2011

IDENTIFICATION OF ACTIVITIES INVOLVED IN CAG/CTG REPEAT INSTABILITY

Nelson Lap Shun Chan
University of Kentucky, nelsonlschan@gmail.com

IDENTIFICATION OF ACTIVITIES INVOLVED
IN CAG/CTG REPEAT INSTABILITY

_____

DISSERTATION

_____

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Medicine
at the University of Kentucky

By

Nelson Lap Shun Chan

Lexington, Kentucky

Director: Dr. Guo-Min Li, Professor of Toxicology

Lexington, Kentucky

2011

ABSTRACT OF DISSERTATION

IDENTIFICATION OF ACTIVITIES INVOLVED
IN CAG/CTG REPEAT INSTABILITY

CAG/CTG repeat instability is associated with at least 14 neurological disorders, including Huntington's disease and Myotonic dystrophy type 1. *In vitro* and *in vivo* studies have showed that CAG/CTG repeats form a stable hairpin that is believed to be the intermediate for repeat expansion and contraction.

Addition of extra DNA is essential for repeat expansion, so DNA synthesis is one of the keys for repeat expansion. *In vivo* studies reveal that 3' CTG slippage with subsequent hairpin formation (henceforth called the 3' CTG slippage hairpin) occurs during DNA synthesis. It is proposed that hairpin tolerance machinery is activated because prolonged stalling of DNA polymerase triggers severe DNA damage. As a means toward studying the hairpin-mediated expansion, we created a special hairpin substrate, mimicking the 3' CTG slippage hairpin, to determine which polymerase promotes hairpin bypass. Our studies reveal polymerase β (pol β) is involved in the initial hairpin synthesis while polymerase δ (pol δ) is responsible for the resumption of DNA synthesis beyond the hairpin (extension step). Surprisingly, we also found that the pol δ can remove the short CTG hairpin by excision of the hairpin with its 3' to 5' exonuclease activity.

Besides repairing the hairpin directly, resolving the hairpin is an alternative pathway to maintain CAG/CTG repeat stability. With limited understanding of which human helicase is responsible for resolving CAG/CTG hairpins, we conducted a screening approach to identify the human helicase involved. Werner Syndrome Protein (WRN) induces the hairpin repair activity when $(CTG)_{35}$ hairpin is formed on the template strand. Primer extension assay reveals that WRN stimulates pol δ synthesis on $(CAG)_{35}/(CTG)_{35}$ template and such induction was still found in the presence of accessory factors. Helicase assay confirms that WRN unwinds CTG hairpin structures.

Our studies provide a better understanding of how polymerases and helicases play a role in CAG/CTG repeat instability. Considering CAG/CTG repeat instability associated disorders are still incurable, our studies can provide several potential therapeutic targets for treating and/or preventing CAG/CTG repeat associated disorders.

KEYWORDS: CTG, Repeat Instability, Hairpin, Helicase, Polymerase.

Nelson Lap Shun Chan_____
Student signature

09/08/2011_____
Date

IDENTIFICATION OF ACTIVITIES INVOLVED
IN CAG/CTG REPEAT INSTABILITY

By

Nelson Lap Shun Chan

Guo-Min Li, PhD
 Director of Dissertation

Liya Gu, PhD
 Director of Graduate Studies

09/08/2011
 Date

This dissertation is dedicated to God

ACKNOWLEDGEMENT

First of all, I would like to express my gratitude to my advisor Dr. Guo-Min Li for his patience, support and encouragement during my PhD studies. I was so privileged to be nurtured into becoming an independent researcher with Dr. Li's guidance. As a young researcher, I got discouraged easily when an experiment fell short of my expectations. However, Dr. Li always sees opportunities from the problems and finds the possible solutions. Hard work and perseverance are two important attitudes I learned from Dr. Li. I am sure these valuable attitudes will be useful to me as I face the many different challenges ahead of me.

Also, I want to thank Dr. Liya Gu for her support and advice on my research. Her expertise in large scale DNA and nuclear proteins extraction saved me a lot of time and troubleshooting. Special thanks to Dr. David Orren and Dr. Amrita Machwe for offering the Werner syndrome protein (K577M) mutant and Bloom's syndrome protein. Their advice on the WRN project also saved me a lot of time and troubleshooting.

My thanks also goes to my committee members, Dr. Howard Glauert, Dr. David Orren, Dr. Martha Peterson and Dr. Andrew Pierce for providing many valuable comments and stimulating ideas for my dissertation. Moreover, I would also like to thank my external examiner, Dr. Stephen Testa, for reading and giving suggestions about my dissertation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 DNA Metabolism and CAG/CTG Repeat Instability

1.1 Introduction

Microsatellites are stretches of DNA in which a sequence of repeating units, typically of 1 to 5 nucleotides, appears several times, such as the dinucleotide $(CA)_n$ or the trinucleotide $(CAG)_n$ (1). Microsatellite repeats are polymorphic, meaning different individuals may have a different number of repeats in each particular satellite in their genome. The number of repeats in a satellite is typically stable in an individual if the number of repeats is below a certain level. However, when the repeat length increases over a particular threshold, it can become unstable and possibly trigger clinical manifestations due to gene malfunction or toxic protein accumulation. It is still not clear why repeat threshold is different for different diseases, but it is hypothesized that cis-elements (DNA content, e.g repeat sequence, flanking sequence) (2)and trans-factors (protein expression) contribute to different threshold(3). More than 30 neurological diseases are associated with microsatellite instability (4), such as Huntington disease and Fragile X syndrome. Since the repeat will be transmitted and expanded when it is passed to the next generation, repeat instability diseases show what is termed anticipitation, meaning the severity of the disease increases with each subsequent generation. Repeat instability can happen in both coding and non-coding regions of DNA, but in coding regions trinucleotide repeats instability is more commonly found than other types of repeats. Since a single amino acid is specified by three nucleotides, di-, tetra- and pentanucleotide repeats in the coding region may result in different amino acid production, causing frameshift mutations (5). Much research has focused on trinucleotide

repeat instability because it contributes to many repeat-associated disorders (4). Different

trinucleotide repeat sequences have been found to form different types of secondary

structures. For example, Watson-Crick base pair mediated hairpins can be formed by

CAG and CTG repeats, Hoogsteen base pair mediated G-quadruplexes can be formed by

CGG repeats and triplexes can be formed by GAA repeats (6). These secondary

structures interfere with the normal DNA metabolism, so it is proposed that repeat

instability is caused by aberrant secondary structure formation in DNA (7).

1.2 Repeat instability in dividing cells

DNA slippage during replication was one of the first mechanisms proposed to

explain repeat instability (8). Secondary structures formed by trinucleotide repeats can

cause misalignment or DNA slippage during DNA synthesis. Repeat expansion or

contraction is resulted when the DNA slippage occurs on the nascent strand or template

strand respectively. Errors during DNA replication are proven to cause repeat instability

in bacteria (9) and yeast (10). *In vitro* studies using the SV40 DNA replication system

reveal that CAG repeat size and location of the SV40 replication origin relative to the

location of the repeat sequence can affect repeat stability in primate (11) and human cells

(12). Location of the repeat sequence relative to the replication origin was also found to

affect repeat instability in bacteria. Bacterial studies reveal that higher repeat instability is

observed when the repeat sequence is closer to the replication origin (13). It is also

observed that small alterations of the distance from the replication origin to the

CAG/CTG repeats may shift contraction to expansion (11). The location of the

replication origin determines the repetitive sequence and secondary structure of the initiation site for lagging strand synthesis. Therefore, this model proposes that affected individuals may have different locations of replication origin, leading to repeat instability (14).

Repeat instability may show strand specific bias during DNA replication. Since A-A base pairing has weaker base stacking than that of T-T, CTG repeats form a more stable structure than secondary structure formed from CAG repeats (15). Therefore, the secondary structure formed by CAG/CTG repeats are different on leading and lagging strands (16). Indeed, higher repeat instability frequency is observed when CTG is used as the template for lagging strand synthesis in bacteria and yeast (9). Also, addition of emetine (a lagging strand synthesis inhibitor) to human cells promotes CAG/CTG instability (17). During lagging strand synthesis, RAD27/FEN1 is responsible for removing the 5' flap. RAD27/FEN1 deletion results in CAG/CTG instability in yeast (18). *In vitro* studies reveal that RAD27/FEN1 fails to remove the 5' flap formed by long CAG repeats and the subsequent ligation of the CAG hairpin on the nascent strand promotes expansion (19). Therefore, secondary structure formation due to prolonged single stranded DNA exposure (20) and/or aberrant lagging strand maturation (21) are proposed to explain higher repeat instability during lagging strand synthesis.

1.2.1 DNA polymerases

DNA polymerase progression is impeded by the secondary structures formed by trinucleotide repeats (22,23). Since the prolonged exposure of single stranded DNA can trigger serious damage to the genome (24,25), it is proposed that cells may utilize a similar damage bypass system to restore DNA synthesis when they come across trinucleotide repeat secondary structures (26). In translesion DNA synthesis (TLS), a low fidelity polymerase is recruited, which inserts several bases past the lesion, after which the polymerase is replaced by the replicative polymerases to continue synthesis (26). Expansion occurs when the hairpin bypass occurs on the nascent strand (27). Since there are a number of TLS polymerases, it is still not clear which polymerase is responsible for the hairpin bypass. Yeast with deletions of polymerase ξ or η did not exhibit any change in repeat instability (26). Polymerase β (pol β) was shown to promote repeat instability during 8-oxo-7,8-dihydro-2'-deoxyguanosine (8-oxoG) repair. It was proposed that pol β promotes repeat expansion when a hairpin is formed at the 5' end (19). However, it is still unknown whether pol β has similar activity to promote CAG/CTG repeat expansion when the hairpin is formed at the 3' end. Identifying the polymerase involvement in hairpin-mediated expansion is another key factor to understand how CAG/CTG repeat instability is developed.

1.3 Repeat instability in non-dividing cells

Patients with repeat instability associated disorders exhibit uncontrolled neuronal cell death due to repeat expansion, indicating that repeat instability can also occur in terminally differentiated cells (3). Therefore, DNA replication cannot be used to explain repeat instability in these non-diving cells (3,27). However, other DNA metabolic processes such as DNA repair and transcription are still active in non-diving cells (28-30), suggesting those processes can promote repeat instability.

1.3.1 Errors generated from DNA repair

In mammalian cells, mismatch repair (MMR), base excision repair (BER), nucleotide excision repair (NER) and double-strand break repair are responsible for correcting errors and maintaining genomic stability (21). However, these repair processes may trigger DNA strand breaks and single-stranded DNA exposure, allowing secondary structure formation of the repeat sequence. Therefore, these protective DNA repair pathways are proposed to cause repeat instability.

1.3.1.1 Mismatch repair

Misincorporation of nucleotides may occur during DNA synthesis. The mismatch repair (MMR) pathway can maintain genomic stability by repairing the misincorporated nucleotides. The MMR pathway relies on MutSα (MSH2-MSH6) and MutSβ (MSH2-MSH3) for mismatch recognition. MutSα is responsible for repairing base-base

mismatches and 1-2 nucleotide insertion/deletion mispairs, while MutSβ is involved in correcting loops consisting of 1 to 12 nucleotides (31). Defects in MMR cause microsatellite instability and predispositions to colorectal and other cancers (31). However, the role of mismatch repair in maintaining CAG/CTG repeat stability is still unclear.

There is considerable evidence that MMR, under certain circumstances, can be mutagenic and lead to CAG/CTG repeat instability. MMR-deficient *E. coli* shows higher CAG/CTG repeat stability than that of the MMR-proficient strain (32). In contrast, MSH2 (33) and MSH3 (34) deficiency can stabilize CAG repeats in transgenic HD (Huntington's disease) mice, indicating that MSH2 or MSH3 triggers repeat instability. *In vitro* binding assays reveal that MutSβ can bind to CAG hairpins, but binding to hairpins inhibits ATP binding and hydrolysis by MutSβ (34). Since ATP hydrolysis by mismatch recognition proteins (e. g., MutSα) is essential for MMR (35), it was proposed that MutSβ may stabilize the CAG hairpin by inhibiting normal MMR and protect the hairpin from other hairpin repair machinery (34). This hypothesis is based on the assumption that MutSβ behaves similarly as MutSα when binding to a heteroduplex. However, a recent study demonstrated that these two mismatch recognition proteins display distinct biochemical and biophysical activities during mismatch recognition. For example, mismatch binding stimulates the ATP binding and hydrolysis activities of MutSα, but inhibits these activities of MutSβ (36).

In contrast to the above, some studies suggest that MMR does not play any role in repeat instability. In one study, deletion of *MSH2*, *MSH3* or *PMS1* did not result in any large CAG/CTG deletions in yeast (37), while in a second study, there was no

endogenous CTG repeat instability observed in human MMR-deficient cells (38).

Moreover, mismatch repair deficient extracts were found to be proficient in *in vitro*

hairpin repair assays (39). Recently, Tian et al. showed that although binding to a

CAG/CTG hairpin reduces MutSβ ATP binding and ATPase activities, there is no

difference in MutSβ affinity for ATP and its ATPase activity when bound to a CAG

hairpin compared to the typical insertion/deletion single-stranded loop substrate (36).

Finally, an excess amount of MutSβ was found not to inhibit hairpin repair, eliminating

the possibility of a MutSβ inhibitory role in the CAG/CTG hairpin repair machinery (36).

Experiments conducted by Pearson's group may explain the above contrasting

findings regarding MMR and CAG/CTG repeat instability. They found that MMR can

both promote and prevent CTG repeat instability, depending upon the hairpin type (40).

MSH2-deficient extract from LoVo cells was proficient in repairing large CTG hairpins

(>20 repeats) but not short CTG hairpins (<3 repeats), supporting the idea that MMR can

repair small loop structures. However, in the same experiment, cell extract with MutSβ

overexpression inhibits hairpin repair on short CTG hairpin, supporting the idea of an

inhibitory role of MutSβ in short hairpin repair. Therefore, Pearson's group concluded

that the structural variation (size) of the hairpin and MutSβ concentration may explain the

differences found by different research groups (40).

Several weaknesses are found in the Pearson's experiment. First, CTG with one

repeat is used in most of the experiments. The size of the CTG repeat is too short for

hairpin formation since two CTG repeats are the minimum requirement for a hairpin

formation (41). It is proposed that MutSβ promotes repeat instability by inhibiting repair

of multiple small CAG/CTG hairpins on a DNA. However, it is still not clear whether the

CAG/CTG expansion is caused by the multiple small hairpins formation or a big hairpin (40). Therefore, Pearson's hypothesis will be consolidated when similar hairpin assays are conducted with various CAG/CTG hairpins.

1.3.1.2 Nuclear Excision Repair (NER)

Nucleotide excision repair (NER) recognizes DNA damage caused by helical distortion of DNA. Since a hairpin loop can also distort the normal helical DNA structure, it is hypothesized that mistakes during NER may result in trinucleotide instability (21). NER consists of global genome repair (GGR) and transcription-couple repair (TCR). TCR recognizes lesions by the stalling of the transcription complex and then recruits Cockayne syndrome proteins A and B (CSA and CSB) while GGR senses the lesion directly by Xeroderma pigmentosum complementation group C (XPC), a protein unique to GGR. Once the lesion is recognized, GGR and TCR share the same downstream pathway (42) for lesion repair.

There is more data supporting a role for TCR than for GGR in promoting repeat instability. Studies reveal that bacterial strains with uvrA or uvrB mutations (NER proteins) exhibit CAG repeat instability only in the presence of transcription (43). Since TCR and GGR shares the same uvrA and uvrB proteins, instability occurs in the presence of transcription indicates TCR plays more important role in triggering repeat instability. Moreover, knocking out XPC (GGR protein) in HD mice did not result in CAG repeat instability (44), and siRNA knockdown of XPC in human cells did not affect CAG repeat

instability (45), suggesting that GGR is not involved in repeat instability. Conversely, siRNA knock down of CSB, a protein involved in TCR, can reduce CAG contraction in human cells (45).

Since siRNA knockdown of CSB and NER downstream components ERCC1 and XPG can also reduce CAG contraction, it is proposed that normal TCR activity on the CAG/CTG hairpin results in repeat instability (45). During transcription, RNAPII induces CAG/CTG secondary structure formation. The stalled RNA polymerase II will recruit TCR repair proteins to remove the hairpin. Since CAG will form a less stable hairpin structure, it is likely that it will branch migrate. It may form a crucifix structure with unequal CAG and CTG number. This crucifix structure will hinder the second round of RNA polymerase II progression, triggering the TCNER. Depending on the unequal repeat number on the transcribed or untranscribed strand, CAG/CTG repeats may expand or contract after the TCR (28). It is more likely for long CAG/CTG repeats to form a hairpin, so the proposed TCR model can also be used to explain the threshold of repeat number in repeat instability associated diseases.

