in a hairpin loop resulted in removal of the hairpin through the coordination among APE1 3'-5' exonuclease activity, FEN1 flap cleavage activity, and LIG I activity, indicating that the removal of a TNR hairpin via the coordination among major BER enzymes during BER in a hairpin loop also serves as one of the mechanisms that shortens expanded TNR tracts. Our results suggest that multiple mechanisms underlie the shortening of expanded TNR tracts induced by oxidative or alkylating DNA damage via BER as a potential therapy for TNR expansion-related neurodegenerative diseases.

In CHAPTER 1, it was demonstrated that BER of an oxidized DNA base lesion in the loop region of a CAG hairpin can result in hairpin removal and prevention or attenuation of expansions. This was the first evidence to indicate that TNR hairpin removal can be accomplished through the coordinated activities of the key BER proteins APE1, FEN1, pol β, and LIG I. We demonstrated a novel role of the 3’-5’ exonuclease activity of APE1 in resolving the upstream 3’-flap of the double-flap intermediate of TNR hairpin repair. This occurred in coordination with removal of the 5’-flap by FEN1, which resulted in resolution of the BER intermediate and complete or partial hairpin removal, thus
preventing or attenuating expansions. This study identified a novel pathway that cells use to prevent TNR expansions. Identifying these pathways is key to identifying novel targets for the therapeutic treatment of TNR expansion-associated diseases.

Base excision repair in a larger CAG hairpin resulted in the production of some shortened expanded products which resulted from partial hairpin removal. This was the result of secondary structures, such as hairpins, that formed in the longer flaps produced by APE1 cleavage of the loop of the larger hairpin. This indicates that FEN1 was unable to cleave the long flap containing a relatively stable hairpin, allowing the extra repeats to become incorporated as an expansion. The activity of FEN1 has been found to be stimulated by several repair cofactors, including PCNA. In CHAPTER 2, we examined the ability of the BER cofactor PCNA to facilitate removal of the extra repeats and prevent TNR expansions.
CHAPTER 2: PROLIFERATING CELL NUCLEAR ANTIGEN PREVENTS TRINUCLEOTIDE REPEAT EXPANSIONS BY PROMOTING REPEAT DELETION AND HAIRPIN REMOVAL

ABSTRACT

Base lesions of DNA and base excision repair (BER) within trinucleotide repeat (TNR) tracts modulate repeat instability through the coordination among the key BER enzymes DNA polymerase β, flap endonuclease 1 (FEN1) and DNA ligase I (LIG I). However, it remains unknown whether BER cofactors can also alter TNR stability. In this study, we discovered that proliferating cell nuclear antigen (PCNA), a cofactor of BER, promoted CAG repeat deletion and removal of a CAG repeat hairpin during BER in a duplex CAG repeat tract and CAG hairpin loop. We showed that PCNA stimulated LIG I activity on a nick across a small template loop during BER in a duplex (CAG)\textsubscript{20} repeat tract promoting small repeat deletions. Surprisingly, we found that during BER in a hairpin loop, PCNA promoted reannealing of the upstream flap of a double-flap intermediate, thereby facilitating the formation of a downstream flap and stimulating FEN1 cleavage activity and hairpin removal. Our results indicate that PCNA plays a critical role in preventing TNR expansions by cooperating with BER enzymes to modulate the structures of dynamic DNA. Our study provides the first evidence that PCNA prevents TNR expansions during BER by promoting repeat deletion and removal of a TNR hairpin.

INTRODUCTION

Over 40 human neurodegenerative diseases are caused by trinucleotide repeat
(TNR) expansions, including Huntington’s disease (CAG/CTG), myotonic dystrophy (CTG/CAG), and Friedreich’s ataxia (GAA/TTC), among others (84, 95). The instability of the DNA results from the formation of non-B form DNA secondary structures such as hairpins, triplexes, sticky DNA, and tetraplexes (102, 105, 114) during DNA replication and repair, recombination and gene transcription (89, 101, 102, 108). The TNR tracts contain tandem purines that are hotspots for the formation of oxidized lesions (150), and repeated cycles of base damage and inefficient DNA base excision repair (BER) lead to cumulative TNR expansions through a “toxic oxidation cycle” that could subsequently result in the onset of disease (95, 142). The BER of oxidized base lesions in a TNR tract has been shown to modulate repeat instability (46, 67, 105, 111, 117, 147) as the result of a loss of coordination between BER proteins and cofactors (46, 67, 111, 117) caused by the formation of secondary structures. We have recently shown that oxidative and alkylated DNA damage within TNR tracts induce large repeat deletions and small expansions through BER, indicating that DNA base lesions and BER also play an important role in mediating TNR deletions (111, 146-148). We have further demonstrated that the location of a lesion within a duplex TNR tract determines whether a TNR deletion or expansion occurs as a result of the formation of secondary structures in either the damaged or template strand (147). A base lesion located at the 5’-end of a TNR tract induced expansions as a result of formation of a large hairpin in the damaged strand which interrupts efficient flap cleavage by flap endonuclease 1 (FEN1) 5’-flap cleavage. A lesion located in the middle of the repeat tract induced deletions as a result of formation of a template hairpin that is bypassed by pol β (147). Furthermore, we found that a bulky oxidized base lesion, 5’,8-cyclo-2’-deoxyadenosine (cdA), in a (CAG)20 template strand caused CTG repeat deletions.
during BER and DNA lagging strand maturation by inducing formation of a small CAG repeat loop that was subsequently skipped over by pol β lesion bypass synthesis (148). We showed that the deletions were dependent on the location of the base lesion (148). Moreover, we have found that BER of a lesion located in the loop of a TNR hairpin leads to hairpin removal and the prevention or attenuation of expansions (155). Hairpin removal occurs when the hairpin is converted into a double-flap intermediate, containing an upstream 3’-flap and a downstream 5’-flap, as a result of APE1 5’-incision of the abasic site. We found that these hairpins were converted to double-flap intermediates that can be resolved by the coordinated actions of FEN1 with the 3’-5’ endonuclease Mus81/Eme1 and the 3’-5’ exonuclease activity of AP endonuclease 1 (APE1), which resolve the downstream and upstream flaps of the intermediates, respectively (155, 222). These findings indicate that the type of base lesion and its location are critical in determining whether BER facilitates or prevents TNR instability.

Efficient BER is mediated by functional coordination among the repair enzymes as well as their coordination with BER cofactors. During long-patch BER, pol β gap-filling synthesis coordinates with FEN1 flap cleavage via a “Hit and Run” mechanism in which pol β gap-filling synthesis creates a flap for FEN1 cleavage that in turn generates an additional single-nucleotide gap which is then filled in by pol β (57). Efficient BER also relies on coordination of the BER core enzymes with repair cofactors such as poly(ADP-ribose) polymerase 1 (PARP1) (71), proliferating cell nuclear antigen (PCNA) (75, 210, 223, 224), X-ray repair cross-complementing 1 (XRCC1) (76-78), and high mobility group box 1 (HMGB1) (79). PARP1 acts as a nick sensor protein and stimulates FEN1 flap cleavage, thereby promoting pol β DNA synthesis (71). XRCC1 physically interacts with
pol β and DNA ligase III (LIG III) to enhance the efficiency of LIG III during single-nucleotide BER (76-78). HMGB1 can stimulate pol β DNA synthesis and FEN1 flap cleavage, increasing the efficiency of long-patch BER (46, 79). Disruption of the coordination between pol β and FEN1 by a CAG repeat hairpin leads to pol β multi-nucleotide gap-filling synthesis and low efficiency of long-patch BER resulting in CAG repeat expansions (67). Thus, efficient BER that is mediated by BER protein-protein interactions and coordination plays a crucial role in preventing TNR expansions. The role of repair cofactors in maintaining trinucleotide repeat stability has yet to be explored.

Proliferating cell nuclear antigen acts as the sliding clamp to increase the processivity of mammalian replication polymerases. PCNA can also interact with other enzymes and proteins involved in DNA replication, DNA repair, and cell cycle control (45, 73, 74, 225-227). PCNA is a toroidal homotrimeric protein that encircles double-stranded DNA and slides along the duplex DNA. PCNA can interact with DNA glycosylases (228), pol β (73), FEN1 (74, 223, 229-231), and DNA ligase I (LIG I) (75, 232), and has been found to localize to the nucleus in response to treatment with DNA alkylating agents such as methyl methanesulfonate (MMS) (233), indicating its crucial role in mediating efficient BER (210, 224, 226). PCNA and FEN1 physically interact through the PCNA-interacting protein (PIP) box at the C-terminus of FEN1 and a hydrophobic cleft on the proximal face of the PCNA trimer (234, 235). The interaction occurs after FEN1 loads onto the free 5’-end of the flap and tracks down to the base. The PCNA protein then slides along the duplex from upstream of the flap to interact with FEN1 at the flap base (74). This interaction stabilizes FEN1 binding to the substrate (74) and stimulates the FEN1 flap cleavage activity (229, 230), thereby facilitating efficient processing of Okazaki fragments and long patch
BER (74, 223, 236). It has been shown that disruption of the interaction between yeast PCNA and the yeast FEN1 homolog, Rad 27, as well as that between yeast PCNA and LIG I, increases CAG repeat expansions and deletions (138), indicating that the interactions between PCNA and FEN1 and LIG I are critical in the maintenance of TNR stability. Moreover, mutations of PCNA that disrupt its interaction with FEN1 have been shown to be associated with DNA repair disorders with neurodegenerative symptoms (237). The results suggest that PCNA plays a crucial role in the maintenance of genome stability. We further hypothesized that PCNA may prevent TNR expansions by coordinating with FEN1 and other repair enzymes during BER in a TNR tract. To test this hypothesis, we explored the effects of PCNA on TNR instability during BER in a duplex CAG repeat tract and in the loop of a CAG repeat hairpin. For the first time, we found that PCNA prevents TNR expansions by promoting CAG repeat deletions and removal of a CAG repeat hairpin during BER. We further demonstrated that the formation of secondary structures in a CAG repeat tract determined whether PCNA can modulate CAG repeat instability by modulating the activities of BER enzymes.

MATERIALS AND METHODS

Materials

The DNA oligonucleotides were synthesized by Integrated DNA Technologies (IDT, Coralville, IA, USA). Deoxynucleotide 5’-triphosphates (dNTPs) were from Fermentas (Glen Burnie, MD, USA). T4 polynucleotide kinase and terminal deoxynucleotidyl transferase were purchased from Thermo Fisher Scientific (Waltham,
Radionucleotides $\gamma^{32}$P ATP (6000 mCi/mmol) and Cordycepin 5'-triphosphate $3'$-$\alpha^{32}$P (5000 mCi/mmol) were purchased from Perkin Elmer Inc. (Boston, MA, USA). Micro Bio-Spin 6 chromatography columns were purchased from Bio-Rad Laboratories (Hercules, CA, USA). All standard chemical reagents were purchased from Sigma-Aldrich (St. Louis, MO, USA) and Thermo Fisher Scientific (Waltham, MA, USA). PCNA was purchased from Enzymax (Lexington, KY, USA). Pol $\beta$ was a generous gift from Dr. Samuel H. Wilson at the National Institute of Environmental Health Sciences/National Institutes of Health. APE1, FEN1, and LIG I were expressed in *Escherichia coli* and purified as described below.

**Oligonucleotide substrates**

Double-flap oligonucleotide substrates were prepared as described previously, with slight modifications (155). Briefly, substrates mimicking the double-flap intermediates with a 5'-Tetrahydrofuran (THF) residue were constructed by annealing an upstream primer containing a 3'-(CAG)$_4$ or (CAG)$_6$ flap and a downstream primer containing a 5'-(CAG)$_3$ or (CAG)$_6$ flap with the template strand containing (CTG)$_7$ or (CTG)$_{10}$, respectively, at a molar ratio of 1:3:3. Substrates containing a (CAG)$_{20}$ duplex with a THF in the first or tenth repeat unit were constructed by annealing a (CAG)$_{20}$-containing oligonucleotide containing a THF lesion with the complementary template strand containing (CTG)$_{20}$ at a ratio of 1:3. A substrate mimicking a duplex (CAG)$_{20}$ repeat intermediate with an abasic lesion in the first repeat pre-cleaved with APE1 and FEN1 was constructed by annealing an upstream primer containing a 3’-CA and a downstream primer containing a
Table 2. Oligonucleotide sequences

<table>
<thead>
<tr>
<th>Oligonucleotides</th>
<th>nt</th>
<th>Sequence (5' -3')</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Downstream</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D1</td>
<td>40</td>
<td>pF CAG CAG CAG CAG CAG CAG TA CGT AGA CTT ACT CAT TGC</td>
</tr>
<tr>
<td>D2</td>
<td>50</td>
<td>pF CAG CAG CAG CAG CAG CAG CAG CAG TA CGT AGA CTT</td>
</tr>
<tr>
<td>D3</td>
<td>76</td>
<td>pAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAG CAG CAG CAG TA CGT AGA CTT ACT CAT TGC</td>
</tr>
<tr>
<td><strong>Template</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>61</td>
<td>GCA ATG AGT AAG TCT ACG TAC TGC TGC TGC TGC TGC TGC</td>
</tr>
<tr>
<td>T2</td>
<td>70</td>
<td>GCA ATG AGT AAG TCT ACG TAC TGC TGC TGC TGC TGC TGC</td>
</tr>
<tr>
<td>T3</td>
<td>100</td>
<td>GCA ATG AGT AAG TCT ACG TA 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 TA CGG ATG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CTA GAT GAC TCG</td>
</tr>
<tr>
<td><strong>Upstream</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U1</td>
<td>40</td>
<td>CGA GTC ATC TAG CAT CCG TA CAG CAG CAG CAG CAG CAG CAG</td>
</tr>
<tr>
<td>U2</td>
<td>49</td>
<td>CGA GTC ATC TAG CAT CCG TA CAG CAG CAG CAG CAG CAG CAG</td>
</tr>
<tr>
<td>U3</td>
<td>100</td>
<td>CGA GTC ATC TAG CAT CCG TA CAF CAG CAG CAG CAG CAG CAG</td>
</tr>
<tr>
<td>U4</td>
<td>100</td>
<td>CGA GTC ATC TAG CAT CCG TA CAG CAG CAG CAG CAG CAG CAG</td>
</tr>
<tr>
<td>U5</td>
<td>22</td>
<td>CGA GTC ATC TAG CAT CCG TA CA</td>
</tr>
</tbody>
</table>

F: tetrahydrofuran, THF
phosphorylated AG(CAG)\textsubscript{18} to a template strand containing (CTG)\textsubscript{20} at a ratio of 1:3:3. Oligonucleotide sequences are listed in Table 2. Substrates were labeled with \textsuperscript{32}P at the 5’- or 3’-end of the damaged strand, upstream or downstream primer, or template strand, as indicated.

**Protein expression and purification**

Flap endonuclease 1 and APE1 were purified as described previously (222). Briefly, FEN1 was expressed in *E. Coli* BL21(DE3). Two liters of lysogeny broth (LB) medium cultures were incubated at 37°C at 225 rpm until OD\textsubscript{600} reached 0.6. The protein expression was induced with 1 mM IPTG for 3.5 hours and harvested by centrifugation at 2600 rpm for 45 minutes at 4°C. Cells were lysed with a French press cell disruptor (Glen Mills, Clifton, NJ, USA) in lysis buffer which contained 30 mM 4-(2-hydroxyethyl)-1-piperazineethane-sulfonic acid (HEPES), pH 7.5, 30 mM KCl, 1 mM dithiothreitol (DTT), 1 mM EDTA, 1 mM phenylmethylsulfonyl fluoride (PMSF), and 0.5% inositol. The supernatant was subjected to purification through a Sepharose Q, CM sepharose and phenyl sepharose column sequentially operated by an AKTA Fast Protein Liquid Chromatography system (FPLC) (GE Healthcare, Piscataway, NJ, USA). The peak fractions were combined and dialyzed into buffer containing 30 mM HEPES, pH 7.5, 30 mM KCl, 0.5% inositol, and 1 mM PMSF. Samples were then loaded onto a 1-ml Mono-S column (GE Healthcare, Piscataway, NJ, USA), and eluted using a linear gradient of KCl (30 mM to 2 M). Purified FEN1 was aliquoted and frozen at -80°C until further use.

AP endonuclease 1 was expressed in *E. Coli* BL21(DE3). Two liters of LB medium cultures were incubated at 37°C at 225 rpm until OD\textsubscript{600} reached 0.6. The APE1 expression
was induced by 0.5 mM IPTG for 3.5 hours. Cells were harvested by centrifugation at 2500 rpm for 30 minutes at 4 °C. Cells were lysed in lysis buffer, which contained 50 mM HEPES, pH 7.5, 30 mM NaCl, 1 mM dithiothreitol (DTT), 1 mM EDTA, and 1 mM PMSF. The supernatant was subjected to sequential purification by Q sepharose, CM sepharose, Mono-S and phenyl sepharose columns operated by an AKTA FPLC. Purified APE1 was aliquoted and frozen at -80 °C until further use.

The LIG I protein was expressed in *E. Coli* BL21(AI). Six liters of LB medium were inoculated with one colony each of the transformed BL21(AI) cells and were incubated overnight without shaking. The cells were then incubated at 37°C at 225 rpm until OD$_{595}$ reached 0.6. LIG I expression was induced by 1 mM IPTG for 24 hours at 15-18 °C. Cells were harvested by centrifugation at 2500 rpm for 30 minutes at 4 °C. Cell pellets were collected and lysed in lysis buffer, which contained 50 mM Tris-HCl, pH 7.5, 50 mM NaCl, 1 mM EDTA, 1 mM DTT, 0.1% Nonidet P-40, 1 mM PMSF, and 1 tablet of cOmplete protease inhibitors (Roche, Indianapolis, IN) and subjected to French Press at 150,000 PSI. The cell lysates were centrifuged at 12,000 rpm for 30 minutes at 4 °C. The supernatant was loaded onto a 20-ml P11 Phosphocellulose column, with fractions eluted using a linear gradient of NaCl (50 mM to 600 mM). The peak fractions were combined and dialyzed into buffer containing 30 mM HEPES, pH 7.0, 30 mM KCl, 0.1% inositol, and 1mM PMSF. Samples were then loaded onto a 10-ml Q sepharose column and eluted using a linear gradient of KCl (30 mM to 2 M). Peak fractions were combined and dialyzed into buffer containing 50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 7 mM 2-mercaptoethanol, 10 mM imidazole, 0.5% inositol, and 1 mM PMSF. Samples were loaded onto a 4-ml Nickel-Nitrilotriacetic acid (Ni-NTA) column, with fractions eluted using a linear gradient
of imidazole (10 mM to 600 mM). Purified LIG I was aliquoted and frozen at -80 °C until further use.

**Reconstituted BER assay**

*In vitro* BER of an abasic site analog, THF, in a (CAG)_{20} duplex or at the 5'-end of the downstream primer of the double-flap substrate was carried out by incubating 25 nM substrate with the indicated concentrations of APE1, FEN1, LIG I, pol β, and PCNA. Substrates containing a (CAG)_{20} duplex with a THF residue were pre-incubated with APE1 at 37 °C for 30 minutes and then subjected to BER reactions. All reactions were carried out in reaction buffer containing 30 mM HEPES, pH 7.8, 40 mM KCl, 5% glycerol, and 0.1 mg/ml bovine serum albumin (BSA), with 5 mM MgCl₂, 2 mM ATP, and 50 μM dNTPs. The 20 μl BER reaction mixtures were incubated at 37 °C for 15 minutes, and reactions were terminated by the addition of 20 μl stopping buffer containing 95% formamide and 10 mM EDTA. Reaction mixtures were then denatured at 95 °C for 10 minutes and separated by 15% urea-denaturing polyacrylamide gel electrophoresis. Substrates and products were detected and analyzed using a Pharos FX Plus PhosphorImager from Bio-Rad Laboratories (Hercules, CA, USA). All experiments were done in triplicate.

**BER enzymatic activity assay**

The activities of FEN1 flap cleavage, pol β DNA synthesis, and LIG I on all substrates were measured by incubating 25 nM substrates with the indicated concentrations of FEN1, pol β, LIG I, and PCNA in reaction buffer containing 30 mM HEPES, pH 7.8,
40 mM KCl, 5% glycerol, and 0.1mg/ml bovine serum albumin (BSA), with 5 mM MgCl₂, 2 mM ATP, and 50 µM dNTPs. The 20 µl BER reaction mixtures were incubated at 37 ºC for 15 minutes, and reactions were terminated by the addition of 20 µl stopping buffer. Reaction mixtures were then denatured at 95 ºC for 10 minutes and separated by 15% urea-denaturing polyacrylamide gel electrophoresis. Substrates and products were detected and analyzed using a Pharos FX Plus PhosphorImager from Bio-Rad Laboratories.

**Probing of hairpin structures and flaps by S1 Nuclease digestion**

The formation of hairpin and flap structures formed by the duplex and double-flap substrates was probed using S1 Nuclease as described previously, with slight modifications as described (155). Briefly, substrates (25 nM) were incubated with the indicated concentrations of S1 Nuclease in its reaction buffer at 37 ºC for 1, 5, 10, and 15 minutes in the absence and presence of PCNA (100 nM). Reaction mixtures were then subjected to proteinase K digestion at 55 ºC for 30 minutes to remove the S1 Nuclease. Reactions were stopped with 15 µl of stopping buffer. Reaction mixtures were then denatured at 95 ºC for 10 minutes and separated by 18% urea-denaturing polyacrylamide gel electrophoresis. Substrates and products were detected and analyzed using a Pharos FX Plus PhosphorImager from Bio-Rad Laboratories.
RESULTS

PCNA slightly stimulates CAG repeat deletion during BER of an abasic site in a duplex (CAG)_{20} repeat tract in the absence of pol β.

Since our previous studies have shown that TNR instability induced by a base lesion is governed by the location of a lesion in a duplex TNR tract during BER (147, 148), we initially determined whether PCNA can facilitate efficient BER of a base lesion located at the first (5’-end) and tenth (middle) repeat of a duplex (CAG)_{20} repeat tract to prevent repeat expansions. We reconstituted BER using the substrates containing a duplex (CAG)_{20} tract with an abasic lesion, the THF residue, at the 5’-end (Figure 2.1, left panel) or in the middle (Figure 2.1, right panel). We found that in the absence of pol β, BER of the base lesion at both locations resulted in the production of a repaired product with a one repeat deletion (Figure 2.1, lanes 4-6 and lanes 13-15). The amount of the deletion products was slightly increased in the presence of 50 nM and 100 nM PCNA (Figure 2.1, lanes 5-6 and lanes 14-15). The presence of pol β resulted in the production of the unexpanded product with the substrates containing an abasic lesion at either the 5’-end or the middle of the (CAG)_{20} repeat tract (Figure 2.1, lane 7 and lane 16). However, the presence of PCNA at 50 nM and 100 nM did not significantly affect the amount of the repaired products (Figure 2.1, lanes 8-9 and lanes 17-18). The results indicate that PCNA slightly stimulated CAG repeat deletions during BER of an abasic lesion in a duplex (CAG)_{20} tract.
Figure 2.1 PCNA stimulates repeat deletion during BER in a duplex (CAG)\textsubscript{20} tract

BER reactions were reconstituted with substrates containing a THF residue at either the first (left panel) or tenth (right panel) CAG repeat unit of a (CAG)\textsubscript{20} duplex tract in the absence and presence of PCNA. Lanes 1 and 10 indicate the substrate alone. Lanes 2 and 11 correspond to the reactions with substrates precut by APE1 (25 nM). Lanes 3 and 12 correspond to the reactions with the APE1 precut substrates and PCNA (100 nM) alone. Lanes 4-6 and 13-15 correspond to BER reactions reconstituted with APE1 (25 nM), FEN1 (5 nM) and LIG I (1 nM) in the absence and presence of PCNA (50 nM or 100 nM). Lanes 7-9 and 16-18 correspond to BER reconstituted with APE1 (25 nM), FEN1 (5 nM), LIG I (1 nM), and pol β (10 nM) in the absence and presence of PCNA. Substrates were \textsuperscript{32}P-labeled at the 3’-end of the damaged strand and are illustrated above each gel.
**Formation of secondary structures in a duplex (CAG)_{20} repeat tract can be modulated by PCNA**

We have previously shown that the formation of template hairpins mediates repeat deletions during BER in the context of a TNR tract (111, 147, 148). To further test whether the secondary structures formed in a duplex (CAG)_{20} repeat tract can be modulated by PCNA, we initially probed for the formation of secondary structures in the template strand of the (CAG)_{20} substrate with an abasic lesion at the 5’-end (Figure 2.2A) or the middle (Figure 2.2B) of the repeat tract using S1 nuclease that specifically cleaves single-stranded DNA. For the damage at the 5’-end, S1 nuclease resulted in cleavage products of 19 nt-23 nt in the absence and presence of PCNA (Figure 2.2A, lanes 3-6 and lanes 10-13) indicating that a small loop containing one CTG formed on the template strand (Figure 2.2A, the scheme below the gels) and PCNA failed to alter the loop structure. S1 nuclease cleavage on the template strand of the substrate with an abasic lesion in the middle resulted in products with 46 nt, 49 nt and 51 nt (Figure 2.2B, lanes 3-6) indicating the formation of a (CTG)_{2} loop on the template strand. In the presence of PCNA, S1 nuclease cleavage on the template strand resulted in products with 51 nt and 52 nt (Figure 2.2B, lanes 10-13), indicating the formation of a loop with one CTG. S1 nuclease cleavage on the upstream strand of the damaged strand of the substrate resulted in cleavage products of 44 nt-48 nt (Figure 2.2C, lanes 3-6) in the absence of PCNA, indicating the formation of a (CAG)_{2} flap. However, in the presence of PCNA, S1 nuclease only resulted in a 48 nt product (Figure 2.2C lanes 10-13), indicating the formation of a flap with one CAG. The results indicated that PCNA promoted reannealing of the upstream repeat flap to the template strand, presumably by sliding towards the downstream strand. The results also indicated
that PCNA altered the structure of a (CTG)$_2$ loop on the template strand and converted it to a smaller (CTG)$_1$ loop.