1.3.1.3 Base Excision Repair (BER) induces repeat instability

DNA nucleotides can be modified by chemicals, radiation and oxidative stress resulting in mutations to the genome. Base excision repair (BER) can repair the modified bases to help maintain genomic integrity (46). However, oxidative stress can induce CAG repeat instability in human HD fibroblasts (47), so it has been proposed that under certain circumstances BER can promote repeat instability. Errors generated from BER can

explain the age-dependent CAG expansion seen in HD patients (47). High oxidative

stress is found in the human brain due to high content of polyunsaturated fatty acids, high

rate of oxygen consumption and relatively low antioxidant capacity (48). 8-oxo-7,8-

dihydro-2'-deoxyguanosine (8-oxoG), a major DNA lesion produced by oxidative stress

(49) can be repaired by BER. *In vivo* and *in vitro* assays reveal that BER of 8-oxoG can

result in repeat instability. Transgenic HD mice with knockdown of 8-oxoguanine DNA

glycosylase (OGG1), an enzyme involved in removing 8-oxoG, reduces CAG expansion

(47). In addition, high amounts of pol β, a BER enzyme, was found to accumulate on the

CAG repeat from striatum where CAG expansion is found to be most prevalent in HD

mice (50). *In vitro* assays demonstrate CAG expansion after BER of 8-oxoG on CAG

repeats. Since long patch BER is involved in strand displacement, it is hypothesized that

in long patch BER, after 8-oxoG removal, strand displacement caused by either pol β or δ

can promote the formation of a 5'flap at a CAG repeat. 5' flap removal may not occur

when the CAG repeat forms a hairpin. Expansion occurs when pol β fills the gap,

followed by nick sealing by DNA Ligase I (Fig. 1.1) (19). Even though the size of the

expansion caused by BER is usually not large, oxidative stress increases in the aging

brain (51). Therefore, it is likely that multiple rounds of BER results in a progressive

expansion in areas containing trinucleotide repeats (27). Accumulation of errors through

BER during aging may help to explain the age dependent CAG expansion in HD patients.

1.3.1.4 Double-strand break repair

Double-strand breaks in DNA can be generated by exogenous and endogenous factors. Ionizing radiation, chemicals and reactive oxygen species can induce double strand break formation. Double-strand breaks can also be formed during replication fork impediment or during DNA synthesis on template with a single strand break (52,53). Since yeast with long CAG/CTG repeats showed higher incidence of double-strand break formation, it was hypothesized that the strand break was either caused by DNA synthesis impediment or by specific CAG/CTG hairpin nucleases (54,55). Double strand breaks in a CAG/CTG repeat region can trigger contraction in bacteria (56) and in yeast (57). Double-strand breaks can be repaired by homologous recombination and non-homologous end-joining (58). HD mice with a deficiency in DNA-PKcs (DNA protein kinase, catalytic subunit), a kinase required for the non-homologous end-joining pathway of DNA repair, do not affect the rate of CAG repeat instability, suggesting the non-homologous end-joining pathway does not contribute repeat instability (59).

Nonetheless, mutations of homologous recombination repair proteins can result in CAG/CTG repeat instability. Specially engineered zinc finger nucleases can introduce double-strand breaks in CAG/CTG repeats, causing CAG repeat contraction *in vivo* (60). However, repeat instability was not observed when cells are coexpressing the zinc finger nucleases and the dominant-negative form of RAD51 (a protein involved in homologous repair), suggesting that homologous recombination repair can be responsible for repeat instability (60). Also, mutation of recA and recB (*E. coli* recombination proteins), can reduce CAG/CTG repeat contractions in *E. coli* (61). Replication slippage during the

strand invasion, gene conversion and single-stranded DNA annealing are proposed

models for explaining repeat instability during homologous recombination repair (62).

1.3.2 Transcription induces repeat instability

Transgenic HD mice harboring an unexpressed transgene containing a CAG

repeat sequence did not exhibit any repeat instability. This contrasted with the repeat

instability seen in mice containing expressed transgenes with CAG repeats. These results

suggest that in some cases transcription may be responsible for repeat instability (63). By

using model systems in which the transcription levels of CAG/CTG repeats can be

controlled in bacterial and human cells, it is possible to study the role of transcription in

repeat instability. Using an IPTG inducible system in bacteria to control transcription

levels, it was demonstrated that active transcription can destabilize a long CAG/CTG

repeat (64). A similar type study using human cells with a Tet-On doxycycline inducible

system also showed that high transcription level can promote CAG repeat instability (45).

Furthermore, it is known that RNA polymerase II stalling during transcription can act as

a primary signal to initiate transcription coupled nucleotide excision repair (TC-NER)

(65). Thus, it was proposed that CAG/CTG hairpins can block RNA polymerase II

progression, triggering TC-NER, leading to repeat instability (66).

1.4 Mechanism of hairpin repair

CAG/CTG hairpin formation is believed to be a mutagenic intermediate that if left unrepaired will lead to repeat instability, Nonetheless, the hairpin repair mechanism is not fully understood. With the use of *in vitro* assays, some characteristics of the hairpin repair mechanism have been identified. Human nuclear extracts are capable of repairing hairpin substrates in a nick-directed manner, with the hairpin being removed by excision (39) and/or incision (67). The DNA polymerase inhibitor, aphidicolin, can abolish hairpin repair, indicating that one or more of polymerases α, δ and ε, are required for the DNA re-synthesis after hairpin removal. In addition, mismatch repair and nucleotide excision repair deficient extracts are still proficient in the hairpin repair, suggesting that the hairpin repair pathway is different from that of mismatch repair and nucleotide excision repair (39). Proteins involved in the incision during hairpin removal are not known yet, but the PCNA inhibitor (p21c) can abrogate the incision activity, indicating that PCNA may play some role with the endonuclease in the hairpin removal step. Since a low repair activity was found when CTG was used as the template for the DNA-resynthesis, it was hypothesized that a helicase may be involved in unwinding the CTG hairpin, promoting DNA synthesis on the CTG template (67).

1.5 Helicases

It is believed that a common feature of all the known metabolic pathways resulting in CAG/CTG repeat instability is CAG/CTG hairpin formation (20). Therefore, factors that can disrupt hairpin formation will help maintain CAG/CTG repeat stability (37). DNA helicases utilize energy to unwind double-stranded DNA and mutations in DNA helicases can result in genomic instability (68). Based on a number of studies, it has been proposed that DNA helicases can promote repeat stability by disrupting CAG/CTG hairpin formation (37).

Suppressor of RAD Six Screen Mutant 2 (SRS2), a yeast 3'to 5' DNA helicase, was shown to be involved in maintaining trinucleotide repeat (TNR) stability using an unbiased yeast mutant screening assay. Yeast containing a point mutant in SRS2 that destroys helicase activity had increased expansion rates of CAG, CTG and CGG repeats, confirming the role of SRS2's helicase activity in maintaining TNR stability (69). *In vitro* helicase assays revealed that SRS2 demonstrates a higher activity and specificity in unwinding CAG and CTG hairpins when compared to several other helicases, indicating that SRS2 may be the preferred helicase involved in preventing repeat expansions (37,70). Pol δ is known to interact with SRS2 through its pol 32 subunit (71). A mutant lacking pol 32 demonstrated an increased rate of CAG/CTG repeat instability compared to the wild type strain and such instability could not be rescued by SRS2 overexpression. From these results, it was proposed that SRS2 prevents CAG/CTG repeat instability by interacting with pol δ to resolve CAG/CTG hairpins during DNA synthesis (69). No human homolog of SRS2 has been identified, but F-Box helicase 1 (FBH1) is proposed to be the functional human orthologue of yeast SRS2 (72). Nonetheless, no

biochemical or genetic assays have been conducted to prove a role for FBH1 in CAG/CTG repeat stability.

Werner Syndrome Protein (WRN), a human 3' to 5' DNA helicase, was identified to prevent repeat instability. WRN unwinds (CGG) hairpins or tetraplexes, and it can enhance polymerase δ synthesis on the CGG repeats (73,74). Cells deficient in WRN showed large DNA deletions (75) and telomere loss (76), suggesting that WRN can somehow play a role in maintaining genomic stability. A mutation in SGS1, the WRN yeast homolog, triggers CTG repeat contraction, especially when CTG repeats are on the lagging-strand template (77). Therefore, it was proposed that SGS1 prevents CTG repeat contraction by resolving CTG hairpins during lagging strand synthesis. However, no data is available about whether or not the human homolog of SGS1, WRN, is also involved in CAG/CTG repeat instability.

1.6 Research Objectives

The complexity of the mechanisms involved in TNR repeat instability is evident by the number of different hairpin repair pathways and the crosstalk among the DNA repair pathways (3,27,28,67). The common points of repeat expansion are the addition of DNA (DNA synthesis) and the hairpin formation. Therefore, we proposed two specific aims to identify what polymerases and helicases are involved in repeat instability and how they act in these pathways.

Specific aim one is to identify DNA polymerase(s) involved in promoting CAG/CTG expansion. Repeat expansion requires the addition of DNA and hairpin formation is associated with hairpin repair and DNA synthesis. Since CAG/CTG hairpins hinder polymerase progression, we hypothesize that cells may treat these hairpins in a manner similar to other types of lesions that trigger translesion synthesis (26). To test this hypothesis, we established assay to screen for different error-prone DNA polymerases to promote repeat expansion. With low processivity of the error-prone polymerases, it is likely that cells may adopt the polymerase switching system to promote repeat expansion. Therefore, once the error-prone polymerase is identified, the assay will be conducted in a purified system containing the replicative polymerase and the target polymerase.

Specific aim two is to identify helicases responsible for unwinding the CAG/CTG hairpins. Our previous research showed that low hairpin repair activity was observed when CTG hairpin was formed in the template strand (CTG slip-in). A-A base pairing has weaker stacking than that of T-T base pairing, CTG forms a more stable hairpin than that of CAG (3,15). In addition, the formation of a stable hairpin on the template strand hinders polymerase progression, so we hypothesize that a helicase activity is involved

during the resynthesis step of CTG slip-in repair (67). To test this hypothesis, different

HeLa nuclear extract fractions were added in the CTG slip-in hairpin repair assay to

screen for potential helicase leading to higher repair activity. After identifying the

helicase, biochemical assays like helicase assay and primer extension assay will be

carried out to characterize how the helicase contributes to enhanced DNA repair activity

CHAPTER 2 Materials and Methods

2.1 Basic techniques

2.1.1 Chemicals and reagents

Affymetrix (Formerly USB): Agarose, Ammonium Sulfate, Ammonium Persulfate, Cesium Chloride (CsCl), EDTA, Ethidium Bromide, Exonuclease V, Heparin Salt, Imidazole, Phenol, Sodium Dodecyl Sulfate (SDS), T4 polynucleotide kinase.

Fisher Scientific: 1- Butanol, Acetic acid, Formamide, Isopropanol, Polyethylene Glycol 8000 (PEG 8000), Potassium Acetate, Potassium Phosphate Dibasic (K_2HPO_4), Potassium Hydroxide (KOH), Potassium Phosphate Monobasic (KH_2PO_4), Sodium Hydroxide (NaOH), Sodium Phosphate Dibasic (Na_2HPO_4), Sodium Phosphate Monobasic (NaH_2PO_4), Tween-20.

Millipore: Amicon Ultra-4 Centrifugal Filter Devices

Nalgene: 0.22μm filters

Perkin Elmer: [γ-^{32}P]-ATP.

Research Products International Corp (RPI): 2XYT Broth, Ampicillin, HEPES, LB Broth, Acrylamide, N,N´-Methylenebisacrylamide, Tetracycline Hydrochloride, Urea, X-ray Film.

Roche: Adenosine Triphosphate (ATP), Deoxynucleotide Triphosphate (dNTP), Dithiothreitol (DTT), Noniodent P-40 (NP-40), Quick Spin Column (TE).

Sigma Aldrich: 2-Mercaptoethanol, Boric Acid, Bromophenol Blue, GenElute™ Gel Extraction Kit, Glycerol, Sodium Chloride (NaCl), Polyvinylpyrrolidone, Sodium Citrate, Sucrose, Tris, Xylene Cyanol.

GE Healthcare: ECL Detection Reagent, Hybond-NX, 5 mL Histrap column, 1 mL Mono-Q column, 1 mL Mono-S column, Phenyl-Sepharose beads, Sephacry S-300 beads, 1 mL SP-Sepharose column, S-Sepharose beads.

Stratagene: *E.coli* BL21 (DE3) pLysS cells, XL1 Blue competent cells.

VWR: Spectra/Por 1 Dialysis Membrane (Dialysis Bag), Ethanol, Methanol.

Integrated DNA Technologies: All the DNA oligonucleotides.

New England Biolab: Restriction enzymes, T4 DNA ligase.

Santa Cruz Biotechnology: WRN antibody, polymerase δ, polymerase ε antibodies.

2.1.2 Primers used in the experiment

Primer Name (Original name)	Sequence	Application
M13GC V6109	CGG ATA ACA ATT TCA CAC AGG	Probe for Southern blot
6135 M13MP18F	CTATGACCATGATTACGAATTC	Probe for Southern blot
mCTG15	A*mC*mG*mACGGCCAGTGCCAAGCTT(CTG)$_{15}$	Hairpin(oligo- based) and helicase assay
mCTG15+2	A*mC*mG*mACGGCCAGTGCCAAGCTT(CTG)$_{15}$GA	Hairpin (oligo- based) 2 extra bp
mCTG15+2MM	A*mC*mG*mACGGCCAGTGCCAAGCTT(CTG)$_{15}$GG	Hairpin (oligo- based) 2 extra bp with mismatch
mCTG15PE+5	A*mC*mG*mACGGCCAGTGCCAAGCTT(CTG)$_{15}$GAATT	Hairpin (oligo- based) 5 extra bp
mCTG15PE+5MM	A*mC*mG*mACGGCCAGTGCCAAGCTT(CTG)$_{15}$GAATG	Hairpin (oligo- based) 5 extra bp with mismatch
MP18C mCTG15PE3m	A*mC*mG*mACGGCCAGTGCCAAGCTT(CTG)$_{14}$*mC*mU*mG	Hairpin (oligo- based) 3' exo resistant
mCTG25PE	A*mC*mG*mACGGCCAGTGCCAAGCTT(CTG)$_{25}$	Hairpin (oligo- based) 25 repeats
pBstNI (C BstNI Cut)	GGGTAACGCCAGGGTTTTC	Digestion primer for primer extension
pBsrBI (MP18C6115 BsrBI)	TTA TCC GCT CAC AAT TCC ACA	Digestion primer for primer extension
Synthesis primer (Mp18 mCtl-19)	A*mG*mU*mCACGACGTTGTAAAAC	Synthesis primer for primer extension
MP18C CAG15	A*mC*mG*mACGGCCAGTGCCAAGCTT(CAG)$_{15}$	Helicase assay
MP18C CAG35PE	A*mC*mG*mACGGCCAGTGCCAAGCTT(CAG)$_{35}$	Helicase assay
MP18C CTG35PE	ACGACGGCCAGTGCCAAGCTT(CTG)$_{35}$	Helicase assay

Table 2.2 Primers used in the experiment. KEY * phosphothiolate bond, m: 2' methyl base

2.1.3 Preparation of buffers

De-ionized distilled water was used to prepare all solutions for the assay. Sterilization was done either by filtration through a 0.22μm filter or autoclaving for 15 min at 121 °C.

2.1.4 Agarose gel electrophoresis

For every 5 μL of sample volume containing DNA, 1 μL of 6X DNA loading buffer (10 mM Tris, pH 7.6, 0.03% bromophenol blue, 0.03% xylene cyanol, 60% glycerol, 60 mM EDTA) was added. Agarose gel electrophoresis was conducted in TAE running buffer (40 mM Tris-Acetate, pH 8.0, 2 mM EDTA). DNA staining was conducted by gentle shaking of the agarose gel in the presence of 100 mL of sterile distilled water containing 0.5 μg/mL ethidium bromide for 5 min. Destaining was conducted by replacing the solution with fresh water with gentle shaking for 10 min. DNA was visualized by a UV transilluminator (Gel logic 112, Kodak).

2.1.5 Urea Denaturing PAGE and Southern blot analysis

For every 1 μL of sample volume containing DNA, 1 μL of 2X SSCP loading buffer (95% formamide, 0.075% xylene cyanol, 0.07, 5% bromophenol blue, 20 mM EDTA) was added. After heating for 5 min at 95 °C, the DNA was kept on ice before loading. Denaturing gels (6%) were prepared by mixing polyacrylamide solution (19:1 (acr:bis)) with 8 M urea in the presence of 1X TBE buffer (40 mM Tris, pH 8.0, 89 mM

boric acid and 2 mM EDTA). Denaturing gel electrophoresis was conducted in 1X TBE buffer at 8 W at room temperature. After electrophoresis, the gel was electrotransferred to a nylon membrane (Hybond-NX) in 1X TBE for 1 hr at 1 mA at 4 ^{o}C in Hoefer TE 42. The membrane was dried and cross-linked using a UV transilluminator box (Fisher Scientific) for 7 min. The membrane was then pre-hybridized with hybridization buffer (50 mM Tris-HCl, pH 7.5, 1 mM EDTA, 1 M NaCl, 2% SDS, 0.5% polyvinylpyrrolidone, 0.2% heparin) for 15 min at 37 ^{o}C in a roller bottle using a Techne-Hybridiser HB-2D. The membrane was then incubated with ^{32}P-labeled oligonucleotide overnight at 37 ^{o}C. The membrane was washed two times with 2X washing buffer (0.03 M sodium citrate, pH 7.0, 0.3 M NaCl, 0.1% SDS) and two times with 1X washing buffer (0.015 M sodium citrate, pH 7.0, 0.15 M NaCl, 0.1% SDS) for 10 min each. The membrane was then dried and exposed to X-ray film.