**Figure 2.2 A bubble forms in the template of a duplex (CAG)$_{20}$ tract during BER**

The formation of secondary structures in the template strand of the duplex (CAG)$_{20}$ substrate with damage at the 5’-end (A) or in the middle (B), was probed using 5U S1 nuclease in the absence (left panel) and presence (right panel) of 100 nM PCNA. (C) The formation of a flap in the upstream strand of the substrate with damage in the middle was also probed with 5U S1 nuclease in the absence (left panel) and presence of PCNA (right panel). Lanes 1 and 8 indicate the substrate alone. Lanes 2 and 9 indicate the substrate precut by APE1. Lanes 3-6 and lanes 10-13 indicate the reactions containing APE1 precut substrate with or without 100 nM PCNA along with 5U S1 nuclease at the time intervals of 1, 5, 10 and 15 minutes. Lanes 7 and 14 indicate size markers whose length is indicated to the right of the gel. Substrates were $^{32}$P-labeled at the 5’-end of the template strand or 3’-end of the damaged strand as illustrated above each gel. The schematic representation of the loops and flaps formed on the template and damaged strands are illustrated below the gels.
Figure 2.2 A bubble forms in the template of a duplex (CAG)$_{20}$ tract during BER

The formation of secondary structures in the template strand of the duplex (CAG)$_{20}$ substrate with damage at the 5’-end (A) or in the middle (B), was probed using 5U S1 nuclease in the absence (left panel) and presence (right panel) of 100 nM PCNA. (C) The formation of a flap in the upstream strand of the substrate with damage in the middle was also probed with 5U S1 nuclease in the absence (left panel) and presence of PCNA (right panel). Lanes 1 and 8 indicate the substrate alone. Lanes 2 and 9 indicate the substrate precut by APE1. Lanes 3-6 and lanes 10-13 indicate the reactions containing APE1 precut substrate with or without 100 nM PCNA along with 5U S1 nuclease at the time intervals of 1, 5, 10 and 15 minutes. Lanes 7 and 14 indicate size markers whose length is indicated to the right of the gel. Substrates were $^{32}$P-labeled at the 5’-end of the template strand or 3’-end of the damaged strand as illustrated above each gel. The schematic representation of the loops and flaps formed on the template and damaged strands are illustrated below the gels.
Figure 2.2 A bubble forms in the template of a duplex (CAG)$_{20}$ tract during BER

The formation of secondary structures in the template strand of the duplex (CAG)$_{20}$ substrate with damage at the 5'-end (A) or in the middle (B), was probed using 5U S1 nuclease in the absence (left panel) and presence (right panel) of 100 nM PCNA. (C) The formation of a flap in the upstream strand of the substrate with damage in the middle was also probed with 5U S1 nuclease in the absence (left panel) and presence of PCNA (right panel). Lanes 1 and 8 indicate the substrate alone. Lanes 2 and 9 indicate the substrate precut by APE1. Lanes 3-6 and lanes 10-13 indicate the reactions containing APE1 precut substrate with or without 100 nM PCNA along with 5U S1 nuclease at the time intervals of 1, 5, 10 and 15 minutes. Lanes 7 and 14 indicate size markers whose length is indicated to the right of the gel. Substrates were $^{32}$P-labeled at the 5'-end of the template strand or 3'-end of the damaged strand as illustrated above each gel. The schematic representation of the loops and flaps formed on the template and damaged strands are illustrated below the gels.
PCNA fails to stimulate FEN1 cleavage activity during BER in a duplex (CAG)\textsubscript{20} repeat tract

PCNA has been shown to directly interact with FEN1 and stimulate its flap cleavage during BER (229, 230). To determine whether PCNA can also stimulate FEN1 flap cleavage activity on a duplex (CAG)\textsubscript{20} repeat tract during BER, we examined the effects of PCNA on FEN1 cleavage during BER in the context of the (CAG)\textsubscript{20} tract with an abasic lesion located at the 5'-end (Figure 2.3, left panel) or in the middle (Figure 2.3, right panel) of the repeat tract. For both substrates, we found that in the absence and presence of 50 nM and 100 nM PCNA, similar amounts of cleavage products were generated by 5 nM FEN1 in the absence of pol β (Figure 2.3, compare lanes 5-6 with lane 4 and lanes 15-16 with lane 14). Similarly, in the presence of pol β, 50 nM and 100 nM PCNA did not significantly alter the amount of FEN1 cleavage products (Figure 2.3, compare lanes 9-10 with lane 8 and lanes 19-20 with lane 18). The results showed that PCNA failed to stimulate FEN1 flap cleavage activity during BER in a duplex (CAG)\textsubscript{20} tract.
PCNA does not alter pol β synthesis activity during BER of an abasic site in a duplex (CAG)\(_{20}\) repeat tract

PCNA has previously been shown to physically interact with pol β (73). To determine whether PCNA can stimulate pol β synthesis during BER in a duplex (CAG)\(_{20}\) tract, we examined pol β DNA synthesis on the (CAG)\(_{20}\) containing substrates (Figure 2.4)
in the absence and presence of 50 nM and 100 nM PCNA. We found that with 2 nM, 5 nM, and 10 nM pol β, similar amounts of pol β DNA synthesis products were formed in the absence and presence of PCNA (Figure 2.4, compare lanes 5-6 with lane 4, lanes 8-9 with lane 7, lanes 11-12 with lane 10, lanes 17-18 with lane 16, lanes 20-21 with lane 19, and lanes 23-24 with lane 22). Thus, the results indicate that PCNA did not affect pol β DNA synthesis during BER in a duplex (CAG)_{20} tract. The results indicated the PCNA-pol β interaction did not alter BER in a duplex (CAG)_{20} tract.
PCNA stimulates LIG I in the context of a small loop on the template strand of a duplex (CAG)\textsubscript{20} repeats tract

Because PCNA slightly stimulated the production of deletion products in the absence of pol \( \beta \) during BER in a duplex (CAG)\textsubscript{20} tract (Figure 2.1), but failed to stimulate
FEN1 and pol β (Figure 2.3 and Figure 2.4), it is possible that the stimulatory effect of PCNA resulted from its stimulation of LIG I activity across a small bubble or hairpin in the template strand, as it stimulates the activity of LIG I in duplex DNA (75). To test this, we examined the effect of 50 nM and 100 nM PCNA on the activity of LIG I (1 nM and 5 nM) on the substrate with a small (CTG)$_1$ repeat loop in the template strand. The substrate mimics the intermediate with a nick across a small (CTG)$_1$ loop on the template strand at the 5’-end that is generated by FEN1 cleavage of the sugar phosphate of an abasic lesion and adjacent nucleotide. The results showed that 50 nM and 100 nM PCNA significantly increased the amount of ligation product resulting from 1 nM and 5 nM LIG I activity (Figure 2.5, compare lanes 3-4 with lane 2 and lanes 6-7 with lane 5). This indicates that PCNA stimulated LIG I to seal a nick across a small CTG loop facilitating the formation of small deletion products during BER in a duplex (CAG)$_{20}$ repeat tract.
Figure 2.5 PCNA stimulates LIG I activity on a nick across a small loop on the template strand

The effect of PCNA on LIG I activity was determined with the substrate containing a nick across a small (CTG)_1 loop in the presence of 1 nM and 5 nM LIG I without or with 50 nM and 100 nM PCNA. The enzymatic reactions were performed in buffer containing 30 mM HEPES, pH 7.8, 5% glycerol, 40 mM KCl, 0.1 mg/ml BSA, 5 mM MgCl₂, 50 μM dNTPs, and 2 mM ATP at 37 °C for 15 minutes. Lane 1 represents the substrate (25 nM) alone. Lanes 2 and 5 correspond to the reactions containing the substrate and 1 nM and 5 nM LIG I, respectively. Lanes 3-4 and lanes 6-7 correspond to the reactions containing 1 nM or 5 nM LIG I with the presence of 50 nM and 100 nM PCNA. The substrate was ^{32}P labeled at the 5’-end of the upstream strand of the substrate.
PCNA stimulates the processing of a double-flap intermediate during BER in a CAG hairpin loop

Our previous studies have shown that during BER of an abasic lesion in the loop region of a CAG repeat hairpin, APE1 5’-incision converts the hairpin into a double-flap intermediate with an upstream 3’-flap and a downstream 5’-flap (155, 222). This intermediate was then processed by the coordinated actions among FEN1 5’-flap cleavage, Mus81/Eme1 3’-5’ endonuclease (155) and APE1 3’-5’ exonuclease (222), leading to removal of the hairpin and prevention of TNR expansions. To test whether PCNA can facilitate the removal of the hairpin by stimulating FEN1 5’-flap cleavage activity for processing of the double-flap intermediate, we reconstituted BER in the absence and presence of PCNA with a substrate containing an upstream 3’-(CAG)₄ flap and a downstream 5’-(CAG)₃ flap with a 5’-THF residue, as well as with a substrate containing a (CAG)₆ flap on both the upstream and downstream strands with a 5’-THF residue. These substrates mimic the double-flap intermediates which are formed during BER of a lesion in a (CAG)₇ and (CAG)₁₂ hairpin, respectively. On the substrate containing short (CAG)₃/(CAG)₄ flaps, we found that the presence of 50 nM and 100 nM PCNA significantly increased the amount of the unexpanded repaired product that has the same length as the template strand during BER without and with pol β (10 nM) (Figure 2.6, left panel, compare lanes 6-7 with lane 5 and lanes 9-10 with lane 8). On the substrate containing long (CAG)₆/(CAG)₆ flaps, PCNA (50 nM and 100 nM) did not significantly affect the production of the unexpanded repaired product in the absence of pol β (Figure 2.6, right panel, compare lanes 16-17 with lane 15). However, in the presence of 10 nM pol β, 50 nM and 100 nM PCNA resulted in a slight increase in the amount of unexpanded
repaired product (Figure 2.6, right panel, compare lanes 19-20 with lane 18). The results indicate that PCNA significantly promoted the resolution of the short double-flap intermediate and moderately facilitated the removal of the long double-flap intermediate leading to the removal of the CAG repeat hairpin during BER.
Figure 2.6 PCNA stimulates processing of double-flap intermediates resulting from a CAG repeat hairpin

The effects of PCNA on processing of the double-flap intermediates resulting from a CAG repeat hairpin were determined by reconstituting BER with double-flap substrates containing (CAG)$_3$/(CAG)$_4$ flaps (left panel) or (CAG)$_6$/(CAG)$_6$ flaps (right panel) that mimic an intermediate generated by APE1 5’-incision in a CAG repeat hairpin loop in the presence and absence of PCNA. Lanes 1 and 11 indicate the size markers of the unexpanded product. Lanes 2 and 12 represent the size markers of the expanded product.
PCNA stimulates FEN1 flap cleavage activity on a double-flap intermediate of hairpin repair

PCNA has been shown to stimulate FEN1 flap cleavage activity to facilitate DNA lagging strand maturation (74, 229) and long-patch BER (223, 230). To determine whether PCNA also promotes the resolution of the double-flap intermediate and subsequent hairpin removal by stimulating FEN1 5’-flap cleavage of the double-flap intermediate, we examined FEN1 cleavage activity on the shorter (CAG)$_3$/(CAG)$_4$ double-flap substrate and the longer (CAG)$_6$/(CAG)$_6$ double-flap substrate in the absence and presence of PCNA (Figure 2.7). We found that PCNA (50 nM and 100 nM) significantly increased the amount of FEN1 cleavage of the downstream flap of the short double-flap substrate independent of pol β (Figure 2.7, left panel, compare lanes 4-5 with lane 3 and lanes 8-9 with lane 7). However, the same concentrations of PCNA only slightly stimulated FEN1 cleavage of the downstream flap of the long double-flap substrate (Figure 2.7, right panel, compare lanes 13-14 with lane 12 and lanes 17-18 with lane 16). The results indicate that PCNA mainly stimulated FEN1 flap cleavage activity on the double-flap intermediate during BER in a small CAG repeat hairpin, promoting hairpin removal and the formation of unexpanded repaired products.
Figure 2.7 PCNA stimulates FEN1 flap cleavage on double-flap intermediates

The stimulation of FEN1 cleavage by PCNA on double-flap substrates that mimic an intermediate during BER in a CAG repeat hairpin was measured by measuring FEN1 activity in the absence and presence of PCNA on substrates containing (CAG)_3/(CAG)_4 flaps (left panel) or (CAG)_6/(CAG)_6 flaps (right panel). Lanes 1 and 10 correspond to the substrate only. Lanes 2 and 11 correspond to reaction mixture containing substrate and PCNA (100 nM) only. Lanes 3-5 and lanes 12-14 correspond to the reactions containing the substrates and FEN1 (0.5 nM or 2.5 nM) in the absence and presence of PCNA (50 nM or 100 nM). Lanes 6 and 15 correspond to reactions containing substrates and pol β (10 nM) in the presence of PCNA (100 nM). Lanes 7-9 and lanes 16-18 correspond to reactions with the substrates, FEN1 (0.5 nM or 2.5 nM) and pol β (10 nM) in the absence and presence of PCNA (50 nM or 100 nM). Substrates were ^32P-labeled at the 3'-end of the damage-containing strand and are illustrated above each gel.
The results further suggest that PCNA managed to slide toward the downstream flap to interact with FEN1, thereby stimulating its flap cleavage activity. We then hypothesized that this could result from reannealing of the upstream strand to the template strand as the upstream strand could be pushed by the sliding of PCNA along the duplex toward the downstream. To test this, we examined the formation of the flap or secondary structures in the upstream strand of the double-flap substrates in the absence and presence of PCNA (Figure 2.8). We found that in the absence of PCNA, S1 nuclease cleavage on the upstream strand of the short double-flap substrate led to the cleavage products of 25 nt, 26 nt, 27 nt, 28 nt, 30 nt, 31 nt, 34 nt, and 35 nt (Figure 2.8A, left panel, lanes 2-5), indicating that the upstream strand adopted a (CAG)₄ repeat flap (Figure 2.8A, the scheme below the gels). In the presence of PCNA, S1 nuclease failed to produce any cleavage products at 1 min (Figure 2.8A, lane 8). However, at 5 min, 10 min and 15 min, the nuclease cleavage only resulted in the products containing 27 nt, 30 nt, 31 nt, and 37 nt (Figure 2.8B, right panel, lanes 9-11). The reduced S1 nuclease cleavage indicates that the presence of PCNA promoted the reannealing of the upstream strand to the template strand, presumably during PCNA sliding toward the downstream strand. This subsequently resulted in the formation of a substrate with an annealed upstream and longer downstream flap allowing PCNA to interact with FEN1 and stimulate its cleavage activity for removing a long CAG repeat flap. Similarly, in the absence of PCNA, S1 nuclease cleavage on the upstream strand of the long (CAG)₆/CAG)₆ double-flap substrate resulted in the products with 22 nt, 25 nt, 28 nt, 31 nt, 34 nt, 37 nt, 40 nt, 43 nt, and 46 nt (Figure 2.8B, left panel lanes 3-5), indicating that the upstream strand formed a long (CAG)₇ flap. In the presence of PCNA, the nuclease cleavage resulted in no cleavage products after 1 minute of incubation with
S1 nuclease (Figure 6B, lane 8), and in products with 22-47 nt after 5-15 minutes of incubation (Figure 2.8B, right panel, lanes 9-11). At all times points in the presence of PCNA, there are large amounts of residual substrate that remain uncleaved by the S1 nuclease, indicating that much of the substrate is annealed and resistant to cleavage by S1 nuclease. The results indicate that PCNA sliding on the duplex DNA pushed the upstream strand to reanneal to the template strand (Figure 2.8B, the scheme below the gels). Moreover, we found that S1 nuclease failed to make any cleavage on the template strand of the double-flap substrates (Figure 2.9), indicating that no secondary structures formed on the template strand to block PCNA sliding.
Figure 2.8 Formation of a flap in the upstream strand of double-flap intermediates

(A) The formation of a flap in the upstream strand of the (CAG)$_3$/CAG)$_4$ double-flap substrate was probed using S1 nuclease in the absence (left panel) and presence (right panel) of 100 nM PCNA. Lanes 1 and 7 indicate the substrate only. Lanes 2-5 and lanes 8-11 indicate reactions containing substrate with or without 100 nM PCNA incubated with S1 nuclease from 1-15 minutes. Lanes 6 and 12 indicate size markers whose length is indicated to the right of the gel. (B) The formation of a flap in the upstream strand of the (CAG)$_6$/CAG)$_6$ double-flap substrate was probed using S1 nuclease. Lanes 1 and 7 indicate the substrate only. Lanes 2-5 and lanes 8-11 indicate the reactions containing the substrate with or without 100 nM PCNA incubated with S1 nuclease from 1-15 minutes. Lanes 6 and 12 indicate size markers whose length is indicated to the right of the gel. Substrates were $^{32}$P-labeled at the 5'-end of the upstream strands as illustrated above each gel. The schematic representations of the flaps formed are illustrated below the gels.
Figure 2.8 Formation of a flap in the upstream strand of double-flap intermediates

(A) The formation of a flap in the upstream strand of the (CAG)_3/(CAG)_4 double-flap substrate was probed using S1 nuclease in the absence (left panel) and presence (right panel) of 100 nM PCNA. Lanes 1 and 7 indicate the substrate only. Lanes 2-5 and lanes 8-11 indicate reactions containing substrate with or without 100 nM PCNA incubated with S1 nuclease from 1-15 minutes. Lanes 6 and 12 indicate size markers whose length is indicated to the right of the gel. (B) The formation of a flap in the upstream strand of the (CAG)_6/(CAG)_6 double-flap substrate was probed using S1 nuclease. Lanes 1 and 7 indicate the substrate only. Lanes 2-5 and lanes 8-11 indicate the reactions containing the substrate with or without 100 nM PCNA incubated with S1 nuclease from 1-15 minutes. Lanes 6 and 12 indicate size markers whose length is indicated to the right of the gel. Substrates were ^32P-labeled at the 5’-end of the upstream strands as illustrated above each gel. The schematic representations of the flaps formed are illustrated below the gels.
PCNA slightly facilitates pol β synthesis activity on the double-flap intermediate

Since our results indicate that PCNA sliding pushed the upstream strand of the double-flap substrates to reanneal to the template strand (Figure 2.8), this may further provide more 3’-termini for pol β DNA synthesis. Thus, it is possible that PCNA may
stimulate pol β DNA synthesis during BER in a hairpin by altering the dynamics of the double-flaps. To test this, we examined pol β DNA synthesis on the (CAG)$_3$/(CAG)$_4$ and (CAG)$_6$/(CAG)$_6$ double-flap substrates (Figure 2.10) in the absence and presence of 50 nM and 100 nM PCNA. We found that PCNA (50 nM and 100 nM) increased the amount of DNA synthesis products on the (CAG)$_3$/(CAG)$_4$ substrate generated by pol β at both 2 nM and 5 nM (Figure 2.10, left panel, compare lanes 4-5 with lane 3 and lanes 7-8 with lane 6). For the (CAG)$_6$/(CAG)$_6$ double-flap substrate, PCNA did not affect the DNA synthesis activity of pol β at 2 nM (Figure 2.10, right panel, compare lanes 12-13 with lane 11). However, it slightly stimulated pol β polymerase activity at 5 nM (Figure 2.10, compare lanes 15-16 with lane 14). Thus, the results indicate that PCNA slightly stimulated pol β DNA synthesis during BER in a double-flap intermediate resulting from a CAG repeat hairpin, and therefore led to efficient BER in the hairpin. This further indicates that PCNA stimulation of pol β DNA synthesis activity was mediated by the reannealing of the upstream flap to the template strand, creating more annealed 3’-termini that were extended by pol β.
120

**DISCUSSION**

In this study, we discovered that PCNA promoted CAG repeat deletions and removal of a CAG repeat hairpin (Figures 2.1 and 2.6), thereby preventing CAG repeat expansion. We showed that PCNA modulated the stability of the repeats by facilitating the activity of LIG I for ligating a nick in the context of a small template CTG loop (Figure

**Figure 2.10 PCNA slightly stimulates pol β DNA synthesis on double-flap intermediates during BER**

The effect of PCNA on pol β DNA synthesis on double-flap intermediates during BER was measured in the absence and presence of PCNA on substrates containing (CAG)₃/(CAG)₄ (left panel) or (CAG)₆/(CAG)₆ (right panel) double-flaps. Lanes 1 and 9 correspond to the substrate only. Lanes 2 and 10 correspond to the reactions containing the substrates and PCNA (100 nM) only. Lanes 3-5 and lanes 11-13 correspond to reactions containing substrate and pol β (2 nM) without or with PCNA (50 nM or 100 nM). Lanes 6-8 and lanes 14-16 correspond to substrates incubated with pol β (5 nM) without and with PCNA (50 nM or 100 nM). Substrates were ³²P-labeled at the 5’-end of the upstream strand and are illustrated above each gel.
as well as by stimulating FEN1 flap cleavage activity and pol β DNA synthesis for the resolution of a double-flap intermediate during BER in a CAG repeat hairpin loop (Figures 2.7 and 2.10). We demonstrated that the stimulatory effects of PCNA on prevention of CAG repeat expansion were governed by interactions among the dynamics of CAG repeats and their template, the formation of secondary structures and flaps, the PCNA sliding and processing of repair intermediates by pol β, FEN1 and LIG I (Figures 2.2, 2.5, 2.7, 2.8 and 2.10). Our results support a model by which PCNA prevents TNR expansions through stimulation of TNR deletions and the removal of a TNR hairpin (Figure 2.11). When a DNA base lesion occurs in either the duplex region or in the hairpin loop, a DNA glycosylase removes the damaged base leaving an abasic site. For BER of a lesion in a duplex TNR tract, APE1 incises the 5’-side of the abasic site allowing DNA slippage and the formation of a TNR loop that prevents PCNA from sliding toward the downstream to interact with pol β or FEN1 (Figure 2.11, the pathway on the left). However, this still allows PCNA to interact with LIG I and stimulate sealing of a nick across a small TNR loop on the template strand facilitating repeat deletion. During BER in a hairpin loop, APE1 5’-incision of the abasic site converts the hairpin into a double-flap intermediate. PCNA sliding pushes the upstream flap to reanneal to the template strand creating a 3’-terminus that can be extended by pol β. The reannealing of the upstream flap to the template further allows PCNA to slide to the downstream to interact with FEN1, thereby stimulating FEN1 cleavage of the downstream flap and ultimately facilitating the removal of the hairpin (Figure 2.11, the pathway on the right).
Figure 2.11 PCNA prevents TNR expansions during BER
Oxidized DNA base lesions induced by DNA damaging agents can occur in a duplex TNR tract (left pathway) or in the loop region of a TNR hairpin (right pathway). During BER within a duplex TNR tract, the formation of a loop structure within the template can block the advance of PCNA toward the downstream, preventing PCNA from interacting with FEN1 (left pathway). During BER of a lesion located in the loop of a TNR hairpin, the hairpin is converted into a double-flap intermediate. PCNA is then loaded from the upstream and slides toward the downstream pushing the reannealing of the upstream flap to the template strand. This creates a longer 5'-flap for FEN1 to be readily loaded. PCNA then interacts with FEN1 and stimulates FEN1 flap cleavage activity, thereby promoting removal of the hairpin and prevention of TNR expansions.
Here, we have identified a new role of PCNA in promoting TNR deletions and the removal of a TNR hairpin to prevent TNR expansions. PCNA has been shown to stimulate FEN1 flap cleavage activity (74). Yet it was unknown whether the stimulatory effect of PCNA would occur in TNR sequences that can form a variety of secondary structures such as hairpins and loops. In this study, we have shown that in a duplex (CAG)$_{20}$ that allowed the formation of a CTG repeat loop on the template strand during BER, PCNA failed to stimulate FEN1 activity or pol β activity (Figures 2.3 and 2.4). However, we found that PCNA exhibited the stimulatory effects on pol β DNA synthesis activity and FEN1 flap cleavage activity (Figures 2.7 and 2.10) on the double-flap substrates that could not form a loop or hairpin on the template strand (Figure 2.9). The results indicate that a loop or hairpin that formed in the template strand blocked the sliding of PCNA along the duplex toward the downstream, thereby preventing its direct interaction with FEN1 at the base of the downstream flap. Interestingly, our results showed that during BER in a TNR hairpin loop, due to the lack of secondary structures in the template strand, PCNA was able to facilitate the reannealing of the upstream strand to the template to stimulate pol β and FEN1 activities, promoting the removal of the TNR hairpin. This further suggests that structurally the domain of PCNA that interacts with the template strand is much more rigid than its domain that interacts with the damaged or upper strand of duplex DNA. This appears to allow the ring shaped protein to readily accommodate a flap structure on the damaged strand rather than a loop or hairpin on the template strand.

Previous studies have shown that LIG I activity can be stimulated by PCNA in duplex DNA of random sequence (75, 232). Here, we found that PCNA stimulated LIG I activity in a TNR duplex across a small CTG repeat loop on the template strand in the
absence of pol β (Figures 2.1 and 2.5). This indicates that PCNA managed to interact with LIG I to stabilize the binding of the enzyme to the nick opposite a small loop on the template strand, stimulating the ligation of the nick. The results further suggest that PCNA sliding toward the downstream was blocked by the loop on the template strand and failed to reach LIG I that bound to the nick generated by pol β DNA synthesis which was further downstream. Our results indicate that PCNA stimulation of LIG I plays an important role in promoting TNR deletion for preventing TNR expansions during BER in a duplex TNR tract.