2.1.6 Native polyacrylamide gel electrophoresis

Native polyacrylamide gel (6%) was prepared by mixing polyacrylamide solution (19:1 (acr:bis)) with 0.5X TBE buffer (20 mM Tris, pH 8.0, 0.445 mM boric acid and 1 mM EDTA) and 2.5 % glycerol. Native polyacrylamide gel electrophoresis was conducted in 0.5X TBE buffer at 120 V at room temperature.

2.1.7 SDS-PAGE and Western blot analysis

One third volume of protein loading buffer (120 mM Tris-HCl, pH 6.8, 3% SDS, 15%, 2-mercaptoethanol, 10% glycerol and 0.075% bromophenol blue) was added to the protein and heated for 5 min at 95 $^{\circ}$C. The heated samples were then resolved in an 8-12% SDS-PAGE gel in running buffer (25 mM Tris-HCl, pH 7.5, 0.2 M glycine, 0.1% SDS) at 160 V at room temperature.

After SDS-PAGE, proteins were electrotransferred onto a nitrocellulose membrane at 400 mA for 1 hr in transfer buffer (25 mM Tris, pH 7.5, and 191 mM glycine, 20% methanol). The membrane was then blocked using 5% non-fat dry milk (NFDM) in TBST (10 mM Tris-HCl, pH 7.5, 0.8% NaCl, 0.1% Tween-20) for 15 min at room temperature. The membrane was then incubated with primary antibody against the target protein in the presence of 5% NFDM in TBST overnight at 4 $^{\circ}$C with gentle rocking. The next day the membrane was washed with TBST with gentle rocking in at room temperature for 10 min each (3 times in total). Secondary antibody was added to the membrane in the presence of 5% NFDM in TBST with gentle rocking for 1.5 hr at room temperature. The membrane was washed with TBST for 3 times (10 min each). Proteins of interest were visualized by using ECL Detection Reagent and the signal was captured on an X-ray film.

2.1.8 ^{32}P T4 polynucleotide kinase (T4PNK) 5' end labeling

Oligonucleotide (10 pmole) was incubated with 20 pmol of [γ-^{32}P] ATP at 6000 Ci/mmol in 50 μL reaction volume containing 0.5 M Tris-HCl, pH 7.6, 100 mM MgCl2,

100 mM 2-mercaptoethanol and 3 units of T4PNK at 37 $^{\circ}$C for 30 min. The reaction was terminated by heating at 65 $^{\circ}$C for 10 min. The labeled DNA (50 µL) was then applied to a dried Quick Spin Column and centrifuged for 4 min at 1100 x g at room temperature. 5' end labeled DNA was collected in the eluate.

2.2 Preparation of nuclear extracts and proteins

2.2.1 Nuclear extracts preparation

2.2.1.1 Cell culture

HeLa S3 cells (National Cell Culture Center) were cultured in Eagle's minimal essential medium (Mediatech), supplemented with 10% FBS (HyClone Laboratories), and kept at 37 $^{\circ}$C. High Five insect cells (Invitrogen) were cultured in TNM-FH (Invitrogen) with 10% FBS and kept at 27 $^{\circ}$C.

2.2.1.2 Nuclear extract preparation

One liter of human cells was collected by centrifugation in a RC-3B centrifuge at 3,200 rpm for 8 min. After that, 20 mL washing buffer (20mM HEPES-KOH, pH 7.5, 5 mM KCl, 0.5 mM MgCl$_2$, 0.2 % Sucrose, 0.5 mM DTT, 1X protease inhibitor containing 100 mM PMSF, 191.5 mM benzamidine, 0.05 g/L pepstatin A, 0.05 g/L leupeptin) was added to resuspend the cells, followed by 5 min centrifugation at 4,500 rpm. The cells were then lysed using a Dounce homogenizer in the presence of 8 mL Hypotonic buffer (20 mM HEPES-KOH, pH 7.5, 0.5 mM MgCl$_2$, 5 mM KCl, 0.5 mM DTT, 1X protease inhibitor). The nuclear pellet was resuspended in 12 mL Extraction buffer (50 mM HEPES-KOH, pH7.5, 10% sucrose, 0.5 mM DTT, 1X protease inhibitor). NaCl (5 M)

was added to make the final salt concentration 155 mM. Nuclear proteins were extracted by gentle rocking at 4 $^\circ$C. After centrifugation at 11,000 rpm for 20 min, ammonium sulfate (0.42 g/mL) was slowly added to the supernatant with gentle stirring. The precipitated nuclear proteins were collected by centrifugation at 11,500 rpm for 20 min. The pellet was resuspended in 1 mL dialysis buffer (25 mM HEPES-KOH, pH7.9, 50 mM KCl, 0.1 mM EDTA, 2 mM DTT, 1X protease inhibitor) and transferred into a dialysis bag for dialysis against 1 liter dialysis buffer at 4 $^\circ$C. Dialysis stopped when KCl concentration reached between 100 mM to 200 mM. After dialysis, the nuclear extracts were centrifuged at 14,000 rpm for 15 min at 4 $^\circ$C. The supernatant was aliquoted, frozen in liquid nitrogen and stored in -80 $^\circ$C.

2.2.2 Phosphocellulose chromatography of HeLa nuclear extract

Phosphocellulose chromatography (P-11) fraction preparation was described previously (78). Briefly, 300 mg HeLa nuclear extract was diluted to 5 mg/mL with buffer p-11$_A$ (25 mM HEPE-KOH, pH 7.5, 0.1 mM EDTA, 1X protease inhibitor). Diluted extract was loaded onto a phosphocellulose column (Whatman P-11, 6 cm by 10 cm^2) equilibrated with buffer p-11$_A$ containing with 50 mM KCl. After washing with 200 mL of buffer p-11$_A$, the column was eluted with 0.8 liter linear gradient of 0.05 to 1.3 M KCl in buffer p-11$_A$ and eluted samples collected by a fraction collector (6 mL/tube). Each tube was dialyzed against 1 liter buffer p-11$_A$ containing 100 mM KCl. Finally, 120 P-11 fractions were then aliquoted, frozen in liquid nitrogen and stored at -80 $^\circ$C.

2.2.3 Purification of polymerase δ and exonuclease mutants

Baculovirus stocks for expressing polymerase δ subunits p12, p50, p66 and p125$_{WT}$ were obtained from Ellen Fanning (Vanderbilt University) and the cDNA for p125$_{D402A}$ was obtained from Yoshihiro Matsumoto (Fox Chase Medical Center). The p125$_{D402A}$ cDNA was modified and cloned into the pFastBac vector for baculovirus stock preparation. The four polymerase δ subunits were co-expressed in High Five insect cells and purified as described (79). Briefly, the insect cells were lysed by Dounce homogenizer in the presence of buffer Pol δ$_{lysis}$ (20 mM Tris-HCl, pH 7.8, 100 mM NaCl, 0.2% NP-40, 20 mM imidazole and 1X proteinase inhibitors). The cell lysate was then centrifuged at 18,000 rpm for 45 min. The supernatant was then loaded onto a 5 mL HisTrap column and eluted with 30 mL linear gradient from 20 mM to 400 mM imidazole of buffer Pol δ$_A$ (20 mM Tris, pH 7.8, 100 mM NaCl, 0.02% NP-40 and 1X proteinase inhibitor). Polymerase δ was eluted in 300 mM imidazole. Fractions containing polymerase δ was then diluted 5 times with buffer Pol δ$_B$ (20 mM Tris pH 7.8, 0.02% NP-40, 10 % glycerol, 1 mM DTT and 1X proteinase inhibitor) and loaded onto a 1 mL Mono Q column. Proteins were eluted in a 30 mL linear gradient of 20 mM to 400 mM NaCl in buffer Pol δ$_B$ and polymerase δ was eluted at 300 mM NaCl. Fractions containing polymerase δ were then diluted 5 times with buffer Pol δ$_B$ and loaded onto a 1 mL Mono S column. Proteins were eluted in a 30 mL linear gradient of 20 mM to 500 mM NaCl and polymerase δ was eluted at 400 mM NaCl. Polymease δ was then aliquoted, frozen in liquid nitrogen and stored in -80 ^{o}C. Purification of polymerase δ exonuclease mutant δ$_{D402A}$ was performed using the same method as used for the wild-type.

2.2.4 Purification of pol β

Baculovirus stocks for pol β were obtained from Yanbin Zhang (University of Miami). Pol β was expressed in High Five insect cells. After virus inoculation, the insect cells were lysed using a Dounce homogenizer in the presence of buffer Polβ_A (25 mM HEPES-KOH, pH 8.0, 20 mM imidazole, 300 mM NaCl, , 10% glycerol, 4 mM 2-mercaptoethanol and 1X proteinase inhibitor containing 100 mM PMSF, 191.5 mM benzamidine, 0.05 g/L pepstatin A, 0.05 g/L leupeptin). The cell lysate was centrifuged at 18,000 rpm for 45 min. The supernatant was then loaded onto a 5 mL HisTrap column and eluted with a 50 mL linear gradient from 20 mM to 240 mM imidazole in buffer Polβ_A. Pol β fractions were pooled and diluted with Polβ_B (25 mM HEPES-KOH, pH 8.0, 10% glycerol, 4 mM 2-mercaptoethanol and 1X proteinase inhibitor) to a final concentration of 120 mM NaCl and loaded onto a 1 mL Mono S column. Proteins were eluted using a 30 mL linear gradient of 120 mM to 500 mM NaCl in buffer Polβ_B and pol β was then eluted using 300 mM NaCl. Fractions containing pol β were concentrated into 250 μL with an Amicon Ultra-4 Centrifugal Filter Device and loaded onto a 24 mL Superdex 200 gel filtration column. Buffer Polβ_C (25 mM HEPES-KOH, pH 8.0, 10% glycerol, 300 mM NaCl, 4 mM 2-mercaptoethanol) was then used to elute pol β with flow rate of 0.5 mL/min. Pol β was eluted in about 26 mL to 28 mL. Pol β was then aliquoted, frozen in liquid nitrogen and stored at -80 $^{\circ}$C.

2.2.5 RFC purification

Baculovirus stocks for replication factor C (RFC) subunits p36 and p40 and cDNA for p37, p38 and p140 were obtained from Dr. Bruce Stillman (Cold Spring

Harbor Laboratory). A hexahistidine tag was added to p38 and p140 previously in our lab. The five subunits of RFC were co-expressed in High Five insect cells and purified as described (80). Briefly, the insect cells were lysed using a Dounce homogenizer in the presence of 30 mL buffer RFC$_A$ (50 mM Na$_2$HPO$_4$, pH8.0, 0.5 M NaCl, 10 mM imidazole, 10% glycerol, 5 mM 2-mercaptoethanol, 0.5% NP-40 and 1X proteinase inhibitor. The cell lysate was then centrifuged at 18,000 rpm for 45 min. The supernatant was then loaded onto a 5 mL HisTrap column and eluted with 30 mL linear gradient from 10 mM to 400 mM imidazole in buffer RFC$_A$. RFC was eluted at about 300 mM imidazole. RFC fractions were then concentrated to 250 µL with an Amicon Ultra-4 Centrifugal Filter Device and loaded onto a 24 mL Superdex 200 gel filtration column. Buffer RFC$_B$ (25 mM HEPES-NaOH, pH 8.0, 150 mM NaCl, 1 mM EDTA, 2 mM DTT) was then used to elute RFC with flow rate of 0.5 mL/min. RFC was eluted in about 22 mL to 24 mL. RFC was then aliquoted, frozen in liquid nitrogen and stored at -80 $^{\circ}$C.

2.2.6 PCNA purification

The profilerating cell nuclear antigen (PCNA) expression plasmid was a gift from Dr. Bruce Stillman (Cold Spring Harbor Laboratory). PCNA was overexpressed in *E.coli* BL21 (DE3) pLysS cells and purified as described (81). Briefly, 2 liters of the *E.coli* overexpressing PCNA was lysed with sonication in the presence of 30 mL buffer PCNA$_A$ (25mM Tris-HCl, pH 7.4, 1mM EDTA, 25 mM NaCl, 0.01% NP-40, 1mM DTT, 1X protease inhibitor containing 100 mM PMSF, 191.5 mM benzamidine, 0.05 g/L pepstatin A, 0.05 g/L leupeptin). The cell lysate was then centrifuged at 18,000 rpm for 45 min. The supernatant was then loaded onto a 10 mL Q-Sepharose column and eluted with a

200 mL gradient from 0.2 to 0.7 M NaCl in buffer PCNA$_A$. PCNA was eluted at 0.4 M

NaCl. PCNA containing fractions were pooled and dialyzed in HAP buffer (25 mM

KPO$_4$, pH 7.0, 0.01 NP-40, 10% glycerol, 5 mM DTT, 1X protease inhibitor containing

100 mM PMSF, 191.5 mM benzamidine, 0.05 g/L pepstatin A, 0.05 g/L leupeptin). After

dialysis, the protein was loaded onto a 10 mL S-Sepharose column . The flow through

fraction containing PCNA was then loaded onto a 20 mL hydroxylapatite column and

eluted with a 500 mL gradient of 25 mM to 500 mM KPO$_4$ in HAP buffer. PCNA was

eluted at about 300 mM KPO$_4$. PCNA fractions were then pooled and dialyzed against

one liter of 1.2 M NaCl in buffer PCNA$_B$ (25 mM KPO$_4$, pH 7.0, 0.01 NP-40, 5 mM DTT,

1X protease inhibitor). The dialysate was then loaded onto a 25 mL Phenyl-Sepharose

column and eluted with a reverse linear gradient from 1.2 M to 0 M NaCl in buffer

PCNA$_B$. Fractions containing PCNA were pooled and dialyzed agains buffer PCNA$_C$ (25

mM KPO$_4$, pH 7.0, 0.01 NP-40, 20% sucrose, 5 mM DTT, 1X protease inhibitor). PCNA

was then aliquoted, frozen in liquid nitrogen and stored at -80 ^{o}C.

2.2.7 WRN purification

WRN baculovirus was obtained from Dr. Lawrence Loeb (University of

Washington). WRN was expressed in High Five insect cells. After virus inoculation, the

insect cells were lysed using a Dounce homogenizer in the presence of 30 mL buffer

WRN$_A$ (150 mM Tri,s pH 8.0, 20 mM imidazole, 10% glycerol, and 50 mM NaCl and

1X proteinase inhibitor containing 100 mM PMSF, 191.5 mM benzamidine, 0.05 g/L

pepstatin A, 0.05 g/L leupeptin). The cell lysate was then centrifuged at 18,000 rpm for

45 min. The supernatant was then loaded onto a 5 mL HisTrap column and eluted with a

30 mL linear gradient from 20 mM to 300 mM imidazole in buffer WRN$_A$. WRN

fractions were pooled and loaded onto a 1 mL SP-Sepharose column equilibrated with

buffer WRN$_B$ (150 mM Tris, pH 8.0, 10% glycerol, and 50 mM NaCl and 1X proteinase

inhibitor). WRN was eluted with a 30 mL linear gradient from 20 mM to 500 mM NaCl

in buffer WRN$_B$. WRN was eluted from the columnat about 400 mM NaCl. WRN

containing fractions were concentrated to 250 μL with an Amicon Ultra-4 Centrifugal

Filter Device and loaded onto a 24 mL Superdex 200 gel filtration column. Buffer WRN$_C$

(150 mM Tris, pH 8.0, 10% glycerol, and 160 mM NaCl and 1X proteinase inhibitor

containing 100 mM PMSF, 191.5 mM benzamidine, 0.05 g/L pepstatin A, 0.05 g/L

leupeptin) was used to elute WRN with flow rate of 0.5 mL/min. WRN was eluted in

about 23 mL to 26 mL. WRN was then aliquoted, frozen in liquid nitrogen and stored at -

80 ^{o}C.

2.3 DNA substrate preparation

2.3.1 Single-stranded and double-stranded DNA preparation

2.3.1.1 Phage stock preparation

M13MP18 plasmid containing CAG/CTG repeats were prepared by former

colleague. They were transformed into the compentent *E.coli* XL-1 Blue1 cells

(Stratagene). Positive colonies were confirmed by sequencing. The positive colonies were

cultured in 50 mL of 2XYT media overnight. The phage stocks were collected by

centrifugation at 18,000 rpm for 20 min and stored at 4 ^{o}C.

2.3.1.2 Single-stranded DNA preparation

XL1-Blue overnight culture (30 mL) was used to inoculate 3 liters of 2XYT and shaken at 220 rpm at 37 $^{\circ}$C. When the OD_{595} reached 0.3, phage stock (3 mL) was added into the culture and the culture incubated for 8 hours at 220 rpm at 37 $^{\circ}$C. For harvesting, the culture was kept on ice for 30 min, followed by 30 min centrifugation at 4,500 rpm at 4 $^{\circ}$C. Single-stranded DNA (ssDNA) was extracted from the supernatant while double-stranded DNA was extracted from the cell pellet.