A direct interaction has been shown between PCNA and pol β (73), although the biological significance of this interaction and/or stimulation of polymerase activity has yet to be identified. Our results show PCNA stimulation of pol β DNA synthesis activity on the double-flap substrates (Figure 2.10). We found that the stimulatory effect was mediated by PCNA sliding that facilitated the reannealing of the upstream flap to the template strand creating more annealed 3’-termini as pol β substrates. This has been supported by the significantly decreased S1 cleavage products on the upstream flaps in the presence of PCNA compared to in the absence of PCNA (Figure 2.8). Our results further indicate that PCNA cooperated with pol β functionally to facilitate the processing of a double-intermediate and removal of a hairpin during BER, thereby preventing repeat expansion.

Our study provides evidence of a BER cofactor, PCNA, playing an important role in facilitating TNR deletion and hairpin removal. Previously, we have also shown the 3’-5’ endo- and exonucleases, such as MUS81/EME1 and APE1, can prevent TNR expansions by promoting TNR hairpin removal (155, 222). This indicates that multiple mechanisms are employed in cells to prevent TNR expansions. Our results further indicate that the
coordination among the BER proteins and cofactors APE1, pol β, FEN1, LIG I and PCNA can promote TNR deletion and hairpin removal, and this serves as an efficient mechanism of preventing TNR expansions and shortening expanded TNR tracts. Our results further demonstrate that BER is a key pathway that can be targeted for the prevention and therapy of TNR expansion-associated neurodegenerative diseases.

In CHAPTER 1 and CHAPTER 2, it was demonstrated that BER in the loop region of a TNR hairpin can lead to hairpin removal and prevention of TNR expansions as a result of coordination between BER proteins and cofactors. This coordination between BER proteins and cofactors facilitates resolution of the double-flap intermediate formed during BER in a hairpin. The results in CHAPTER 1 and CHAPTER 2 indicate that FEN1 5’-flap cleavage activity in coordination with APE1 3’-5’ exonuclease activity facilitates removal of the 3’-flap and 5’-flap of the double-flap intermediate, respectively, and that PCNA facilitates removal of the hairpin by modulating the structure of the double-flaps to create a downstream flap for FEN1 cleavage and a substrate with a 3’-termini for pol β synthesis, thereby stimulating FEN1 removal of the 5’-flap. We have further shown that BER of a base lesion in a TNR duplex can lead to small TNR deletions as a result of ligation across a small template loop that is facilitated by the BER cofactor PCNA. However, studies have previously focused on the role of oxidized and alkylated DNA damage on TNR instability, with no investigation of the role that modifications of the sugar residue may play in governing TNR instability during BER. Modifications of the sugar residue can influence which pathway of BER is initiated in response to DNA damage and can influence how the lesions interact with key BER proteins. In CHAPTER 3, we investigated the effects of
DNA sugar modifications on BER of a lesion in a TNR hairpin and TNR duplex using synthesized DNA oligonucleotides and purified human BER proteins to carry out *in vitro* reconstitution of BER and assays of BER enzymatic activity.
CHAPTER 3: AN OXIDIZED ABASIC LESION INHIBITS DNA BASE EXCISION REPAIR RESULTING IN DNA STRAND BREAKS

ABSTRACT

DNA base lesions and base excision repair (BER) play a role in modulating trinucleotide repeat (TNR) instability that is associated with human neurodegenerative diseases and cancer. We have previously shown that BER of a model abasic lesion, tetrahydrofuran (THF), in the loop region of a TNR hairpin can lead to hairpin removal and the prevention of TNR expansions. However, it is unknown whether modifications of the sugar in an abasic lesion may affect BER in a TNR tract. In this study, we explored the effects of three types of abasic lesions on BER in a TNR tract, an unmodified native sugar, reduced sugar (THF), and the oxidized sugar 5’-(2-phosphoryl-1,4-dioxobutane) (DOB). DOB irreversibly inhibits DNA polymerase β (pol β) by forming a covalent bond with the pol β dRP lyase domain. We examined the effects of modifications to the sugar residue on pol β synthesis and FEN1 cleavage as well as their effects on BER efficiency. We found that the DOB lesion significantly inhibited pol β synthesis activity and slightly inhibited FEN1 cleavage activity, as compared to a flap containing a THF. We showed that decreased activities of repair proteins during BER of an oxidized abasic lesion decreased the overall efficiency of BER resulting in accumulation of BER intermediates with a DNA break. This is the first evidence that an oxidized abasic lesion inhibits BER in a TNR tract.
INTRODUCTION

Trinucleotide repeat (TNR) expansions are associated with over 40 human neurodegenerative diseases, including Huntington’s disease (CAG/CTG), Friedreich’s ataxia (GAA/TTC), Fragile X syndrome (CGG/CCG), and Myotonic Dystrophy (CTG/CAG), among others (84, 95). TNR instability results from the formation of DNA secondary structures such as hairpins, triplexes, quadruplexes, and sticky DNA (102, 105, 114), which form during DNA replication, repair, and recombination, and gene transcription (89, 101, 102, 108). Base excision repair (BER) within TNR tracts has been linked to repeat instability (46, 117), which results from a loss of coordination between key BER proteins and cofactors (46, 67, 105, 111, 117, 147). The formation of hairpins in a TNR tract can block flap endonuclease 1 (FEN1) 5’-flap cleavage, resulting instead in alternate flap cleavage in which FEN1 removes fewer nucleotides (67). When the addition of nucleotides by DNA polymerase β (pol β) exceeds the number of nucleotides removed by FEN1 an expansion occurs (67, 147). Alternatively, when hairpins form in the template strand of a TNR tract, pol β can skip over the hairpin, resulting in addition of fewer nucleotides than are removed by FEN1 cleavage and resulting in a deletion (111, 148). The tandem purines found in TNR tracts are hotspots of oxidative DNA damage (150), and repeated cycles of oxidative damage and DNA repair can lead to an accumulation of expansions through a “toxic oxidation cycle” (95, 117, 142). We have previously shown that both oxidative and alkylating DNA damage can result in TNR instability, and that the location of the lesion in the TNR tract determines the type of instability that results (111, 146, 147). Damage at the 5’-end of a TNR tract was found to induce expansions as a result of hairpin formation in the downstream region of the damaged strand, while damage in the
middle of the repeat tract induced deletions as a result of hairpin formation in the template strand (147). The formation of a cyclodeoxypurine lesion, 5’, 8-cyclo-2’-deoxypurine, in the template strand was also found to induce deletions by inducing formation of a small loop that is skipped over by pol β during BER (148). Thus, lesion location and lesion type both play determining roles in mediating TNR instability.

We have previously shown that BER of a lesion in the loop of a TNR hairpin leads to hairpin removal and the prevention of TNR expansions (155, 222). This is accomplished by the formation of a double-flap intermediate following AP endonuclease 1 (APE1) 5’-incision of the abasic site in the hairpin loop and the coordinated activity of FEN1 and a 3’-5’ nuclease such as APE1 or Mus81/Eme1, which remove the 5’ and 3’-flaps, respectively (155, 222). Previous studies have shown that guanines in the single-stranded loop of a TNR hairpin are particularly susceptible to the formation of oxidized lesions such as 8-oxo-guanine (8-oxoG) (142, 150), and that lesions formed in the stem region or adjacent to a hairpin are preferentially relocated into the loop (154). This indicates that oxidized lesions preferentially accumulate in the hairpin loop. However, whether modifications of the sugar residue in an abasic lesion have any effect on repair efficiency in a TNR tract has not previously been explored.

Many types of DNA damage are subject to the BER pathway, including oxidation, alkylation, and deamination, all of which result in formation of an abasic site as the first step of BER. Abasic sites are the most common DNA lesions, occurring at a rate of ~10,000 abasic sites per cell per day (1), and are highly mutagenic (81, 238-240). Thus, their efficient repair is key in the maintenance of genomic stability. Failure of BER machinery to complete the repair process can result in the accumulation of repair intermediates that
contain ssDNA breaks, further leading to genomic instability and/or cytotoxicity. In this study we explore whether modifications of the sugar residue in the abasic site affect the activity of BER proteins and BER efficiency in a TNR tract, and thus affect genomic stability. We used synthesized DNA oligonucleotides that mimic the intermediates formed during BER in a TNR hairpin and TNR duplex with a native abasic site, reduced abasic site (Tetrahydrofuran, THF), or the oxidized abasic site 5’-(2-phosphoryl-1,4-dioxobutane) (DOB) at the 5’-end of the downstream flap. Damage that results in a native sugar in the abasic site can be repaired via short-patch BER. Following formation of an abasic site, APE1 5’-incision of the abasic site creates a downstream deoxyribosephosphate (dRP) (241) that is removed by the dRP lyase activity of pol β (242). Damage such as oxidation to the sugar residue makes the abasic site refractory to pol β dRP lyase activity and results instead in long-patch BER in which pol β must carry out strand displacement synthesis to complete the repair process in coordination with FEN1 removal of a 5’-flap. The nature of TNRs allows DNA slippage, creating a gap greater than 1 nucleotide that forces repair into the long-patch pathway regardless of damage to the sugar residue. Native abasic sites can result from spontaneous base loss via depurination or from removal of a damaged base by a specific glycosylase protein such as 8-oxo-guanine DNA glycosylase 1 (OGG1). THF mimics a reduced abasic site and is conventionally used as an abasic site analog in biochemical studies. Oxidized abasic sites can result from exposure to exogenous environmental toxicants, ionizing radiation, chemotherapeutic drugs, or endogenous reactive oxygen species that result from cellular energy metabolism. The DOB lesion specifically can result from exposure to ionizing radiation or potent antitumor antibiotics such as neocarzinostatin (243, 244) as a result of radical formation following hydrogen
atom abstraction from the C5 position of the deoxyribose ring and occurs concomitantly with a single-strand break (245, 246), leaving the DOB abasic lesion at the 5’-end downstream of the break. These sugar modifications are compared in Figure 3.1. DOB forms a permanent covalent crosslink with Lys72 (and Lys84 to a lesser extent) in the dRP lyase domain of pol β (247), which is responsible for Schiff base formation during the β elimination reaction (248), resulting in a DNA-protein crosslink, irreversible inhibition of pol β dRP lyase following ~ 4 turnovers of the enzyme, and a decrease in pol β DNA synthesis activity (38, 247, 249, 250). Here we explore the effects of sugar modification on the activity of BER proteins during repair in a TNR tract and the downstream effects of crosslink formation on BER efficiency.

Figure 3.1 Abasic DNA lesions with modifications to the sugar residue
A native sugar (left), reduced sugar (middle), and oxidized sugar (right).

Modifications of the sugar in an abasic site can affect how the lesion interacts with the repair proteins and, thus, may affect the efficiency of BER. Because the DOB lesion
can react with pol β forming a DNA-protein crosslink, we hypothesized that this would reduce pol β synthesis activity in the TNR tract by removing available functioning pol β from solution and might inhibit loading of FEN1 onto the flap. We further hypothesized that this would interrupt coordination between the BER proteins and decrease BER efficiency by preventing formation of a ligatable nick for Ligase I (LIG I) activity. We explored how modifications of the sugar residue affect BER efficiency in a TNR tract by comparing enzymatic activities and BER efficiency during repair of the 3 types of abasic lesions. We found that the DOB lesion was still able to crosslink with pol β when the lesion was located at the end of a 5’-flap. This crosslink greatly inhibited pol β synthesis activity and slightly inhibited FEN1 cleavage activity as compared to the native sugar and reduced sugar. The interruption of efficient repair protein activities resulted in a decrease in repair efficiency which resulted in an accumulation of DNA breaks and could lead to genomic instability. This is the first evidence that an oxidized abasic lesion inhibits BER during repair in a TNR tract.

MATERIALS AND METHODS

Materials

The DNA oligonucleotides containing the DOB lesion were provided by Dr. Marc Greenberg (John’s Hopkins University, Baltimore, MD) and were prepared as described previously (251). All other DNA oligonucleotides were synthesized by Integrated DNA Technologies (IDT, Coralville, IA, USA). T4 polynucleotide kinase and terminal deoxynucleotidyltransferase were purchased from Thermo Fisher Scientific (Waltham,
MA, USA). Radionucleotides $[\gamma^{-32}\text{P}]$ ATP (6000mCi/mmol) and Cordycepin 5’-triphosphate 3’-$[\alpha^{-32}\text{P}]$ (5000 mCi/mmol) were purchased from Perkin Elmer Inc. (Boston, MA, USA). Deoxynucleotide 5’-triphosphates (dNTPs) were from Fermentas (Glen Burnie, MD, USA). Micro Bio-Spin 6 chromatography columns were purchased from Bio-Rad Laboratories (Hercules, CA, USA). All standard chemical reagents were purchased from Thermo Fisher Scientific (Waltham, MA, USA) and Sigma-Aldrich (St. Louis, MO, USA). Pol $\beta$ was a generous gift from Dr. Samuel H. Wilson at the National Institute of Environmental Health Sciences/National Institutes of Health. FEN1 and LIG I were expressed in *Escherichia coli* and purified as described below.