For single-stranded DNA extraction, NaCl (36 g/L) and polyethylene glycol 8,000 (50 g/L) were added to the supernatant of the culture. After 60 min stirring at room temperature, the supernatant was centrifuged for 30 min at 18,000 rpm at 4 $^{\circ}$C. The phage pellet was resuspended in 30 mL 10 mM Tris-HCl, pH 8.0. CsCl (0.4348 g for every 1 g of phage solution) was added to the resuspended phage pellet. After 16 hours of centrifugation at 45,000 rpm at 25 $^{\circ}$C, phage particles appeared as a viscous layer. The phage layer was then extracted with a syringe and dialyzed against 1 liter of TE buffer (10mM Tris-HCl, pH 8.0, 1 mM EDTA) three times (buffer changed every 4 hr). An equal amount of phenol solution was added to the phage solution with vigorous mixing, followed by 5 min centrifugation at 8,000 rpm at room temperature. The upper layer was removed to a clean tube and the phenol extraction repeated three more times. After phenol extraction, the single-stranded DNA was dialyzed against 1 liter TE buffer, pH 8.0, three times at 4 $^{\circ}$C (buffer changed every 4 hr). Single-stranded DNA was stored at 4 $^{\circ}$C.

2.3.1.3 Double-stranded DNA preparation

Double-stranded DNA was extracted from the cells pellet mentioned in 2.3.1.2. The pellet was resuspended in 60 mL cold Solution I (25 mM Tris-HCl, pH 8.0, 10 mM EDTA, 0.9% glucose) and kept on ice for 10 min. Then, 120 mL of Solution II (0.2 N NaOH and 1 % SDS) was added to the mixture with gentle stirring, followed by 90 mL of Solution III (3 M KOAc, 2 M CH_3COOH) addition with gentle stirring. The cell lysate was then centrifuged for 30 min at 4,500 rpm at 4 ^{o}C. The supernatant filtered through 4 layers of cheesecloth was and then mixed with 0.6 volume of isopropanol on ice. The mixture was then centrifuged for 30 min at 4,500 rpm at 4 ^{o}C. DNA appeared in the form of a white pellet. The pellet was washed with 50 mL of 70% ethanol followed by 15 min centrifugation at 4,500 rpm at room temperature. The DNA pellet was air-dried for 10 min and resuspended in 20 mL TE buffer (10mM Tris-HCl, pH 8.0, 1 mM EDTA). The resuspended solution was weighed and mixed with CsCl (1.05 g/g) and ethidium bromide (50 μLμL/g). Gradient centrifugation was performed for 18 hr at 45,000 rpm at 25 ^{o}C. After centrifugation, supercoiled double-stranded DNA appeared as a red band in the centrifuge tube. The double-stranded DNA was then extracted from the tube using a syringe. An equal volume of butanol was added to the solution containing the double stranded DNA with gentle mixing. After 5 min centrifugation at 1,000 rpm at room temperature, the clear upper layer was transferred to a clean tube and the butanol extraction repeated three more times. After four rounds of butanol extraction, the double-stranded DNA was dialyzed against 1 liter TE buffer, pH 8.0, at 4 ^{o}C (buffer changed every 4 hr, three times in total). Double-stranded DNA was stored at 4 ^{o}C.

2.3.2 Hairpin substrate preparation (double-stranded hairpin substrate)

Circular hairpin substrate with a 5' nick was constructed based on the differences in repeat number between the circular single-stranded DNA and the linearized double-stranded DNA (summarized in Table 4.1). Linearized double-stranded DNA was generated by BglI digestion for three hours at 37 ^{o}C. DNA purification was conducted by phenol extraction and TE dialysis as described in 2.3.1.2.

Double-stranded DNA denaturation was performed by adding 1 mg of linearized double-stranded DNA and 0.25 µg of circular single-stranded DNA into a 30 mL reaction mixture containing 50 mM Tris-HCl, pH 7.6, 100 mM NaCl, 1 mM EDTA and 0.3 N NaOH for 5 min at room temperature. The solution was then neutralized by adding 3 mL 2.9 N CH_3COOH to the reaction mixture. The salt concentration was adjusted by adding 1.35 mL 3 M KCl and 3.7 mL 1.0 M potassium phosphate buffer (K_2HPO_4/KH_2PO_4), pH 7.5. The DNA annealing reaction was initiated by heating the reaction mixture for 30 min at 65 ^{o}C, followed by slowly cooling down to 37 ^{o}C. The annealed DNA was stored at 4 ^{o}C. Annealing efficiency was determined by 1% agarose gel electrophoresis. Hydroxylapatite (HAP) column chromatography was used to remove the single stranded DNA. Two g of hydroxylapatite resin was added into a column (2.5 cm diameter) and was equilibrated in 30 mM K_2HPO_4/KH_2PO_4, pH 6.9. The annealing reaction product was loaded onto the column with a flow rate of 10 mL/ h. Single-stranded DNA was removed by washing the column with 6 column volumes of 30 mM K_2HPO_4/KH_2PO_4, pH 6.9, followed by 2 column volumes of 160 mM K_2HPO_4/KH_2PO_4, pH 6.9. Hairpin substrates and the linearized double-stranded DNA were eluted by 3 column volumes of 420 mM

K_2HPO_4/KH_2PO_4, pH 6.9, and collected by a fraction collector (1 mL/ tube). The separation efficiency was monitored by 1% agarose gel electrophoresis.

Fractions from the HAP column containing the double-stranded DNA were combined. Butanol extraction (mentioned in 2.3.1.3) was used to concentrate the DNA to about 2 mL, and then the DNA was dialyzed against 1 liter TE buffer as described in section 2.3.1.3. Exonuclease V digestion was conducted in the reaction buffer (66.7 mM Glycine, 5 mM $MgCl_2$, 8.3 mM 2-mercaptoethanol, 0.5 mM ATP, and exonuclease V (0.2 U/μg DNA)) for 2 hr at 37 ^{o}C. Exonuclease V digestion efficiency was monitored by the disappearance of the linearized double-stranded DNA by agarose gel electrophoresis. One time of phenol extraction (mentioned in 2.3.1.2) and butanol extraction (mentioned in 2.3.1.3) were used to purify and concentrate the DNA to 200 μL.

Size exclusion column chromatography was conducted to purify the hairpin substrate from the free nucleotides. An S-300 column was made by slowly pouring 50 mL sephacryl S-300 beads into a glass column (45 cm long and 1.2 cm diameter). The S-300 column was then equilibrated with 3 column volumes of TES buffer (10 mM Tris-HC,l pH 7.6, 1 mM EDTA, and 0.3 mM NaCl). The concentrated DNA was loaded onto the column, followed by TES buffer elution (10 mL/hr). Fractions were collected with a fraction collector with 1 mL per tube. Separation efficiency was monitored by 1% agarose gel electrophoresis. Fractions containing the pure hairpin substrate were combined and dialysis against 1 liter TE buffer for 3 times (buffer change every 4 hr). Hairpin substrate concentration was determined by the OD_{260} absorbance at room temperature. The whole process of substrate preparation was described in Fig. 2.1A and B.

2.3.3 Hairpin substrate preparation (oligo-based)

The 5' exonuclease resistant oligonucleotide containing 15 CTG repeats (described in Fig. 2.1) was annealed with the circular single stranded DNA containing 10 CAG repeats in a 100 μL reaction containing 1.67 M NaCl. Based on the repeat number difference, a substrate with 5 CTG hairpin was formed (Fig. 2.1C). BbvI and T7 endonuclease I as described previously (16) confirmed that our substrates contained a 3' CTG hairpin slippage structure.
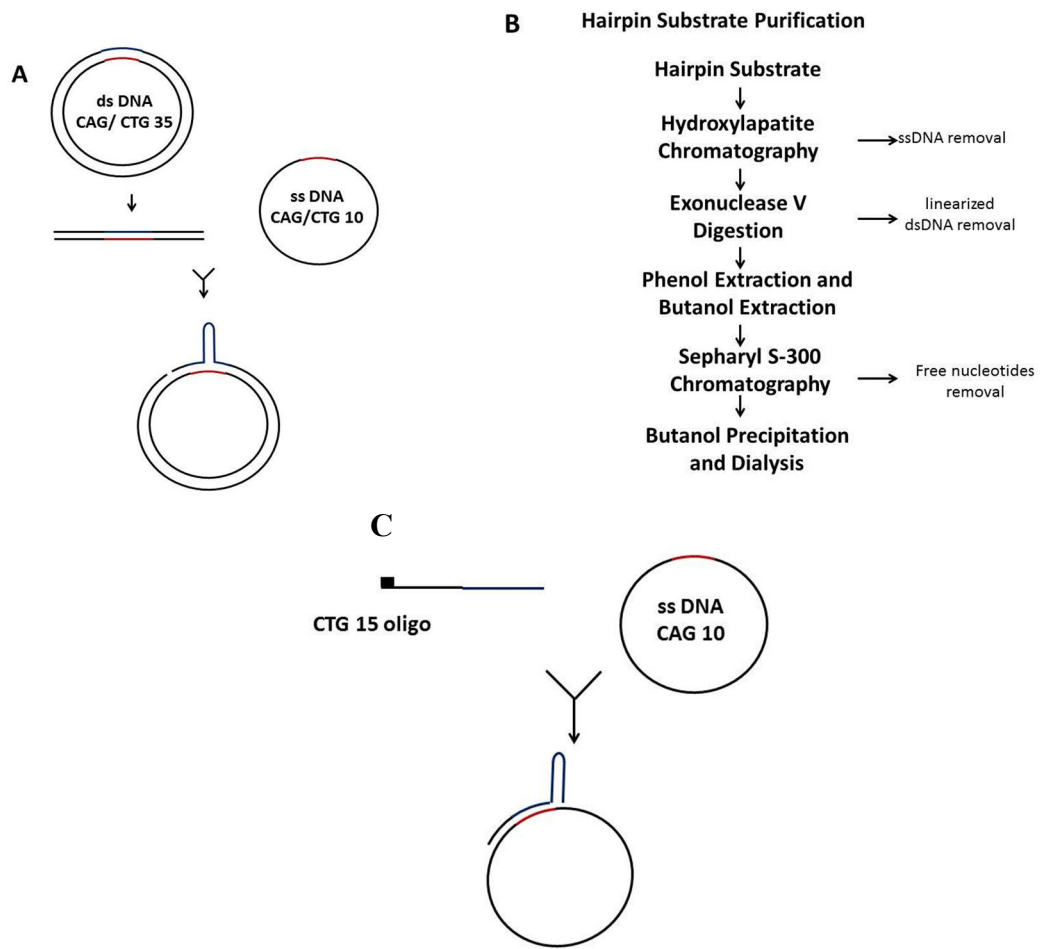
Figure 2.1 Different DNA substrates preparation. A) 5' hairpin substrate was formed by annealing a single-stranded DNA to linearized longer repeat double-stranded DNA. Based on the repeat size difference, a CAG/CTG hairpin was formed. B) Hairpin substrate purification. With the use of column chromatography and enzyme digestion, the hairpin substrate can be purified. C) Oligo-based hairpin substrate preparation. The 5'exonuclease resistant oligonucleotide with 15 CTG repeats was annealed with the single-stranded DNA with 10 CAG repeats to form a hairpin substrate with a CTG hairpin formed at the 3'end.

2.4 *In vitro* hairpin repair assay (or hairpin primer extension assay)

The hairpin repair assay was performed as described in (67). Hairpin substrate (100 ng) was incubated with 130 μg of nuclear extract or proteins in a 40 μL reaction (110 mM KCl, 20 mM Tris-HCl, pH 7.6, 5 mM MgCl$_2$, 1 mM glutathione, 1.5 mM ATP, 0.1 mM of each dNTP, 0.05 mg/mL BSA) at 37 ^{o}C for 30 min. After that, 60 μL proteinase K solution (0.67% SDS, 2.5 mM EDTA, 20 mg/mL proteinase K) was then added to digest the proteins and stop the reaction. An equal amount of phenol solution was added to the reaction mixture with vigorous mixing. After 5 min centrifugation at 13,200 rpm, 80 μL of the aqueous upper layer was extracted. 120 μL of TE buffer was added into the original tube containing phenol solution for back extraction. After phenol extraction, 1/10 volume of 3 M sodium acetate and 10 fold of 100% ethanol were added to the extracted aqueous solution with vigorous mixing. After that, the reaction mixture was kept at -80 ^{o}C for 15 min. The solution was then centrifuged at 14000 rpm for 15 min at 4 ^{o}C. DNA was precipitated at the bottom and the solution was removed followed by addition of 500 μL 70% ethanol. After 5 min centrifugation at 13,200 rpm at room temperature, the solution was removed and the tube containing DNA was dried in vacuum centrifuge for 5 min. The dried DNA was resuspended in 10 μL ddH$_2$O and digested with 0.3 unit of BglI and BsrBI (New England Biolabs) in a 20 μL reaction for two hr. The hairpin repair efficiency was scored by the Southern blot mentioned in section 2.1.5. Quantification was done by Kodak MI SE.

The hairpin extension reaction was basically the same except the hairpin substrate described in section 2.3.3 was used.

2.5 Primer extension assay

2.5.1 Linearized single-stranded DNA preparation for primer extension

Single-stranded DNA preparation was described in section 2.3.1.2. Circular single-stranded DNA (10 μg) was mixed with two digestion primers (pBsrBI and pBstNI, 10 pmole each, see primer list) in a 100 μL reaction containing 1.67M NaCl. The reaction mixture was heated at 95 ^{o}C for 30 min followed by slowly cooling down to room temperature. The annealed DNA was then digested with 10 units of BsrBI at 37 ^{o}C and BstNI at 55 ^{o}C for two hours, respectively. The linearized single-stranded DNA was recovered by phenol extraction, followed by ethanol precipitation (described in section 2.4, Fig. 2.1). The linearized DNA was then dried in a vacuum centrifuge. DNA was dissolved in 100 μL of ddH$_2$O and the concentration was estimated by OD$_{260}$.

2.5.2 Primer extension assay

The synthesis primer (see primer list) was 5' end-labeled with ^{32}P γ-ATP (described in 2.1.7). The annealing reaction was conducted by heating 1 pmole of the labeled primer and 1 μg of linearized template (described in 2.3.4) at 95 ^{o}C for 30 min, followed by slowly cooling the reaction to room temperature in the presence of 167 mM NaCl. The primer-template substrate was then incubated with 60 fmol pol δ in the presence or absence of 0.35 pmole WRN, in a 10 μL reaction buffer containing 20 mM Tris, 7.6, 50 μg/mL BSA, 1 mM glutathione, 5 mM MgCl$_2$, 1.5 mM ATP, 110 mM KCl and 0.2 mM each of the four dNTPs at 37 ^{o}C for 20 min. For the reaction with AMP-PNP, ATP was replaced by 2 mM AMP-PNP. The reaction was terminated by addition of 10

μL 2XSSCP (described in 2.1.4). Reaction products were separated on a 6% sequencing gel and visualized using a phosphorimager (Molecular Dynamics, Inc).

2.6 Helicase assay

Intrastrand CTG hairpin formation was conducted by heating 0.8 pmol of $5'^{32}P$ end labeled (as described in 2.1.7) pCTG15 or pCTG35 (Table 1) at 95 ^{o}C for 15 min in an 80 μL reaction containing 167 mM NaCl, followed by slowly cooling the reaction to room temperature. The helicase reaction was performed by incubating 1 μL (0.08 pmole) of labeled hairpin with 0.06 pmole cold pCAG15 (Table 1) in the presence or absence of WRN in 20 μL reactions containing 20 mM Tris, 7.6, 50 μg/mL BSA, 1 mM glutathione, 5 mM $MgCl_2$, 110 mM KCl and in the presence or absence of 1.5 mM ATP at 37 ^{o}C. The reaction was terminated by addition of 20 μL of 16% sucrose. The reaction products were resolved using 6% native polyacrylamide gel electrophoresis and visualized using a phosphorimager.

2.7 Hairpin removal assay (oligo-based)

The substrate for oligo-based hairpin removal was basically the same as described in section 2.3.3, except the radiolabeled oligonucleotide was used for annealing. HeLa nuclear extract (50 μg) or purified proteins in the hairpin removal buffer (110 mM KCl, 20 mM Tris-HCl, pH 7.6, 5 mM $MgCl_2$, 1 mM glutathione, 1.5 mM ATP, 0.05 mg/mL BSA) and 37.5 μM aphicodicolin were incubated at 37 ^{o}C for 30 min. The reaction products were resolved in a 15% urea denaturing polyacrylamide gel, followed by phosphorimager detection.

2.8 Nucleotide incorporation assay

[32]P radiolabeled hairpin substrate mentioned in section 2.7 was incubated with pol β or δ exonuclease mutant in a 40 μL reaction (110 mM KCl, 20 mM Tris-HCl, pH 7.6, 5 mM MgCl$_2$, 1 mM glutathione, 1.5 mM ATP, 0.1 mM of various combinations of dNTP, 0.05 mg/mL BSA) at 37°C for 15 min. After proteinase K digestion, the reaction products were recovered by phenol extraction, followed by ethanol precipitation. The DNA was then dried and dissolved in 4 μL 1XSSCP. The reaction products were resolved using an 8% sequencing gel, followed by phosphorimager detection.

CHAPTER 3 (CAG)$_n$/(CTG)$_n$ hairpin removal and escape by DNA polymerases during DNA Synthesis

3.1 Introduction

Expansion of CAG/CTG trinucleotide repeat (TNR) causes at least 14 neurological and neurodegenerative disorders, including Huntington disease and myotonic dystrophy (20,82). These diseases have different thresholds for the number of repeats at which the TNR stretch becomes unstable. Once the repeat threshold is passed, the repeats can, in some cases, expand by up to several thousand units (83), which leads to cell dysfunction and degeneration by altering the expression of the affected genes or the function of the affected gene products (84). Patients with longer repeats usually exhibit a decrease in the age of onset of the disease and an increase in its severity (3). However, the mechanisms that promote TNR expansion are poorly understood.

Hairpin formation within CAG/CTG repeats is associated with repeat instability. TNR expansion can occur when a hairpin is formed in the nascent strand during DNA synthesis, while repeat contraction occurs when the hairpin is formed in the template strand (85). *In vitro* studies have revealed that CAG and CTG repeats (as long as more than two) can form very stable hairpin structures (41,86-88). Liu et al. (17) recently demonstrated that the CAG/CTG hairpin formation also occurs *in vivo*, in a manner dependent on DNA replication. Recent biochemical studies have revealed that human cells possess a DNA hairpin repair system that catalyzes error-free removal of CAG/CTG hairpins in a nick-dependent manner (39,67). Regardless of the CAG/CTG hairpin location, the hairpin repair system always targets the nicked (i.e., newly synthesized) DNA strand for incisions, mainly using structure-specific endonucleases, followed by

DNA synthesis using the continuous (parental) strand as a template. However, no DNA removal intermediate was observed in the continuous strand (39,67,89).

CAG/CTG hairpin formation is believed to occur via DNA strand slippage during DNA metabolic processes that introduce single-stranded DNA formation within or near the repeat region (90). These processes include DNA replication (17,23,91), repair (27), and recombination (77,92). The expansion nature of the repeats supports hairpin formations in the nicked nascent strand, e.g., in the case of DNA replication, a CAG/CTG hairpin can easily be formed within a repeat-containing Okazaki fragment, where strand slippage can occur in a 5'to 3' or 3'to 5' orientation. A common feature of these DNA metabolic processes is that they all involve DNA synthesis, a reaction catalyzed by DNA polymerases. However, little is known about the role of DNA polymerases in TNR expansions.

At least 15 mammalian DNA polymerases have been identified (93). These polymerases play distinct roles in genome maintenance, with a few of them functioning in the replication of the genome and the majority of them participating in DNA repair and translesion DNA synthesis (TLS). Replicative DNA polymerases, e.g., polymerase (pol) δ and pol ε, possess a 3'to 5' proofreading exonuclease activity and are essentially error-free. In contrast, DNA polymerases involved in TLS contain no proofreading activity and are highly mutagenic (93). Despite their distinct roles in DNA metabolic processes, recent evidence suggests that these polymerases collaborate to deal with bulky DNA lesions during DNA synthesis in a reaction called polymerase switching (94). When DNA synthesis by replicative DNA polymerases is blocked by a bulky DNA lesion, the low fidelity polymerase will replace the replicative DNA polymerase to bypass the lesion.

After that, the low fidelity polymerase will be switched back to the replicative

polymerase to resume the high-fidelity DNA synthesis. Since CAG/CTG hairpin results

in abnormal DNA structure formation, it is likely that the CAG/CTG hairpin on the

nascent strand is treated as a huge DNA lesion for DNA synthesis. Nonetheless, how

DNA polymerases handle DNA hairpins is still unknown.

In this study, we constructed a series of CTG hairpin substrates that simulate the

hairpin structures in the nascent strand during DNA synthesis, and examined HeLa

nuclear extracts and several DNA polymerases for their ability to process these hairpin

structures. Subthreshold repeat length was used so as to study whether there is factor

promoting the subthreshold-mediated expansion. We demonstrate here that depending on

whether the hairpin contains a 3' tail, pol δ can either remove or retain the hairpin.

However, pol β promote repeat expansion with limited DNA synthesis after the hairpin.

Surprisingly, in the presence of both pol β and pol δ, the hairpin-retained product is ten

times more than that generated by pol β or pol δ alone. This synergistic stimulation

suggests that polymerase switches between pol β or pol δ occur, which promote the

hairpin retention, i.e., TNR expansion.

3.2 Results

3.2.1 Characterization of hairpin substrate

To explore how human cells process a 3' slipped CAG or CTG hairpin during DNA synthesis, DNA hairpin Substrate I (oligo-based substrate, Fig. 3.1A and B) was designed. We employed the T7 endonuclease I and BbvI cutting system as demonstrated by Dr. Pearson's group to determine the structure of our hairpin substrate (16). T7 endonuclease I is specific for cleaving of the hairpin junction (16). Our result showed a single cutting indicating only one hairpin is formed (Fig. 3.2A, lane 1). BbvI is a restriction enzyme recognizing "GCAGC", cutting of a distance of 12 nt away from the restriction site. Since it can recognize "CAG" sequence repeatedly on the double-stranded CAG/CTG repeat, a ladder pattern will be generated in the 5'-^{32}P radiolabeled double stranded DNA with (CAG/CTG) repeat (Fig. 3.2B, lane 1). Theoretically, it will generate a ladder from 11bp to 32bp on a 5' end-labeled (CAG/CTG)$_{10}$ double stranded DNA. However, we just observed a ladder pattern from 11 to 29 on the double-stranded (CAG/CTG)$_{10}$. The double-stranded (CAG/CTG)$_{10}$ was formed by annealing the oligo with the single stranded DNA. It is possible that BbvI will not recognize the very first recognition site because the restriction site was just too close to the single stranded DNA region.

When compared with the double-stranded (CAG/CTG)$_{10}$ BbvI cutting pattern, if there is any hairpin in the doubled strand (CAG/CTG)$_{10}$ region, it will generate different pattern than that of the double-stranded DNA with (CAG/CTG)$_{10}$. Our hairpin substrate showed similar ladder pattern from 11bp to 26bp, indicating the hairpin did not formed at the (CAG)$_{10}$ region. Since the 29bp band is missing, it is suggested that the hairpin is formed at the 3' end (Fig. 3.2B, lane 2).

44
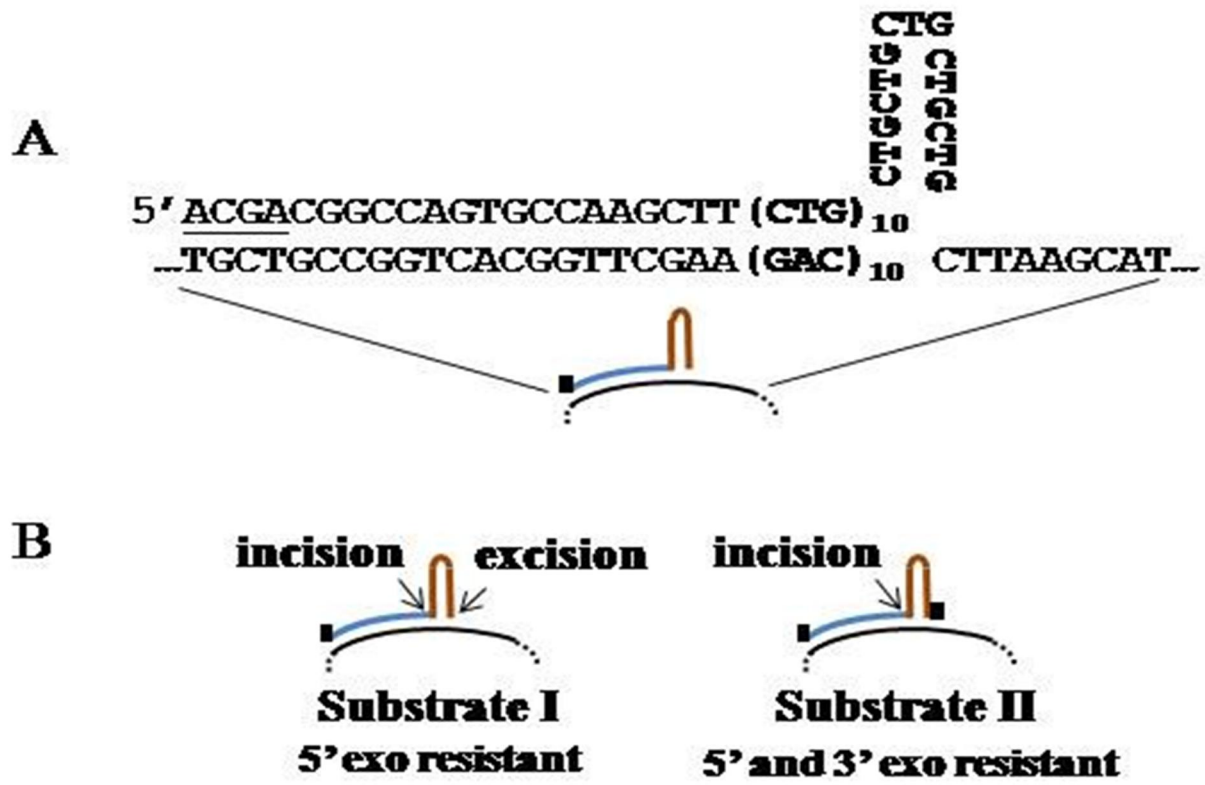
Figure 3.1 Hairpin repair assay. A) The hairpin substrate is modified to be resistant to 5' exonuclease activity (the underlined nucleotides are 2' O-methyl bases linked by phosphothioate bonds). The hairpin substrate was formed by annealing the oligonucleotide containing a $(CTG)_{15}$ repeat to circular single stranded DNA containing a $(CAG)_{10}$ repeat. B) Two different substrates were used in the hairpin repair assay. Substrate I is the hairpin substrate resistant to the 5' exonuclease degradation, while Substrate II is resistant to both 5' and 3' exonuclease degradation.
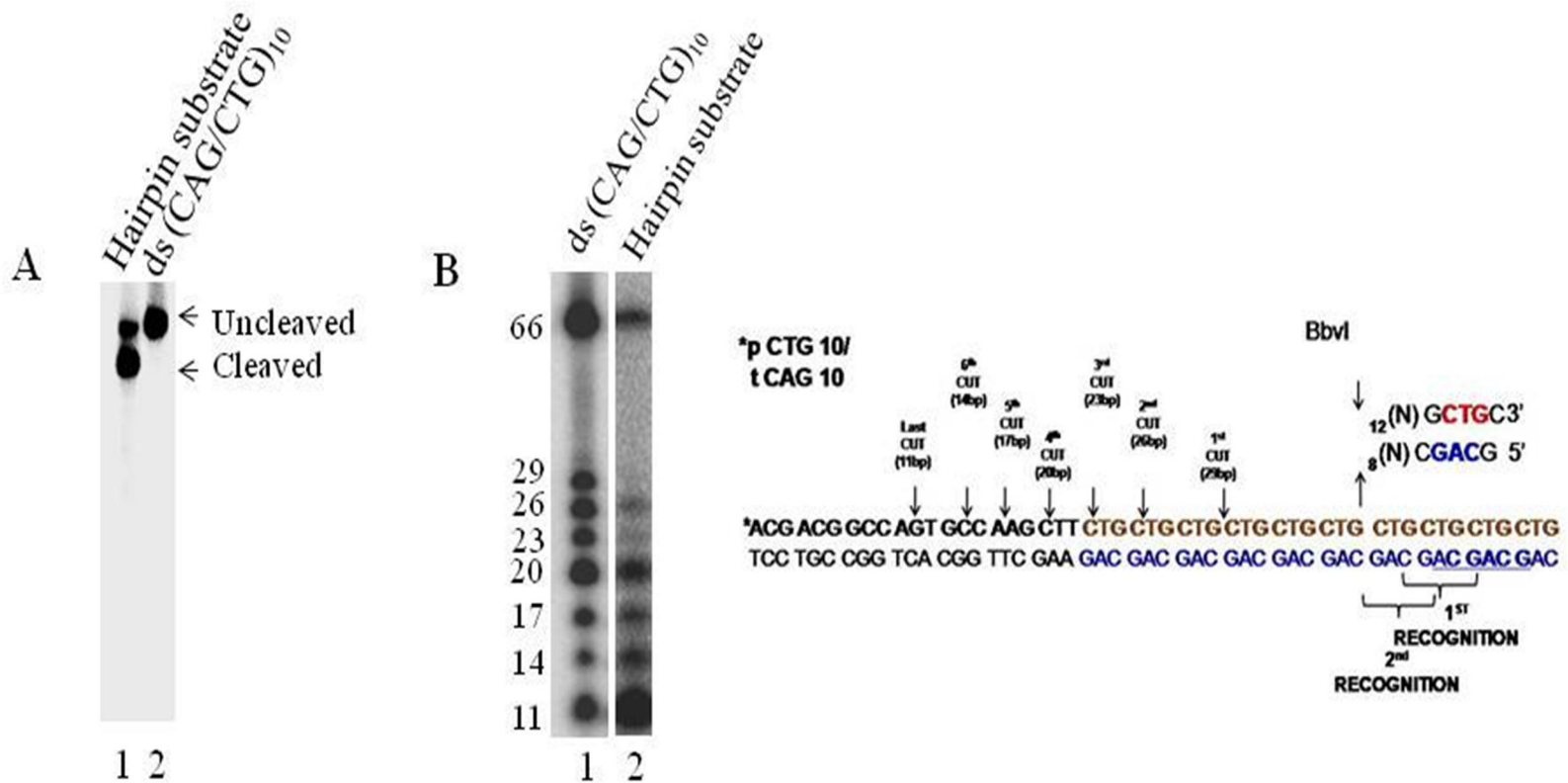
Figure 3.2 Characterization of the (CTG)₅ hairpin. A) T7 endonuclease I digestion confirms our substrate contains a single hairpin. B) BbvI cutting confirms our hairpin substrate form a hairpin at the 3' end. 5' radiolabeled double-stranded DNA or hairpin substrate was incubated with T7 endonuclease I or BbvI at 37 °C for two hr. Results were resolved in a 15 % 8 M urea denaturing gel.

3.2.2 Polymerase β promotes CTG repeat expansion in nuclear extract-catalyzed DNA synthesis.

Unlabeled substrate I was incubated with HeLa nuclear extract under conditions supporting DNA synthesis. The reaction products were detected by Southern blot analysis using a labeled probe specifically recognizing the downstream sequence of the newly synthesized strand (see Fig. 3.3A). When the hairpin is removed, the product will be 15 bp shorter than if it is not. Thus, the hairpin removal efficiency can be monitored by their relative mobility in gel electrophoresis. Incubation of DNA Substrate I with HeLa nuclear extracts yielded a major product and a minor product (Fig. 3.3B, lanes 5 and 6). Because the minor and major products migrated in the positions corresponding to the molecular markers with and without the CTG hairpin (Fig. 3.3B, lanes 7 and 8), respectively, it is determined that the vast majority of the DNA substrate underwent hairpin removal, while a small fraction of the substrate escaped the hairpin removal. Given the fact that the hairpin is located right at the 3' end of the primer and that repeat expansions require DNA synthesis, the observed hairpin escape (retention) is likely due to direct incorporations of nucleotides to the 3' end of the hairpin. To test this possibility, DNA polymerase δ, β, μ, or η was added to HeLa extracts in the hairpin-primer extension reactions. The results show that in comparison with HeLa extract alone, addition of the individual DNA polymerases, except pol η, stimulated the production of the hairpin escaped species. Interestingly, the strongest stimulation occurred in the reaction with pol β (Fig. 3.3C, lane 3), and less with error-prone TLS polymerases. These results suggest that pol β promotes CAG/CTG expansion during DNA synthesis.
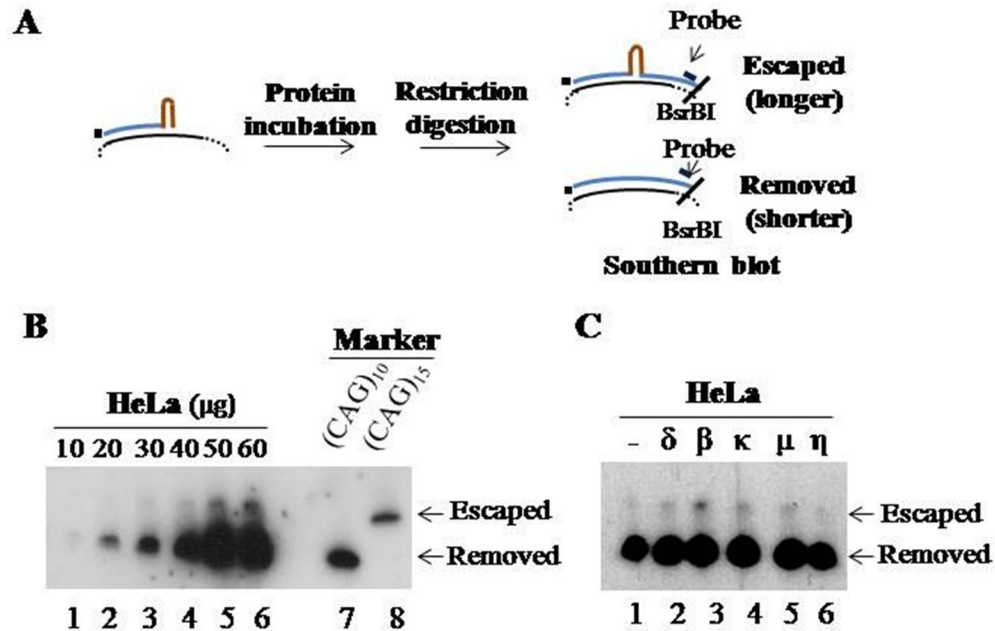
Figure 3.3 Pol β enhance hairpin escaped in the presence of HeLa nuclear extracts. A) Hairpin primer extension assay. Hairpin substrate was incubated with HeLa nuclear extract or purified proteins at 37 °C for 30 min. After BsrBI digestion, repair efficiency was scored by Southern blotting using the probe targeting the downstream region of the newly synthesized strand. Products from which the hairpin was removed (Removed) are shorter than products which retained the hairpin sequence (Escaped).(Detailed method mentioned in 2.3.3) B) High concentration of HeLa nuclear extract promotes escaped repair. Hairpin primer extension reaction was carried out in the increasing amount of HeLa nuclear extract. Markers were generated using the hairpin repair assay procedure using 30 μg of HeLa nuclear extract except the primer did not contain any (CTG) repeats. $(CAG)_{10}$ and $(CAG)_{15}$ represent single-stranded template containing 10 and 15 CAG repeats respectively. C) Pol β stimulates escaped repair in the presence of HeLa nuclear extract. The hairpin repair assay was conducted with the addition of different polymerases to 30 μg of HeLa nuclear extract. Used in the assays were 600 fmol pol δ, 130 fmol pol β, 100 fmol pol κ, 180 fmol pol μ, or 60 fmol pol η each with 150 ng of Substrate I.