**Oligonucleotide substrates**

Oligonucleotide substrates were prepared as described previously (155). Briefly, substrates mimicking the BER intermediates with a 5’-THF, 5’-Uracil, or 5’-DOB residue were constructed by annealing an upstream primer containing a 3’-(CAG)$_3$ flap or a 3’-primer with no repeats and a downstream primer containing a 5’- (CAG)$_2$ flap with the template strand containing (CTG)$_5$ or (CTG)$_2$, respectively, at a molar ratio of 1:3:3. Oligonucleotide sequences are listed in Table 3. The Uracil-containing substrates were incubated with Uracil DNA glycosylase (UDG) to create an abasic site with a native sugar residue prior to each experiment. The DOB-containing substrates were provided by Dr. Marc Greenberg (John’s Hopkins University, Baltimore, MD) with o-nitrobenzyl protecting groups as indicated in Figure 3.2 (251). The DOB lesion was formed through photolysis (Figure 3.2 (251)) by exposing the annealed substrate to 365 nm UV for 20
minutes immediately prior to each experiment. Substrates were labeled with $^{32}$P at the 5’-end of the upstream strand or 3’-end of the downstream strand, as indicated.
### Table 3. Oligonucleotide sequences

<table>
<thead>
<tr>
<th>Oligonucleotides</th>
<th>nt</th>
<th>Sequence (5’ -3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Downstream</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D1</td>
<td>32</td>
<td>U CAG CAG CAG CAG ATG CCC GCC CCA CTC ATT GC</td>
</tr>
<tr>
<td>D2</td>
<td>32</td>
<td>F CAG CAG CAG CAG ATG CCC GCC CCA CTC ATT GC</td>
</tr>
<tr>
<td>D3</td>
<td>32</td>
<td>D CAG CAG CAG CAG ATG CCC GCC CCA CTC ATT GC</td>
</tr>
<tr>
<td><strong>Template</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>59</td>
<td>GCG CAA TGA GTG GGG CGG GCA TCT GCT GCT GCT GTA CGG GCG CTA GGC GAC TC</td>
</tr>
<tr>
<td>T2</td>
<td>50</td>
<td>GCG CAA TGA GTG GGG CGG GCA TCT GCT GTA CGG GCG CTA GGC GAC TCG CG</td>
</tr>
<tr>
<td><strong>Upstream</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U1</td>
<td>34</td>
<td>CGA GTC GCC TAG CGC CCG TAC AGC AGC AGC A</td>
</tr>
<tr>
<td>U2</td>
<td>20</td>
<td>CGA GTC GCC TAG CGC CCG TA</td>
</tr>
</tbody>
</table>

F: tetrahydrofuran, THF  
D: 5’-(2-phosphoryl-1,4-dioxobutane), DOB
Protein expression and purification

FEN1 was purified as described previously (222). Briefly, FEN1 was expressed in E. Coli BL21(DE3). Two liters of lysogeny broth (LB) medium were inoculated with one colony each of the transformed BL21(DE3) cells and were then incubated overnight without shaking. The cells were incubated at 37˚C at 225 rpm until OD\textsubscript{600} reached 0.6. The protein expression was induced with 1 mM IPTG for 3.5 hours and harvested by centrifugation at 2600 rpm for 45 minutes at 4˚C. Cells were lysed with a French press cell disruptor (Glen Mills, Clifton, NJ, USA) in lysis buffer which contained 30 mM 4-(2-hydroxyethyl)-1-piperazineethane-sulfonic acid (HEPES), pH 7.5, 30 mM KCl, 1 mM

Figure 3.2 Preparation of DOB-containing substrates via photolysis (modified from 251)
DOB-containing substrates were prepared by exposure of o-nitrobenzyl protected
dithiothreitol (DTT), 1 mM EDTA, 1 mM phenylmethylsulfonyl fluoride (PMSF), and 0.5% inositol. The cell lysates were centrifuged at 12,000 rpm for 30 minutes at 4 °C. The supernatant was loaded onto a 10-ml Sepharose Q column operated by an AKTA Fast Protein Liquid Chromatography system (FPLC) (GE Healthcare, Piscataway, NJ, USA). The flow-through was collected and dialyzed into buffer containing 30 mM HEPES, pH 7.5, 30 mM KCl, 0.5% inositol, and 1 mM PMSF, and subsequently loaded onto a 5-ml CM sepharose column (Bio-Rad Laboratories, Hercules, CA, USA), with fractions eluted using a linear gradient of KCl from 30 mM to 2 M. Peak fractions were combined and dialyzed into buffer containing 30 mM HEPES, pH 7.5, 0.5% inositol, 1.7 M (NH₄)₂SO₄, and 1 mM PMSF. Samples were then loaded onto a 2-ml phenyl sepharose column, with fractions eluted using a linear gradient of (NH₄)₂SO₄ from 1.7 M to 0 M. The peak fractions were combined and dialyzed into buffer containing 30 mM HEPES, pH 7.5, 30 mM KCl, 0.5% inositol, and 1 mM PMSF. Samples were then loaded onto a 1-ml Mono-S column (GE Healthcare, Piscataway, NJ, USA), and eluted using a linear gradient of KCl (30 mM to 2 M). Purified FEN1 was then aliquoted and frozen at -80 °C until further use.

LIG I was expressed in *E. Coli* BL21(AI). Six liters of LB medium were inoculated with one colony each of the transformed BL21(AI) cells and were incubated overnight without shaking. The cells were then incubated at 37°C at 225 rpm until OD₅₉₅ reached 0.6. LIG I expression was induced by 1 mM IPTG for 24 hours at 15-18 °C. Cells were then harvested by centrifugation at 2500 rpm for 30 minutes at 4 °C. The supernatant was discarded and cell pellets were lysed in lysis buffer, which contained 50 mM Tris-HCl, pH 7.5, 50 mM NaCl, 1 mM EDTA, 1 mM DTT, 0.1% Nonidet P-40, 1 mM PMSF, and 1 cOmplete tablet of protease inhibitors (Roche, Indianapolis, IN) and subjected to French
Press at 150,000 PSI. The cell lysates were centrifuged at 12,000 rpm for 30 minutes at 4°C. The supernatant was then loaded onto a 20-ml P11 Phosphocellulose column, with fractions eluted using a linear gradient of NaCl (50 mM to 600 mM). The peak fractions were combined and dialyzed into buffer containing 30 mM HEPES, pH 7.0, 30 mM KCl, 0.1% inositol, and 1mM PMSF. Samples were loaded onto a 10-ml Q Sepharose column and eluted using a linear gradient of KCl (30 mM to 2 M). Peak fractions were combined and dialyzed into buffer containing 50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 7 mM 2-mercaptoethanol, 10 mM imidazole, 0.5% inositol, and 1 mM PMSF. Samples were then loaded onto a 4-ml Nickel-Nitrilotriacetic acid (Ni-NTA) column, with fractions eluted using a linear gradient of imidazole (10 mM to 600 mM). Purified LIG I was aliquoted and frozen at -80°C until further use.

**Reconstituted BER assay**

*In vitro* BER of a (a) native sugar produced by removal of uracil, (b) the abasic site analog tetrahydrofuran (THF), and (c) the oxidized DOB lesion, at the 5’-end of the downstream primer of the substrates was carried out by incubating 50 nM substrate with the indicated concentrations of FEN1, LIG I, and pol β. The native sugar was produced by pre-incubating the uracil-containing substrate for 30 minute with Uracil DNA Glycosylase (UDG). The DOB-containing substrate was subjected to photolysis at 365 nm UV light for 20 minutes to remove the protecting group and form the DOB lesion (Figure 3.2), immediately prior to carrying out the reaction. All reactions were carried out in reaction buffer containing 30 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), pH 7.8, 50 mM KCl, 0.5% inositol, and 0.1 mg/ml bovine serum albumin (BSA), with 5
mM MgCl₂, 2 mM ATP, and 50 µM dNTPs. The 20 µl BER reaction mixtures were incubated at 37 °C for 15 minutes and reactions were terminated by the addition of 20 µl stopping buffer containing 95% formamide and 10 mM EDTA. Reaction mixtures were then denatured at 95 °C for 10 minutes and separated by 15% urea-denaturing polyacrylamide gel electrophoresis. Substrates and products were detected and analyzed using a Pharos FX Plus PhosphorImager from Bio-Rad Laboratories (Hercules, CA, USA). All reactions were done in triplicate.

**BER enzymatic activity assays**

The activities of FEN1 flap cleavage and pol β DNA synthesis on all substrates were measured by incubating 50 nM substrates with the indicated concentrations of FEN1 and pol β in reaction buffer containing 30 mM HEPES, pH 7.8, 50 mM KCl, 0.5% inositol, and 0.1 mg/ml bovine serum albumin (BSA), with 5 mM MgCl₂ and 50 µM dNTPs. The native sugar was produced by pre-incubating the uracil-containing substrate for 30 minutes with UDG. The DOB-containing substrate was subjected to photolysis at 365 nm UV light for 20 minutes to remove the protecting group and form the DOB lesion, immediately prior to carrying out the reaction. For the pol β DNA synthesis assay, the 20 µl BER reaction mixtures were incubated at 37 °C and aliquots were removed at time points ranging from 2-30 minutes. Reactions were terminated by the addition of 20 µl stopping buffer containing 95% formamide and 10 mM EDTA. For the FEN1 5’-flap cleavage activity assay, samples were pre-incubated with 2.5 nM pol β for 1-5 minutes to allow crosslink formation prior to incubation with FEN1 for 15 minutes at 37 °C. Reactions were quenched
with 200 nM NaBH₄ on ice for 30 minutes. Reaction mixtures were then denatured at 95 °C for 10 minutes and separated by 15% urea-denaturing polyacrylamide gel electrophoresis. Substrates and products were detected and analyzed using a Pharos FX Plus PhosphorImager from Bio-Rad Laboratories (Hercules, CA, USA).

RESULTS

An oxidized abasic lesion inhibits pol β synthesis activity during BER

We first explored the effect of sugar modifications in an abasic lesion on pol β synthesis activity during BER in a TNR tract. During repair of an oxidized DOB lesion, the dRP lyase domain of pol β has been found to form a covalent crosslink with the DOB lesion (38, 247, 249, 250). Whether this crosslink can still form when the DOB lesion occurs in the loop of a hairpin resulting in lesion formation at the end of a 5’-flap has yet to be explored. To determine whether modifications to the sugar residue in an abasic lesion affect pol β synthesis activity during BER in a TNR hairpin and TNR duplex, and also whether the pol β-DOB crosslink would form with the DOB at the end of a 5’-flap, we tested pol β synthesis activity on substrates that mimic the intermediates of TNR hairpin repair and repair in a TNR duplex, that contained a native sugar, reduced sugar (THF), or oxidized sugar (DOB) at the 5’-end of the downstream flap. We used a (CAG)₃/(CAG)₂ double-flap substrate (Figure 3.3) to mimic the intermediate formed from APE1 cleavage of a (CAG)₅ hairpin. We used a substrate containing a nick with a downstream (CAG)₂-flap (Figure 3.4) to mimic the intermediate formed during repair of a TNR duplex. We
tested pol β (2.5 nM) on both the double-flap and single-flap substrates during a time course ranging from 2-30 minutes. We found that during repair of the double-flap substrate, pol β synthesis was decreased on the DOB-containing substrate compared to the native sugar- or THF-containing substrates (Figure 3.3, compare lanes 13-18 with lanes 1-6 and lanes 7-12). This indicates that the pol β-DOB crosslink stills forms when the DOB lesion is located at the free end of a 5’-flap, resulting in a decrease in subsequent pol β synthesis activity. No difference in activity was observed between the native sugar- and THF-containing substrates (Figure 3.3, compare lanes 1-6 and lanes 7-12), indicating that these lesions have no effect on pol β synthesis activity.
Figure 3.3 The oxidized abasic lesion DOB inhibits pol β DNA synthesis activity during BER in a small TNR hairpin loop

The effect of the type of abasic lesion on pol β synthesis activity on a short (CAG)$_3$/(CAG)$_2$ double-flap intermediate formed during BER of a (CAG)$_3$ hairpin was examined by measuring pol β synthesis activity on the substrate containing a native sugar (left panel), THF (middle panel), or DOB lesion (right panel). Lanes 1, 7, and 13 indicate the substrate only. Lanes 2-6, lanes 8-12, and lanes 14-18 indicate the substrate incubated with pol β (2.5 nM) for a time course of 2, 5, 10, 15, and 30 minutes. Substrates were $^{32}$P-labeled at the 5'-end of the upstream strand and are illustrated above each gel.
For the nicked (CAG)$_2$ 5’-flap substrate that mimics repair in a CAG repeat duplex, pol β synthesis activity was also found to be greatly decreased in the substrate containing the DOB lesion compared to the native sugar- and THF-containing substrates (Figure 3.4, compare lanes 13-18 with lanes 1-6 and lanes 7-12), consistent with results from the double-flap substrate. This indicates that the crosslink still forms between the DOB and dRP lyase domain of pol β, as evidenced by the decreased pol β activity. No difference in pol β synthesis activity was observed between the native sugar- and THF-containing substrates (Figure 3.4, compare lanes 1-6 with lanes 7-12), indicating that these lesions have no effect on pol β synthesis activity. Overall, these results indicate that an oxidized sugar residue in an abasic site can inhibit pol β synthesis activity. Specifically, the formation of the crosslink between pol β and the oxidized DOB lesion at the end of the 5’-flap results in decreased pol β activity as a result of a decrease in the amount of free pol β available for this reaction.
Figure 3.4 The oxidized abasic lesion DOB inhibits pol β DNA synthesis activity during BER in a TNR duplex

The effect of the type of abasic lesion on pol β synthesis activity on a substrate containing a single (CAG)$_2$ downstream flap intermediate formed during BER of a (CAG)$_n$ duplex was examined by measuring pol β synthesis activity on the substrate containing a native sugar (left panel), THF (middle panel), or DOB lesion (right panel). Lanes 1, 7, and 13 indicate the substrate only. Lanes 2-6, lanes 8-12, and lanes 14-18 indicate the substrate incubated with pol β (2.5 nM) for a time course of 2, 5, 10, 15, and 30 minutes. Substrates were $^{32}$P-labeled at the 5'-end of the upstream strand and are illustrated above each gel.
An oxidized abasic lesion inhibits FEN1 cleavage activity during BER

After confirming that the pol β-DOB crosslink still forms in the double-flap intermediate of BER in a TNR hairpin and in the intermediate of BER in a TNR duplex, we sought to determine whether modifications of the sugar residue in an abasic site affect FEN 5’-flap cleavage, and more specifically, whether the Pol β-DOB crosslink at the end of the 5’-flap inhibits FEN1 loading and subsequent cleavage activity. We therefore tested FEN1 cleavage activity on the (CAG)₃/(CAG)₂ double-flap substrate (Figure 3.5) which mimics the intermediate formed from APE1 cleavage of a (CAG)₅ hairpin and also on the substrate containing a nick with a downstream (CAG)₂-flap (Figure 3.6) to mimic the intermediate formed during BER in a TNR duplex. This was accomplished by pre-incubating the substrates with pol β (2.5 nM) for 1 minute or 5 minutes to allow the pol β-DOB crosslink to form. The substrates were then incubated with FEN1 (1 nM) to measure 5’-flap cleavage activity. For the (CAG)₃/(CAG)₂ double-flap substrate, we found that FEN1 was able to carry out cleavage on all 3 lesion-containing substrates (Figure 3.5). However, for the DOB-containing substrate, FEN1 cleaved significantly shorter flaps as compared to the native sugar- and THF-containing substrates (Figure 3.5, compare lanes 9-12 with lanes 1-4 and lanes 5-8), indicating a reduced efficiency of FEN1 cleavage in the presence of the pol β-DOB crosslink. This likely indicates that the reduced concentration of free pol β available to carry out strand displacement synthesis prevents formation of a longer 5’-flap for FEN1 cleavage. Surprisingly, we also found that for the native sugar-containing substrate, FEN1 cleaved fewer nucleotides following 5 min pre-incubation with pol β than in the THF-containing substrate (Figure 3.5, compare lane 4 with lane 8). This may indicate that the transient crosslink formed between the native sugar and the Lys72
residue in the dRP lyase domain of pol β during the dRP lyase reaction (248) acts to slightly inhibit FEN1 cleavage activity.

For the substrate containing a nick with a downstream (CAG)$_2$-flap to mimic the intermediate formed during BER in a TNR duplex, we again found slight inhibition of FEN1 cleavage activity, as fewer nucleotides were removed by FEN1 than in the presence of the THF lesion (Figure 3.6, compare lanes 11-12 with lanes 3-4 and lanes 7-8). The number of nucleotides cleaved by FEN1 in the presence of the native sugar was also slightly reduced compared to cleavage in the presence of a THF residue (Figure 3.6, compare lane 4 to lane 8). This again may indicate that the transient native sugar- pol β intermediate acts to slightly inhibit FEN1 cleavage activity. Overall, these results indicate that an oxidized sugar residue in an abasic lesion can affect FEN1 cleavage. As FEN1 was able to cleave flaps containing all lesions with similar efficiency in the absence of pol β (Figures 3.5 and 3.6, compare lanes 2, 6, and 10), this likely results from inhibition of pol β strand displacement synthesis that prevents formation of a longer 5’-flap for FEN1 cleavage. The DNA-protein crosslink formation in the presence of the oxidized lesion thus acts to interrupt the activities of both pol β and FEN1 which are crucial for efficient BER.
Figure 3.5 An oxidized abasic lesion inhibits FEN1 5'-flap cleavage activity during BER in a small TNR hairpin loop

The effect of the pol β-DOB crosslink on FEN1 cleavage activity on a short (CAG)$_3$/(CAG)$_2$ double-flap intermediate formed during BER of a (CAG)$_5$ hairpin was explored by measuring FEN1 cleavage activity on the substrate containing a native sugar (left panel), THF (middle panel), or DOB lesion (right panel). Lanes 1, 5, and 9 indicate the substrate only. Lanes 2, 6, and 10 indicate substrate incubated with FEN1 (1 nM) only. Lanes 3, 7, and 8 indicate the substrate pre-incubated with pol β (5 nM) for 1 min prior to incubation with FEN1 (1 nM). Lanes 4, 8, and 12 indicate substrates pre-incubated with pol β (5 nM) for 5 min prior to incubation with FEN1 (1 nM).
Figure 3.6 An oxidized abasic lesion slightly inhibits FEN1 5'--flap cleavage activity during BER in a TNR duplex

The effect of the pol β-DOB crosslink on FEN1 cleavage activity on a substrate containing a (CAG)$_2$ downstream flap formed during BER in a TNR duplex was explored on the substrate containing a native sugar (left panel), THF (middle panel), or DOB lesion (right panel). Lanes 1, 5, and 9 indicate the substrate only. Lanes 2, 6, and 10 indicate substrate incubated with FEN1 only. Lanes 3, 7, and 9 indicate the substrate pre-incubated with pol β for 1 min prior to incubation with FEN1. Lanes 4, 8, and 12 indicate substrates pre-incubated with pol β for 5 min prior to incubation with FEN1.
An oxidized abasic lesion inhibits efficient base excision repair in a TNR tract

To determine whether the reduced pol β and FEN1 activities resulting from an oxidized sugar would affect the efficiency of BER in a TNR tract, we reconstituted BER using the (CAG)$_3$/(CAG)$_2$ double-flap substrate (Figure 3.7) which mimics the intermediate formed from APE1 cleavage of a (CAG)$_5$ hairpin and also on the substrate containing a nick with a downstream (CAG)$_2$-flap (Figure 3.8) to mimic the intermediate formed during BER in a TNR duplex, containing the native sugar, THF, or DOB lesions at the 5’-end of the downstream flap. We found that during BER of a DOB lesion in the (CAG)$_3$/(CAG)$_2$ double-flap substrate (Figure 3.7), the amount of repaired product was greatly reduced compared to BER of a native sugar or THF lesion (Figure 3.7, compare lanes 13-14 with lanes 3-4 and lanes 8-9, percentage of repaired product is shown below each gel). This indicates that the oxidized lesion, which inhibits both pol β synthesis activity and FEN1 cleavage activity, inhibits the formation of repaired product by preventing the formation of a ligatable nick suitable for ligation by LIG I. This occurs as the protein-DNA crosslink at the 5’end of the downstream region is unsuitable for ligation, and also as FEN1 cleavage of a downstream flap leaves a gap that is not filled by the DNA synthesis activity of the limited remaining pol β. This prevents the formation of repaired product and results instead in an accumulation of repair intermediates containing a DNA break.
Figure 3.7 An oxidized abasic lesion decreases repair efficiency during BER in a small TNR hairpin loop

The effect of abasic lesion type on BER efficiency during BER in a small TNR hairpin loop was examined by reconstituting BER with the short (CAG)$_3$/(CAG)$_2$ double-flap intermediate formed during BER in a (CAG)$_5$ hairpin containing a native sugar (left panel), THF (middle panel), or DOB lesion (right panel). Lanes 5, 10, and 15 indicate synthesized size markers as indicated to the left of the gel. Lanes 1, 6, and 11 indicate the substrate only. Lanes 2, 7, and 12 indicate the substrate incubated with FEN1 (1 nM) only. Lanes 3, 8, and 13 indicate reaction mixtures incubated with FEN1 (1 nM) and LIG I (2.5 nM). Lanes 4, 9, and 14 indicate reaction mixtures incubated with FEN1 (1 nM), LIG I (2.5 nM), and pol $\beta$ (1 nM). Graphs indicating the percentage of repaired products are shown below each gel. Each experiment was done in triplicate. Substrates were $^{32}$P-labeled at the 5’-end of the upstream strand and are illustrated above each gel.
For the substrate containing a nick with a downstream (CAG)₂-flap (Figure 3.8) which mimics the intermediate formed during BER in a TNR duplex, we also found that oxidation of the sugar resulted in a significant decrease in the formation of repaired product (Figure 3.8, compare lanes 13-14 with lanes 3-4 and lanes 8-9, percentage of repaired product is shown below each gel), which indicates that oxidation of the sugar results in decreased BER efficiency and the accumulation of BER intermediates containing DNA breaks. These results indicated that oxidation of the sugar residue contributes to genomic instability by facilitating formation of DNA breaks via BER and preventing completion of repair, resulting in accumulation of DNA breaks.
Figure 3.8 An oxidized abasic lesion decreases repair efficiency during BER in a TNR duplex

The effect of abasic lesion type on BER efficiency during BER in a TNR duplex was examined by reconstituting BER with on a substrate containing a single (CAG)$_2$ downstream flap intermediate formed during BER of TNR duplex containing a native sugar (left panel), THF (middle panel), or DOB lesion (right panel). Lanes 5, 10, and 15 indicate synthesized size markers as indicated to the left of the gel. Lanes 1, 6, and 11 indicate the substrate only. Lanes 2, 7, and 12 indicate the substrate incubated with FEN1 (2.5 nM) only. Lanes 3, 8, and 13 indicate reaction mixtures incubated with FEN1 (2.5 nM) and LIG I (2.5 nM). Lanes 4, 9, and 14 indicate reaction mixtures incubated with FEN1 (2.5 nM), LIG I (2.5 nM), and pol β (1 nM). Graphs indicating the percentage of repaired products are shown below each gel. Each experiment was done in triplicate. Substrates were $^{32}$P-labeled at the 5’-end of the upstream strand and are illustrated above each gel.
The decrease in repaired product observed during BER of an oxidized abasic site in the absence of pol β (Figure 3.7, compare lane 13 with lanes 3 and 8 and Figure 3.8, compare lane 13 with lanes 3 and 8, percentage of repaired product is shown below each gel) further indicates that the activity of LIG I is also affected by the oxidized sugar, as FEN1 was shown to be able to remove the 5’-DOB- containing flap as efficiently as a 5’-native sugar-containing flap and 5’-THF- containing flap in the absence of pol β (Figure 3.5, compare lane 10 with lanes 2 and 6, and Figure 3.6, compare lane 10 with lanes 2 and 6).