3.2.3 Repeat expansion involves concerted actions of pol β and pol δ.

To determine if pol β is responsible for CAG/CTG expansion during DNA synthesis, the hairpin-primer extension reactions were conducted in a purified protein system containing Substrate I, pol β, replication factor C (RFC), and proliferating cellular nuclear antigen (PCNA). As a control, we performed the assay with the high-fidelity pol δ in the presence of RFC and PCNA. Surprisingly, pol δ produced two products that are similar to those found in HeLa extracts, i.e., a major product representing the hairpin removal and a minor product corresponding to the hairpin retention (Fig. 3.4A, lane 1). This result suggests that pol δ is capable of both removing and "tolerating" the $(CTG)_5$ hairpin. When the reaction was carried out in the pol β-containing system, we observed trace amount of the hairpin-escaped product, instead, we saw a major band, which is smaller than the hairpin-removed product (Fig. 3.4A, lane 2). This result appears to be different from what was observed in HeLa extracts supplemented with pol β (see Fig. 3.3C, lane 3).
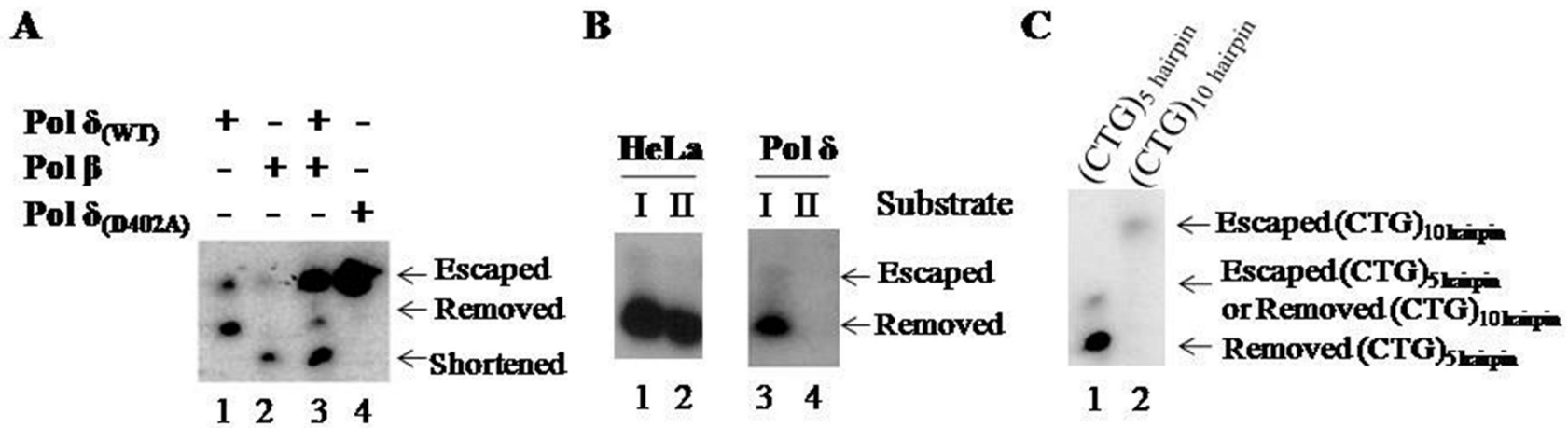
Figure 3.4 Repeat expansion involves concerted actions of pol β and pol δ. A) Pol β and pol δ showed synergetic effect on promoting hairpin escaped while pol δ exonuclease activity is involved in hairpin removal. Pol δ exonuclease mutant (D402A) could only generate the escaped repair product but not the hairpin removed product while pol δ wild-type generated hairpin removal products. B) Pol δ initiates the hairpin repair from the 3' end. Substrate I is resistant to the 5' exonuclease activity while Substrate II is resistant to both 5' and 3' exonuclease activity (Fig. 3.1A). HeLa nuclear extract can remove the hairpin from both 5' and 3' exonuclease resistant substrates, while pol δ could not repair the hairpin substrate with 3' resistant to exonuclease activity. C) Pol δ exonuclease activity can only remove small CTG hairpin. $(CTG)_{10}$ hairpin was generated by annealing the oligonucleotide with 25 CTG repeats against single stranded DNA with 15 CAG repeats. Hairpin primer extension reaction was conducted in the presence of either 30 μg of HeLa nuclear extract or purified proteins (110 fmol RFC, 2 pmol PCNA, 600 fmol pol δ (or pol δ $_{(D402A)}$) at 37 °C for 30 min. Southern blotting was used to detect the products.

The surprising results prompted us to hypothesize that the observed hairpin escape/retention in HeLa extract supplemented with pol β is likely derived from joint efforts by pol β and a highly processive polymerase such as pol δ, where the weakly processive pol β incorporates a few nucleotides to the 3' end of the hairpin and the highly processive polymerase takes over the remaining DNA synthesis, leading to the hairpin escape. This hypothesis was tested in the defined system that contained both pol β and pol δ (Fig. 3.4A, lane 3). Under these conditions, we indeed observed the hairpin-escaped product, which accounts for 75% of the three products detected in the reaction. The other two products appear to be the hairpin-removed one (5%) and the shorter product (20%) specifically generated by pol β alone (see below for description of this product). These observations suggest that pol β and δ together synergistically promote hairpin retention, i.e., repeat expansion during DNA synthesis.

3.2.4 Pol δ removes small hairpin via its proofreading exonuclease activity.

Since pol δ possesses an intrinsic 3'to 5' proofreading nuclease activity (95), we believe that the proofreading activity is responsible for the $(CTG)_5$ hairpin removal. This idea was first tested using a pol δ mutant that contains a D to A substitution at residue 402 (D402A). The substitution inactivates the exonuclease activity but not the polymerase activity of pol δ (95). Unlike the reaction with pol δ, which produced both the hairpin-removed and hairpin-escaped products, the reaction with pol δ D402A generated only the hairpin-escaped species (Fig. 3.4A, lane 4), confirming that the 3'to 5' proofreading activity of pol δ is indeed responsible for the hairpin removal.

To further test the involvement of the pol δ proofreading activity in hairpin removal and to rule out the possibility that the removal is due to an contaminated endonuclease in the pol δ preparation, we incubated wild type pol δ with Substrate II (see Fig. 3.1A and B), which is similar to Substrate I but it contains four 2' O-methyl bases linked with phosphothiolate bonds at 3' end of the hairpin, thereby preventing the substrate from 3' to 5' exonuclease digestion (96). As expected, HeLa nuclear extract can efficiently remove the hairpin structure in Substrate II (Fig. 3.4B, lane 2), possibly by endonuclease involvment. However, incubation of Substrate II with the purified system containing wild type pol δ generated no products (Fig. 3.4B, lane 4), indicating hairpin cannot be removed by pol δ. Taken together, our data shown here support the idea that the pol δ 3' to 5' proofreading exonuclease activity is responsible for the hairpin removal in purified system.

We then examined pol δ for its ability to remove CTG hairpins formed with different number of repeats. Among reactions containing different hairpin substrates, the hairpin-removed product was only observed in the reaction with the $(CTG)_5$ hairpin (Fig. 3.4C, lane 1), suggesting that pol δ cannot remove a larger (or more stable) hairpin. A simple explanation is that CAG/CTG hairpins formed with 10 or more repeats adapt a stable ternary structure that prevents nuclease attacks. It is also noted that in the reactions containing larger hairpin substrates, we did not observe any the hairpin-removed products, and there was no increased production of the hairpin-escaped species (Fig. 3.4C, lanes 1 and 2). This is different from the reaction jointly catalyzed by pol δ and pol β, where reduced hairpin-removal is associated with increased hairpin-escape (Fig. 3.4A, lane 3). These observations strongly suggest that pol δ has a very limited ability to initiate DNA

synthesis using a hairpin as a primer, which could be the role that pol β plays in this reaction.

3.2.5 Pol β initiates DNA synthesis regardless of hairpin and bubble structures in the primer and template strands.

To determine if pol β is capable of initiating DNA synthesis using a hairpin primer, the hairpin synthesis reaction was conducted by incubating ^{32}P-labeled Substrate I, (CTG)$_5$, (at the 5' end of the primer strand) with the pol β synthesis system in the presence of different combinations of dNTPs. The resulting products were analyzed in polyacrylamide gel electrophoresis. The results show that pol β could effectively incorporate correct or sometimes incorrect bases at the 3' end of the hairpin, depending on the availability of nucleotides (Fig. 3.6A). Relatively extensive incorporations by pol β were also seen when 3 or all 4 dNTPs were provided (Fig. 3.6A, lanes 12 to 16). These data strongly indicate that pol β can initiate DNA synthesis using the CTG hairpin as a primer.

The observation of the exclusive hairpin-escaped product in the pol δ-(D402A)-catalyzed system indicates that the pol δ mutant can utilize the hairpin structure as a primer. Indeed, we observed active incorporations of bases at the 3' end of the hairpin by the mutant polymerase (Fig. 3.6B). As expected, pol δ-(D402A) is more processive than pol β, as judged by the fact that more slowly migrating molecules are seen in the pol δ-(D402A) reaction (compare Fig. 3.6A with 4B). Surprisingly, we show that pol δ-(D402A) is less faithful than pol β because the former enzyme incorporates more incorrect nucleotides into the elongation chain (Fig. 3.6B). Under the normal

circumstances, the mis-incorporation triggers its removal by the 3'to 5' proofreading activity of wild type pol δ, which may concomitantly remove its associated 5' sequence (i.e., the hairpin) upon activation, explaining why the majority of the products in the pol δ-catalyzed reaction is the hairpin-removed species.

A product smaller than the hairpin-removed band was seen in the pol β-catalyzed reaction (Fig. 3.4A, lanes 2 and 3). After a series of analyses, we figured out that the product still contained the hairpin structure, but is lacking around 20 nucleotides immediately downstream of the hairpin, as the product could not be detected by a probe targeting the region immediately after the hairpin (Fig. 3.5B). The cause for the shortening is likely due to the fact that a CAG sequence downstream of the CAG repeats in the template strand pairs with the 3' end CTG sequence in the primer (hairpin) strand, leading to the formation of a bubble structure in the template strand (Fig. 3.5A). Since this shorter product was not observed when HeLa nuclear extract (Fig. 3.3C, lane 3), pol δ or pol δ (D402A) (Fig. 3.4A, lanes 1 and 4) was used, it is likely that a bubble structure in the immediate template sequence is highly inhibitory to pol δ mediated extension. Thus, the observation of the shorter product in the pol β-catalyzed reaction suggests that pol β is tolerant to the template bubble structure.
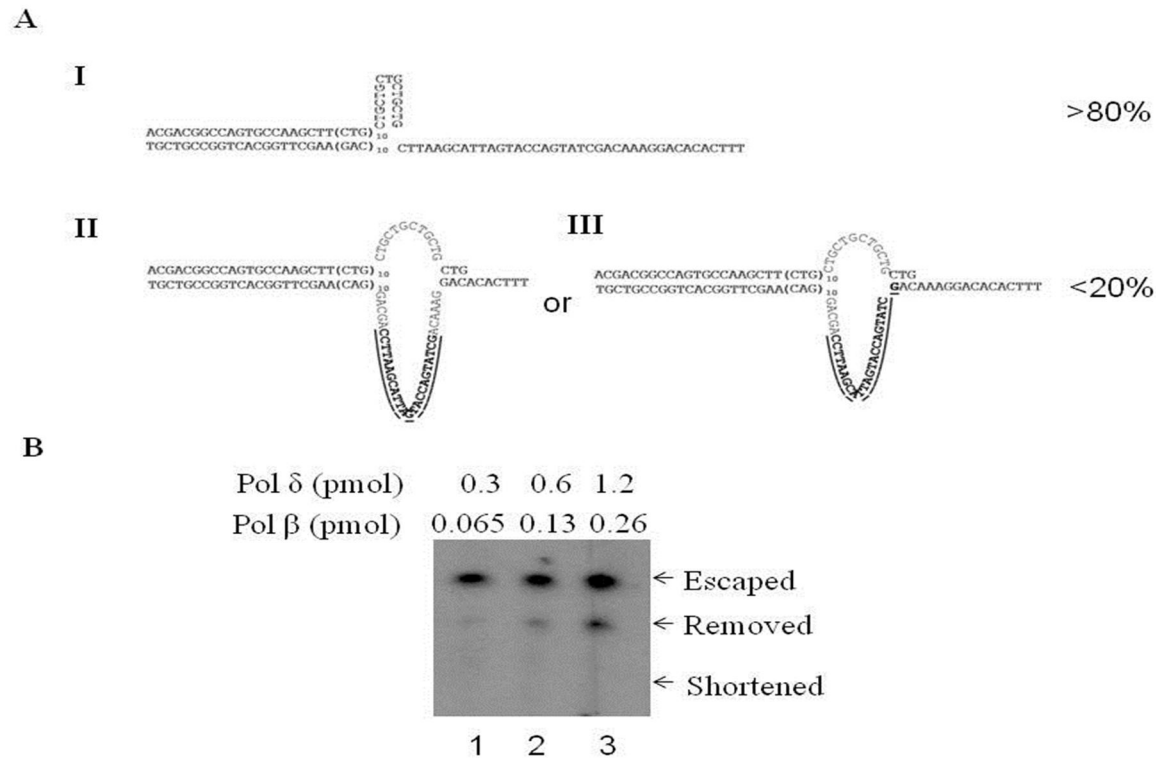
Figure 3.5 Postulated "bubble" structure of the shortened product generated by pol β . A) Prostulated bubble structure. Since there are two downstream CAG sequences in the template which can anneal with the CTG of the hairpin there are two potential structures (II and III) which can form in this manner. Based on the endonuclase cutting and the repair assay results produced by pol β (Fig 3.4A), we can deduce that the majority of the substrate formed the hairpin at the 3' end (Structure I) while less than 20 % of substrate will form a bubble structure (Structure II and III). B) Identification of the shortened band with different probes. Hairpin primer extension was conducted as described in section 2.4. 5'-^{32}P probe targeting the underline bold sequence was used. Our result showed that the shortened band cannot be detected by the probe, suggesting some of the underlined region is missing. Detailed method was described in 2.3.3.
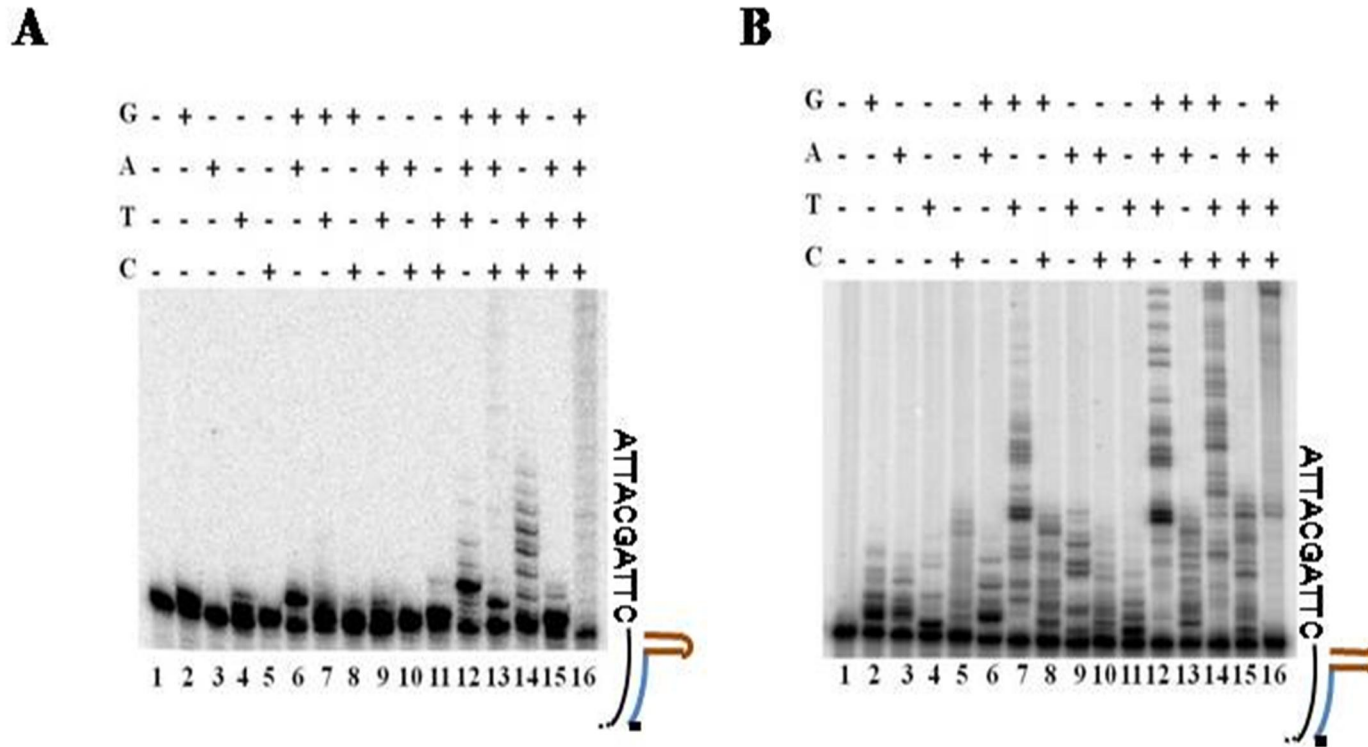
**Figure 3.6 Pol β initiates DNA synthesis regardless of hairpin and bubble structures in the primer and template strands.** Pol β has a higher fidelity of filling in the correct bases during hairpin-mediated expansion than the pol δ exonuclease mutant. ^{32}P end labeled hairpin substrate was incubated with pol β or pol δ mutant in the presence of different combinations of nucleotides at 37 °C for 15 min. The hairpin extension assay was conducted using A) 130 fmol pol β, or B) 600 fmol pol δ exonuclease mutant D402A. Products were resolved on a 10% 8M urea polyacrylamide denaturing gels, followed by phosphorimager detection.

3.2.6 Hairpin removal or retention activity of pol δ depends on the immediate 3' sequence of the hairpin.

To determine the mechanism by which pol δ promotes hairpin escape during DNA synthesis, we tested pol δ ability to process hairpin substrates that carry a different number of bases (2 and 5) with or without a mismatch at the 3' end of the hairpin (see Fig. 3.7A). These substrates mimic the hairpin products generated by pol β as described above or hairpins formed within TNR sequences via strand slippage during DNA synthesis. As expected, processing of Substrate I (i.e., no 3' tail after the hairpin) by pol δ generated the hairpin-removed major product and the hairpin-escaped minor product (Fig. 3.7B, lane 1). Incubation of the same substrate containing a perfectly paired tail of 2 nucleotides at the 3' end of the hairpin rendered the system to switch the ratio of these two products (Fig. 3.7B, lane 2), i.e., the major and minor products are now the hairpin-escaped and hairpin-removed species, respectively. However, when there was a mismatch at the last place of the 2-nucleotide tail (Fig. 3.7B, lane 3), pol δ produced less hairpin-escaped product and more hairpin-removed product as compared with the same substrate without a mismatch at the 3' end (Fig. 3.7B, compare lane 2 with lane 3). These results suggest that the 3' mismatch near the hairpin triggers the hairpin removal by the 3' to 5' proofreading nuclease of pol δ. Interestingly, when the 3' tail reaches 5 nucleotides or more, pol δ could only promote hairpin-escape, regardless of the presence or absence of a mismatch at the 3' end (Fig. 3.7B, lanes 4 and 5), indicating the mismatch location can determine whether pol δ carries out the hairpin removal or hairpin escape.
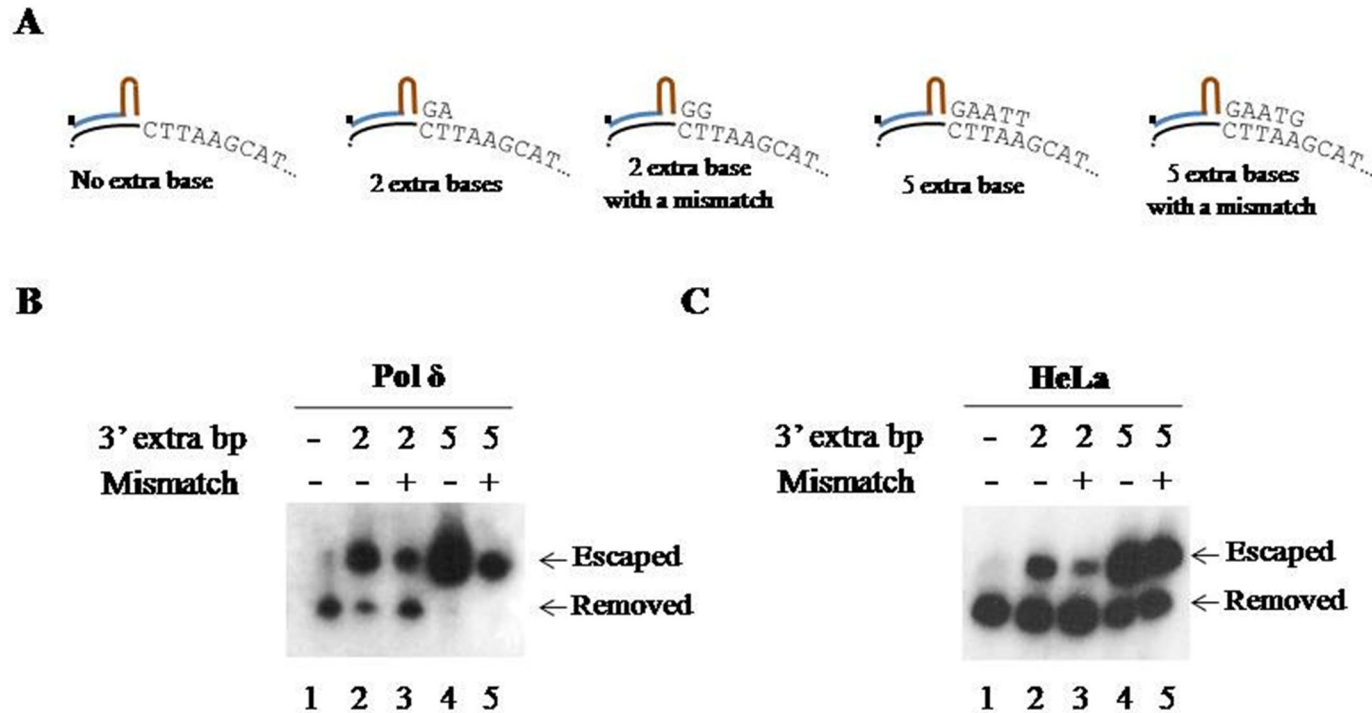
Figure 3.7 Hairpin removal or retention activity of pol δ depends on the immediate 3' sequence of the hairpin. A) Different numbers of perfectly matched sequence were added to the 3' end of (CTG)$_5$ hairpin. B) Two extra bases with no mismatch can induce hairpin escaped by pol δ. C) HeLa has higher ability to promote repair escape when the 3' extra sequence is longer. The hairpin primer extension assay was conducted at 37 °C for 30 min with B) 600 fmol pol δ, 110 fmol RFC, 2 pmol PCNA or C) 30 μg HeLa nuclear extract. Southern blotting was used to visualize the results.

Similar analyses were performed using HeLa nuclear extracts (Fig. 3.7C). Compared with products generated from the pol δ system, a striking difference is that HeLa extracts generated more hairpin-removed products. This is likely due to the endonucleolytic removal of the hairpin by the DNA hairpin repair system in HeLa extracts as described previously (67). In reactions with five perfectly matched nucleotides after the hairpin, no hairpin-removed products were detected in the pol δ system. However, it is noted that not all hairpins were removed by the repair system, as there is still significant amount of hairpin-escaped products in each reaction (Fig. 3.7C).

3.3 Discussion

As a major contributor to CAG/CTG repeat expansion, hairpin formation within the repeats is associated with DNA replication (97) and repair (47). Because DNA expansion requires DNA synthesis, DNA polymerases must play a major role in this process. However, little is known about the mechanism by which DNA polymerases promote CAG/CTG expansion. In this study, we provide strong evidence that DNA polymerases can remove or retain a CAG/CTG hairpin formed in the nascent DNA strand during DNA synthesis, depending on the hairpin structure and the DNA polymerases involved in the synthesis reaction.

Several surprising findings were made in this study. First, we show that the 3'- 5' exonuclease activity of pol δ is capable of removing the hairpin primer if it contains no complementary 3' sequences (Fig. 3.4). Second, among DNA polymerases tested, the polymerase involved in base excision repair, pol β, is the most active enzyme to promote hairpin retained when added to nuclear extracts (Fig. 3.3C). However, we failed to detect the full-length hairpin-escaped product in the purified system catalyzed by pol β alone (Fig. 3.4A), suggesting involvement of additional factor(s) in promoting the hairpin expansion. This additional factor was found to be pol δ, as the hairpin-escaped species is the dominant product when purified pol δ was included in the pol β system (Fig. 3.4A), indicating that the hairpin escape is catalyzed by a collaborative effort of these two polymerases.

Although the mechanism by which pol δ and pol β collaborate to promote TNR expansion requires further investigatios, it is possible that the hairpin escape by pol β and

pol δ involves polymerase switching, a concept originally established for the template strand lesion bypass by TLS polymerases (94), and recently adapted to propose possible TNR expansion (27,97). A model describing the switch between pol β and pol δ is depicted in Fig. 3.8. We hypothesized that a hairpin formation through 3' slippage in the nascent strand blocks pol δ polymerization activity, which leads to the recruitment of pol β to the site for translesion synthesis that adds several nucleotides to the 3' end of the hairpin, followed by re-recruitment of pol δ to DNA synthesis with high fidelity and efficiency. We believe that it is the polymerase switch that leads to the hairpin escape and repeat expansion. This model requires that (i) pol β is capable of adding nucleotides to the hairpin primer; and (ii) pol δ can use pol β-generated products for DNA synthesis, but not excision. Our experiments shown in Figs. 3.6 and 3.7 confirm these are indeed the cases.

Pol β, a key enzyme in base excision repair, exhibits a very high overall mutation rate *in vitro* (98,99). The high error rate of pol β has been attributed to its lack of an intrinsic proofreading exonuclease activity (100). Previous studies have shown that pol β can bypass a number of different types of lesions in the template strand, including cisplatin adducts, cyclobutane pyrimidine dimers and 6-4 photoproducts (101). Our data shown here reveals that pol β can also "extend" a $(CAG)_n/(CTG)_n$ hairpin lesion, and that the "hairpin tolerance" occurs in the primer strand and leads to expansion of the repeats. We also show that pol β can catalyze hairpin extension even though the template strand contains a bubble near the site of the synthesis (Fig. 3.5). These observations suggest that pol β can tolerate a variety of unusual DNA lesions/structures in both the template and primer strands during DNA synthesis. However, it appears that the pol β promotes

CAG/CTG repeat instability by cooperating with a highly processive polymerase, such as pol δ.

We demonstrate that pol δ can be either error-free or error-prone when using a $(CAG)_n/(CTG)_n$ hairpin primer for DNA synthesis. The error-free processing relies on the pol δ 3'to 5' proofreading activity for the hairpin removal, as a pol δ mutant defective in the proofreading activity fails to remove the hairpin (Fig. 3.4A). For pol δ to conduct the error-free processing, the hairpin primer must contain no additional 3' sequence or a 3' sequence with 1 or 2 mispaired nucleotides. However, when pol δ encounters a hairpin primer carrying a complementary 3' sequence with 2 or more correct nucleotides, the polymerase no longer uses its proofreading activity to remove the hairpin, but incorporates nucleotides to the 3' ends, leading to an error-prone processing and potential CAG/CTG repeat expansion if the escaped hairpin is not removed by the hairpin repair system (Fig. 3.8). These 3' tail-containing hairpin primers can be derived either directly from 3' strand slippage within CAG/CTG repeats that carry several complementary nucleotides at the 3' end or from pol β-processed hairpins after polymerase switching (see Fig. 3.8). Therefore, whether or not pol δ conducts error-free or error-prone synthesis to a CAG/CTG hairpin primer is dependent on the hairpin structure.
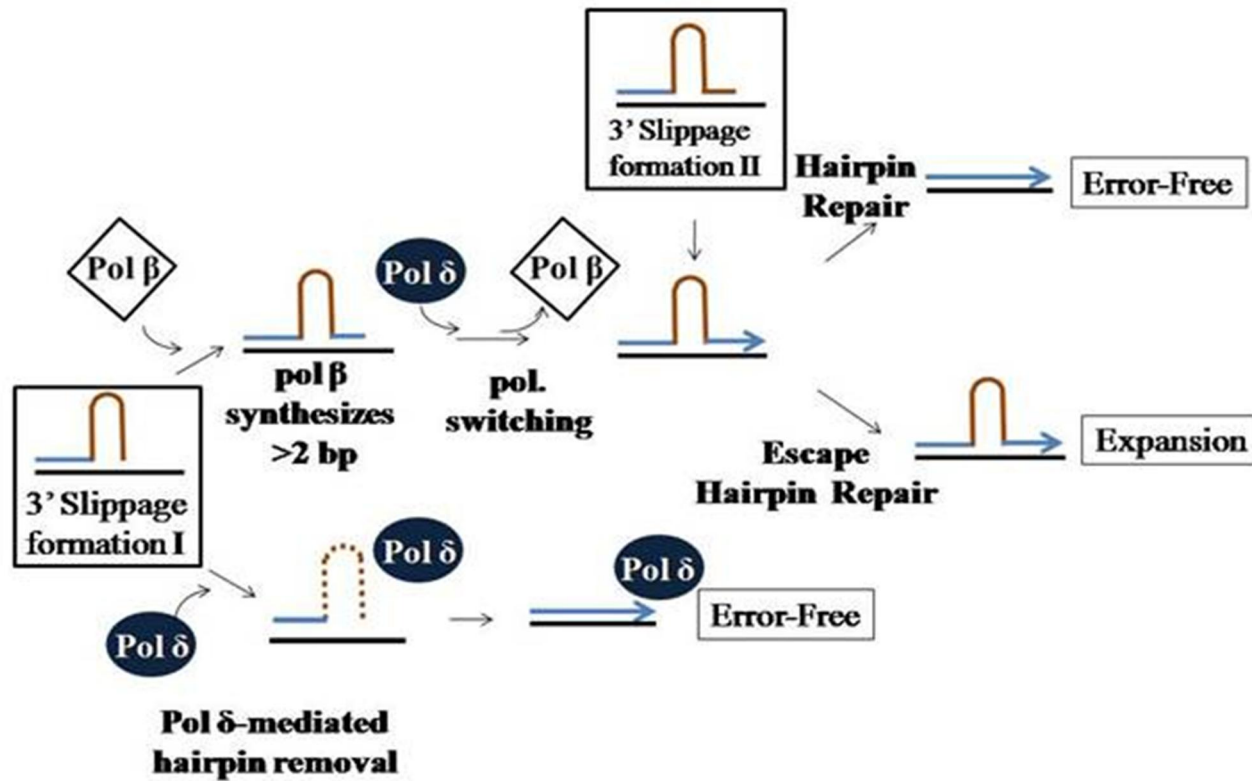
Figure 3.8 Proposed model for polymerase switching in hairpin repair and expansion. When pol δ binds to the 3' slippage hairpin, it can remove the hairpin using its 3'to 5' exonuclease activity.Expansion is promoted by pol incorporating at least 2 correct bases after the 3' slippage hairpin, followed by the extension step by pol δ. The CAG/CTG hairpin repair pathway can act as a secondary system to maintain repeat stability by removing the extended hairpin caused by pol β. Failure of repairing the hairpin on the nascent strand results in repeat expansion.