**DISCUSSION**

In this study, we explored the effects of modifications of the sugar residue in an abasic site on the activity of BER proteins and BER efficiency in a TNR tract. We found that an oxidized lesion significantly inhibited pol β synthesis activity (Figures 3.3 and 3.4) and slightly inhibited FEN1 5’-flap cleavage activity (Figures 3.5 and 3.6). This resulted in a decrease in BER efficiency in the presence of the oxidized sugar (Figures 3.7 and 3.8). Previous studies have shown that pol β only achieves insignificant amounts of DNA strand displacement synthesis when acting on a DOB-containing DNA duplex of random sequence, due to efficient pol β inactivation by crosslink formation with the lesion (247). The C4-oxidized abasic site (C4-AP) has also been found to form crosslinks with the lyase domain of pol β, but does so less efficiently than DOB (247). C4-AP inhibits pol β less efficiently (~7 turnovers compared to ~4 turnovers for DOB) due to the reduced electrophilicity of C4-AP compared to DOB, and as a result C4-AP can be efficiently repaired by BER because there is sufficient strand displacement synthesis to remove the lesion in conjunction with FEN1 cleavage (247, 250). This supports a model (Figure 3.9)
in which oxidation of a sugar in a TNR tract results in the formation of a DOB lesion in conjunction with a single-strand DNA break. This creates a repair intermediate which contains the DOB at the 5’-end of the downstream flap. The DOB forms a covalent crosslink with the lyase domain of pol β, inactivating the enzyme. FEN1 is able to load onto the flap and remove the cross-linked DNA, though with reduced efficiency, however, the remaining pol β is insufficient to fill in the resulting gap. This prevents the formation of a ligatable nick that can be sealed by LIG I and thus results in an accumulation of DNA breaks, leading to cytotoxicity and genomic instability. Furthermore, our results indicate a decreased efficiency of BER even in the absence of pol β. Because FEN1 was shown to be able to remove the lesions with equal efficiency in the absence of the protein-DNA crosslink (Figures 3.5 and 3.6), this indicates that the activity of LIG I is also inhibited by the oxidized lesion. Similar to the active site of the pol β dRP lyase domain, LIG I also contains a key Lys residue in the active site (Lys568) (252, 253), which presents the possibility that LIG I could also be modified through reaction with the DOB, removing active LIG I from solution and preventing efficient ligation of repaired product (Figure 3.9). This hypothesis remains to be explored. As pathways of DNA break repair, such as homologous recombination, have also been linked to TNR instability, this could further contribute to the instability of TNR tracts, genomic instability, and cytotoxicity.
An oxidized abasic lesion inhibits efficient BER in a TNR tract

Exposure to ionizing radiation or antitumor drugs can result in the formation of an oxidized sugar in a TNR tract, which occurs concomitantly with a single-strand DNA break. The attempt by pol $\beta$ to remove the abasic lesion through its dRP lyase activity results in formation of a pol $\beta$-DOB crosslink at the end of the 5’-flap. FEN1 is able to load onto the 5’-flap and cleave, although with reduced efficiency, resulting in removal of the cross-linked protein. However, the limited amount of remaining pol $\beta$ results in inefficient gap-filling synthesis and is unable to create a ligatable nick, thus resulting in the accumulation of DNA breaks. We propose that LIG I may also become crosslinked to the oxidized lesion, removing sufficient LIG I from solution and preventing LIG I from efficiently sealing the nick, resulting in accumulation of DNA breaks.

Figure 3.9 An oxidized abasic lesion inhibits efficient BER in a TNR tract

Exposure to ionizing radiation or antitumor drugs can result in the formation of an oxidized sugar in a TNR tract, which occurs concomitantly with a single-strand DNA break. The attempt by pol $\beta$ to remove the abasic lesion through its dRP lyase activity results in formation of a pol $\beta$-DOB crosslink at the end of the 5’-flap. FEN1 is able to load onto the 5’-flap and cleave, although with reduced efficiency, resulting in removal of the cross-linked protein. However, the limited amount of remaining pol $\beta$ results in inefficient gap-filling synthesis and is unable to create a ligatable nick, thus resulting in the accumulation of DNA breaks. We propose that LIG I may also become crosslinked to the oxidized lesion, removing sufficient LIG I from solution and preventing LIG I from efficiently sealing the nick, resulting in accumulation of DNA breaks.
SUMMARY

Base excision repair (BER) is a critical pathway in the maintenance of genomic stability and is the major pathway of repair of small base lesions such as oxidized and alkylated lesions, deaminated bases and ssDNA breaks. Oxidative and alkylated DNA damage have been found to play a role in facilitating TNR instability that is associated with human neurodegenerative diseases and cancer, indicating a critical role of BER in mediating TNR instability. Instability has been linked to the formation of DNA secondary structures, such as hairpins, that interrupt the coordination of activities between key BER proteins, preventing efficient repair and resulting in instability in the length of the TNR tracts.

Expansions in TNR tracts are associated with over 40 human neurodegenerative diseases, including Huntington’s disease and myotonic dystrophy, while deletions are associated with certain types of cancers such as ovarian and prostate cancer. While many studies have confirmed the role of BER in mediating TNR instability, the mechanisms TNR expansion and contraction, as well as mechanisms used by cells to prevent this instability, are not well understood. Furthermore, while research has indicated that the location of DNA damage in a TNR tract governs whether expansions or deletions occur, how the type of damage might affect the repair process has not previously been explored. Identification of cellular pathways for preventing TNR expansions could help identify novel targets for the therapeutic treatment of this family of TNR instability-associated diseases, for which there are currently no cures and only limited treatment options. To identify cellular pathways of preventing TNR expansions and the role of BER in modulating TNR
instability, this Ph.D. dissertation research project focuses on identifying the interactions of BER proteins and cofactors in the removal of extra CAG repeats by exploring how different types of DNA damage and DNA structures affect the activities of BER proteins and the outcome of BER in a TNR tract. We explored the role of DNA damage and BER in modulating TNR instability using synthesized DNA oligonucleotides containing DNA damage in a TNR duplex tract or TNR hairpin loop and reconstituted BER using purified human repair enzymes and cofactors.

Initially, we explored how the key BER enzymes could cooperate to promote removal of a TNR hairpin and prevention of TNR expansions. Hairpins have been experimentally shown to form both in vitro and in vivo in TNR tracts, and guanines located in the loop of a hairpin have been found to be particularly susceptible to the formation of oxidized lesions such as 8-oxoG, indicating that hairpins with oxidative damage in the loop are likely to accumulate in TNR tracts. We found that interactions between the key BER enzymes, APE1, FEN1, pol β, and LIG I could resolve the intermediates formed during BER of an oxidized lesion in the loop of a TNR hairpin, leading to hairpin removal and prevention or attenuation of expansions. During BER in a TNR hairpin loop, APE1 incision at the abasic site converts the hairpin into a double-flap intermediate with an upstream 3’-flap and a downstream 5’-flap. This intermediate was found to be resolved by the coordinated actions of the 3’-5’ exonuclease activity of APE1 and the 5’-flap cleavage activity of FEN1, which removed the upstream and downstream flaps, respectively. This study identified a critical role for the 3’-5’ exonuclease activity of APE1, which had previously remained elusive. Furthermore, APE1 was found to stimulate the activity of LIG I to specifically facilitate completion of hairpin removal and prevention of expansions.
Thus, while BER in a TNR tract can lead to TNR expansions in the context of a TNR duplex, this project provides further evidence that damage location plays a role in governing TNR instability through BER. We showed that BER in the loop of a TNR hairpin leads to hairpin removal and prevention of TNR expansions through the coordinated action of the key BER enzymes.

While this research showed that the key BER enzymes can facilitate TNR hairpin removal to prevent TNR expansions, removal of a small hairpin was found to be much more efficient, with BER in the loop of a larger hairpin leading to some expanded repaired products that were shorter than the original hairpin-containing DNA, but still longer than the template strand as a result of partial hairpin removal. In the second part of this research project, we explored the role of the BER cofactor PCNA in facilitating more efficient hairpin removal to prevent TNR expansions. PCNA is known to interact with the key BER proteins during repair in non-repeat duplex DNA, and in doing so to facilitate more efficient LP-BER. In this study we explored the role of PCNA in facilitating efficient BER during repair in a TNR duplex and in the loop of a TNR hairpin. We found that PCNA plays a critical role in promoting repeat deletion and removal of a TNR hairpin during repair. During BER in a TNR duplex, we found that PCNA was blocked from interaction with FEN1 by small bubbles formed in the template strand that prevented PCNA from sliding downstream along the duplex to interact with FEN1 at the flap base. However, we found that following FEN1 removal of a 5’-lesion and adjacent nucleotide, PCNA stimulated the activity of LIG I across a small template loop to promote removal of CAG repeats resulting in small TNR deletions. During BER in the loop of a TNR hairpin, we found that PCNA facilitated hairpin removal and prevention of TNR expansions. We
discovered that PCNA facilitated hairpin removal in several ways. By loading upstream of the damage and sliding along the DNA toward the downstream, PCNA was found to facilitate the annealing of the upstream flap of the double-flap intermediate of hairpin repair, forming a longer downstream 5’-flap. This created more 3’-termini as a substrate for pol β synthesis, resulting in an increase in pol β synthesis activity, and also allowed PCNA to interact with FEN1 at the base of the downstream 5’-flap to stimulate FEN1 activity for more efficient removal of the TNR-flap, preventing the extra repeats from being incorporated as an expansion. Therefore, PCNA was found to both promote CAG repeat removal during repair in a TNR duplex and facilitate hairpin removal during repair in the hairpin loop, indicating a role for PCNA in preventing TNR expansions. This study also indicates that the structure of DNA in a TNR tract plays a role in governing whether proper coordination between repair enzymes and cofactors is maintained to modulate TNR instability.

The results from the first two parts of this dissertation project indicate that interactions between BER enzymes and cofactors can modulate TNR expansions by facilitating deletion of repeats and hairpin removal. These interactions were found to be dependent on the structure of the DNA, supporting previous studies finding that lesion location plays a role in governing TNR instability as a result of different DNA structures formed within the repeats. While lesion location is indicated in governing TNR instability, the effect of lesion type on TNR instability has not previously been explored. No research has previously determined how modifications of the sugar residue in an abasic lesion might affect this process. In the third part of this research project, we explored the effect of sugar modification in an abasic site on the activities of BER enzymes and BER efficiency during
repair in a TNR duplex and the loop of a TNR hairpin. Using intermediates formed during BER in a CAG repeat duplex and a CAG hairpin loop containing a native, reduced, or oxidized sugar at the 5’-end of a downstream flap, we explored the effect of lesion type on the activities of the BER enzymes and found that an oxidized lesion significantly inhibited pol β synthesis activity and slightly inhibited FEN1 flap cleavage activity, as compared to the native and reduced sugars. The inhibition of pol β synthesis activity resulted from crosslink formation between the oxidized abasic site, DOB, and Lys72 in the active site of the dRP lyase domain of pol β. This indicated that the DNA-protein crosslink was still able to form when the lesion was located at the end of a 5’-flap. This decrease is pol β strand displacement synthesis then resulted in reduced FEN1 cleavage as a longer 5’-flap for FEN1 cleavage failed to form. We then analyzed the effect of modified enzymatic activities on the efficiency of BER in both substrates and found that the reduced FEN1 and pol β activities resulted in decreased BER efficiency on both substrates, resulting in accumulation of BER intermediates containing ssDNA breaks. The accumulation of DNA breaks would lead to genomic instability, and as other mechanisms of break repair such as HR are also linked to TNR instability, BER of the oxidized lesion may also contribute to TNR instability. We further found that even in the absence of the DNA-pol β crosslink there was a decrease in the efficiency of BER, indicating that the LIG I is also likely inhibited by the oxidized lesion. Thus, in this project, we found that modifications of the sugar in an abasic site do play a role in mediating BER efficiency and TNR instability, as we found that an oxidized sugar inhibits efficient coordination of BER protein activities and BER efficiency, leading to an accumulation of DNA damage and DNA breaks.
By exploring the role of DNA damage and BER in modulating TNR instability, we have found that both DNA structure and the type of DNA damage play a role in governing TNR instability. This research has advanced our understanding of molecular mechanisms and cellular pathways of preventing TNR instability and the role of BER in modulating TNR expansions. We have identified molecular mechanisms and cellular pathways of removing TNR repeats and TNR hairpins that can be used as a basis for identifying novel targets for the therapeutic prevention and treatment of TNR instability-associated diseases.

These studies have defined the role of DNA base lesions and BER in modulating TNR expansions through interactions of BER proteins and cofactors. However, recent studies indicate that there is likely crosstalk between repair pathways, including BER, MMR, and NER. Other repair proteins such as the MMR complex MSH2/MSH3 have been found to stabilize TNR secondary structures, and key BER proteins such as glycosylases and pol β have been found to interact with MMR machinery both physically and functionally. Future studies exploring how repair crosstalk pathways are involved in TNR expansions will help identify novel mechanisms of modulating TNR instability and models for targeting treatment of TNR instability diseases. Furthermore, so far our research has utilized synthesized DNA oligonucleotides that are restricted both in length and in structure, while in vivo, DNA is organized into chromatin and contains additional modifications such methylation that may play a role in mediating TNR instability. Thus, future studies exploring these mechanisms in vivo will be needed to create a more comprehensive model for the role of BER in modulating TNR instability.
REFERENCES


VITA

JILL BEAVER

2002-2006  NC Teaching Fellow
            University of North Carolina
            Chapel Hill, NC

2006      B.S., Biology
            University of North Carolina
            Chapel Hill, NC

2007      M.A., Teaching
            University of North Carolina
            Chapel Hill, NC

2007-2012 Wake County Public School System
            Raleigh, NC

2012-2013 Chemistry Teaching Assistant of the Year Award
            Florida International University
            Miami, FL

2014      Biomedical Research Initiative Summer Research Award
            Florida International University
            Miami, FL

2014-2016 Doctoral Candidate and FIU RISE Research Fellow
            Florida International University
            Miami, FL

PUBLICATIONS AND PRESENTATIONS


Beaver, J.M. and Liu, Y. Prevention of trinucleotide repeat expansion by DNA bases excision repair cofactors. DNA Repair and Genome Stability Symposium, December 16, 2015, Miami, FL.


Beaver, J.M. and Liu, Y. Proliferating Cell Nuclear Antigen prevents trinucleotide repeat expansion. 18th Annual Biomedical and Comparative Immunology Symposium, March 3-4, 2016, Miami, FL.