Although pol β exhibits the most potent hairpin-escape activity among polymerases examined, several TLS polymerases also show more or less hairpin-escape activity (Fig. 3.3C). In addition, we have not tested many other low fidelity TLS polymerases, which may also facilitate CAG/CTG expansions by participating in polymerase switching and incorporating nucleotides to the hairpin primer. Given the fact that mammalian cells possess a number of low fidelity DNA polymerases, including pol β and at least 15 different TLS polymerases, the propensity for a 3' slipped hairpin to escape during DNA synthesis is very high, especially when one or more such low fidelity polymerases are overexpressed. Therefore, the 3' slippage-formed hairpin and its subsequent processing by pol β (or a TLS polymerase) and pol δ may represent a major source for CAG/CTG repeat expansion.

In summary, we have identified a novel mechanism for CAG/CTG repeat expansion, which likely involves a hairpin formation within the repeat units via 3' slippage in the nascent DNA strand, DNA polymerase switching from pol δ to pol β, translesion synthesis by pol β, and polymerase re-switching from pol β to δ for high-fidelity synthesis. This process allows the hairpin to be retained, leading to CAG/CTG expansion. However, further studies are required to confirm the mechanism in vivo, particularly in diseases caused by CAG/CTG repeat expansions.

CHAPTER 4 Werner Syndrome Protein (WRN) resolves CAG/CTG hairpins

4.1 Introduction

CAG/CTG repeat instability is associated with many neurological diseases, including Huntington's disease, Spinocerebellar ataxia type 7 (SCA7) and Myotonic dystrophy type 1 (3,102). CAG/CTG repeats are polymorphic (103,104), meaning that in the genome of different individuals the number of trinucleotide repeats (repeat number) present in a particular stretch of DNA varies. Healthy individuals have a relatively stable CAG/CTG repeat number which changes little during their lifetime. However, in instances where the repeat number increases to a certain threshold, which varies depending on the disease involved, an increase in the number of trinucleotide repeats (repeat expansion) present in a stretch of DNA occurs over time, triggering the development of disease symptoms (3). The expanded repeat may result in cell death due to DNA metabolism interference, toxic protein accumulations or gene malfunction (105). Once the symptoms have developed, nothing can stop the disease progression. Therefore, preventing the repeat regions from expanding is the key to tackling repeat instability associated disorders.

Hairpin formation during DNA replication is one of the favored mechanisms for triggering CAG/CTG repeat instability (8). CAG/CTG repeat regions can hinder DNA polymerase progression during replication if an intra-strand hairpin structure in the repeat region forms (106). Depending on which strand the hairpin forms, DNA replication on CAG/CTG repeat can lead to repeat expansion or contraction. Bacterial (9) and yeast (10) studies reveal that DNA replication of CAG/CTG repeat regions can be error prone resulting in repeat instability. It is also shown that a high frequency of repeat contraction

65

occurs when the CTG repeat-formed hairpin is used as a template for DNA synthesis during DNA repair and/or replication (21). Since DNA repair is still active in non-dividing cells (30), errors from DNA repair can explain repeat instability in non-dividing cells.

The common point of all the DNA metabolic processes triggering CAG/CTG repeat instability is the hairpin formation, therefore, hairpin removal/resolution can prevent repeat instability. Recent studies have revealed that human cells possess a hairpin repair (HPR) system that catalyzes error-free removal of CAG/CTG hairpins in a nick-dependent manner (67). Interestingly, regardless of the strand location of the CAG/CTG hairpins, the HPR system always targets the nicked (i.e., nascent) DNA strand for incisions, mainly using structure-specific endonucleases (67). If the hairpin is located in the nicked strand, the repair system removes the hairpin either by making dual incisions flanking the heterology or by a combination of nick-directed excision and flap endonucleolytic cleavage, which leaves a small single-strand gap. If the hairpin is located in the template strand, incisions occur opposite the hairpin, followed by hairpin unwinding, which generates a relatively large single-strand gap. In either case, the gap is filled by replicative DNA polymerases using the continuous strand as a template (67).As a result, the HPR system ensures TNR stability (See Fig. 4.1).
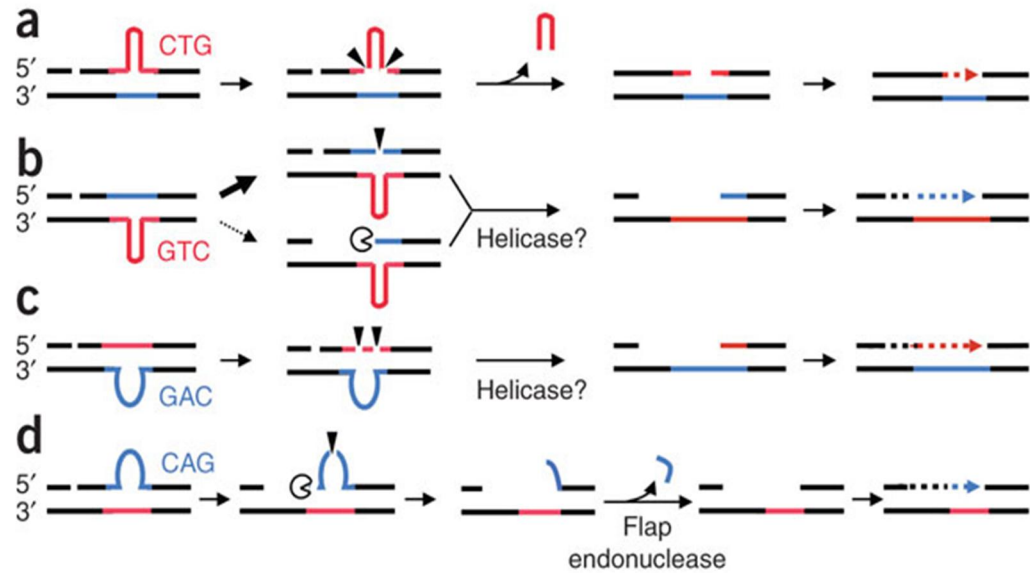
Figure 4.1 Proposed model for hairpin repair. Hairpin repair can be divided into DNA removal and synthesis. DNA removal can be done by excision and incision on the nicked strand. DNA synthesis occurs using the continuous strand as a template. Published data from Hou C, **Chan NL**, Gu L, & Li GM (2009) Incision-dependent and error-free repair of (CAG)(n)/(CTG)(n) hairpins in human cell extracts *Nat Struct Mol Biol* 16(8):869-875.

Interestingly, low repair efficiency was observed when a CTG hairpin was used as a template for re-synthesis during hairpin repair (Fig. 4.2) (67). Since CTG repeats form a more stable hairpin than CAG repeats because of the stacking energy (15,16), it is hypothesized that the low repair efficiency of the CTG hairpin is due to polymerase impediment by the CTG repeat-formed secondary structure. Therefore, a CAG/CTG hairpin-unwinding helicase should enhance the hairpin repair activity (67).

Substrate	ssDNA	Linearized dsDNA	Hairpin description
C-(CTG) (Slip-out)	(CAG)$_{10}$	(CTG)$_{35}$	25 **CTG** hairpin on **nicked** strand
C-(CAG) (Slip-out)	(CTG)$_{10}$	(CAG)$_{35}$	25 **CAG** hairpin on **nicked** strand
V-(CAG) (Slip-in)	(CAG)$_{35}$	(CTG)$_{10}$	25 **CAG** hairpin on **continuous** strand
V-(CTG) (Slip-in)	(CTG)$_{35}$	(CAG)$_{10}$	25 **CTG** hairpin on **continuous** strand

Table 4.1 Hairpin substrate description. The repeat difference of the single- and double-stranded DNA results in different CAG/CTG hairpin formation.
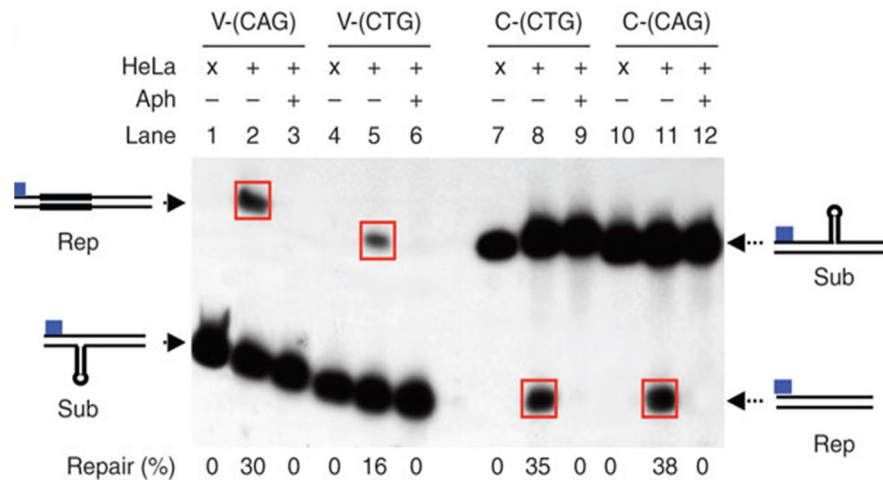
Figure 4.2 Low hairpin repair activity was observed in CTG slip-in substrate. Low hairpin repair activity was observed when CTG hairpin was formed at the continuous strand (CTG slip-in). The repaired bands are shown as bracket. Aph: Aphilidocolin, a DNA polymerase inhibitor. Detailed methods mentioned in 2.3.2. Published data from Hou C, **Chan NL**, Gu L, & Li GM (2009) Incision-dependent and error-free repair of (CAG)(n)/(CTG)(n) hairpins in human cell extracts *Nat Struct Mol Biol* 16(8):869-875.

Indeed, previous studies have implicated DNA helicases in maintaining TNR stability, presumably by resolving hairpins. Deletion of the SRS2 helicase from a yeast strain resulted in CAG/CTG repeat instability. *In vitro* studies reveal that SRS2 has high activity and specificity for unwinding CAG/CTG repeats (69,70). In addition, deletion of SGS1 can cause repeat contraction when CTG is used as the template for lagging strand synthesis, suggesting SGS1 can also unwind CTG hairpins in the template strand during DNA synthesis (77). However, no human homolog of SRS2 has been identified (107), while WRN and BLM are the human homologs of SGS1 (108). Identifying human DNA helicases that can resolve CAG/CTG hairpins will provide molecular basis to elucidate the HPR pathway, as well as the etiology of diseases caused by CAG/CTG repeat instability.

To test the hypothesis that a DNA helicase is involved in HPR, we conducted the *in vitro* HPR assay to screen for factors from HeLa nuclear extract that stimulate the repair of a CTG hairpin formed in the template strand (CTG slip-in). This analysis identified the WRN helicase as one of such factors. A primer extension assay confirmed that WRN stimulated polymerase δ (pol δ) synthesis on a CAG/CTG repeat template, while a helicase assay demonstrated that WRN resolved CTG hairpins. Our results suggest that WRN helps maintain CAG/CTG repeat stability by promoting DNA synthesis by resolving CAG/CTG repeat hairpins in the template DNA strand.

4.2 Results

4.2.1 A partially purified HeLa activity stimulates CTG hairpin repair

To identify factors that stimulate the repair of CTG slip-in hairpin substrate (CTG hairpins formed in the template strand), we screened HeLa nuclear activities on a phosphocellulose p-11 column as previously described (78,89). Results showed that p-11 fractions 39 to 47 stimulated the CTG hairpin repair (Fig. 4.3A, lanes 3 to 5). Western blotting was conducted to determine the presence of several known DNA repair activities, including DNA helicases and DNA polymerases, in these fractions. Interestingly, we found that the distribution of WRN correlated very well as the stimulation activity in these fractions (Fig. 4.3B, lanes 3 to 6). To confirm WRN's role in stimulating the hairpin repair, recombinant 0.35 pmol WRN was added into the reaction containing 50 μg of HeLa nuclear extract. Indeed, addition of recombinant WRN could stimulate the hairpin repair assay (Fig. 4.3C, lanes 2 and 3). Since fraction 47 also showed stimulation, it is possible that other stimulatory factors are also involved in stimulating the CTG hairpin repair.
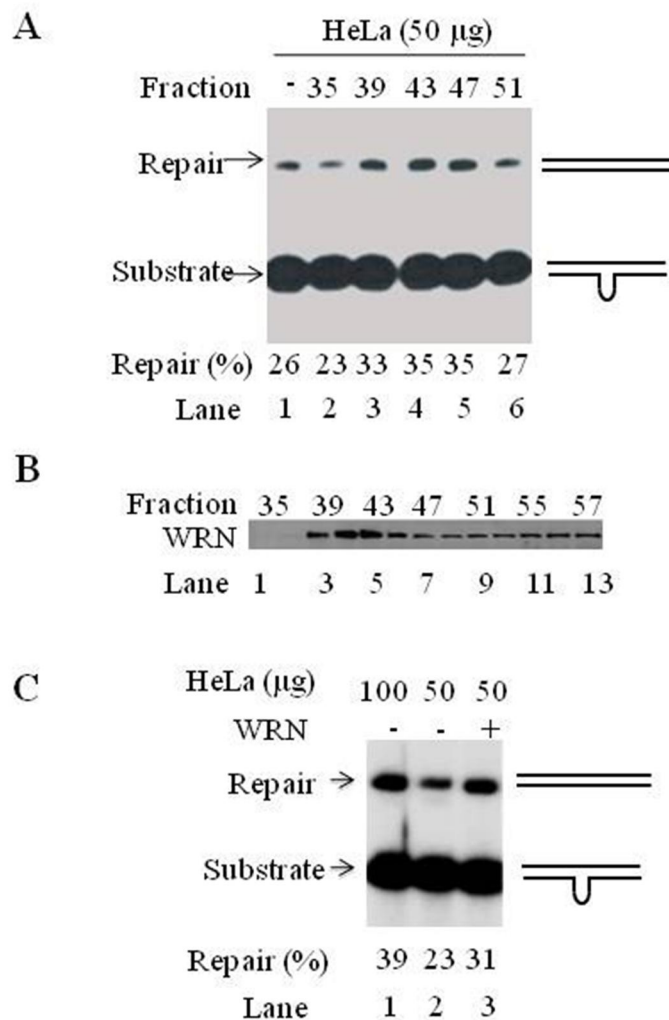
Figure 4.3 Identification of WRN in stimulating hairpin repair on CTG slip-in substrate. A) Repair of the $(CTG)_{35}$ slip-in substrate could be stimulated by some p-11 fractions. Hairpin repair reaction (double-stranded hairpin substrate, described in 2.3.2) was conducted in the presence of 50 μg HeLa nuclear extract and p-11 fractions. Since the repaired product is longer than that of the substrate, the repaired product appeared in the higher position than that of the substrate. B) Distribution of WRN in p-11 column fractions. WRN has peak distribution from fractions 39 to 43 in the p-11 column. Western blot was conducted using antibody against WRN. C) Purified WRN protein (0.4 pmol) can stimulate repair of the $(CTG)_{35}$ slip-in substrate. Hairpin repair was conducted at 37 ^{o}C for 30 min. Southern blot was used for scoring the hairpin repair efficiency

4.2.2 WRN promotes pol δ-catalyzed DNA synthesis on different CAG/ CTG templates

It has been previously shown that WRN interacts with pol δ (74). Based on this known interaction, we hypothesize that WRN stimulates the CTG slip-in hairpin repair by promoting DNA synthesis. To test this hypothesis, we conducted a primer extension assay using a single stranded DNA containing 35 CTG repeats as a template ($(CTG)_{35}$) (Fig. 4.4A). Our results show that WRN stimulates DNA synthesis on $(CTG)_{35}$ templates (Fig. 4.4B, lanes 1 to 4). Since WRN belongs to the RecQ helicase family (109), two other members of this family, RecQ1 and BLM, were also tested in the primer extension assay. Unlike WRN, RecQ1 and BLM could not stimulate DNA synthesis on $(CTG)_{35}$ template (Fig. 4.4B, lanes 5 to 12, Fig. 4.4C). In addition, we showed that WRN also stimulated pol δ-catalyzed DNA synthesis in the presence of PCNA and RFC, two important accessory factors of DNA polymerases (Fig. 4.6B, lanes 3 and 6).

In order to see whether WRN can stimulate various length of CAG/CTG repeat, we conducted the primer extension with different repeat length of CAG/CTG. Our results showed that WRN indeed can enhance all the CAG/CTG repeat synthesis (Fig. 4.6A). These observations suggest that WRN has higher substrate specificity for CTG hairpin than other RecQ helicases.

A

B stNI   B srBI

Primer annealing
and restriction digestion

Linearized SS CAG/CTG template

^{32}P -> ^{32}P labeled synthesis primer
annealing

CAG/CTG

^{32}P Protein incubation

CAG/CTG

B

WRN   RecQ1   BLM
(pmol)   (pmol)   (fmol)

0 0.1 0.2 0.5   0 0.1 0.2 0.5   0 30 60 125

*
Full length
$(CTG)_{35}$

1 2 3 4   5 6 7 8   9 10 11 12
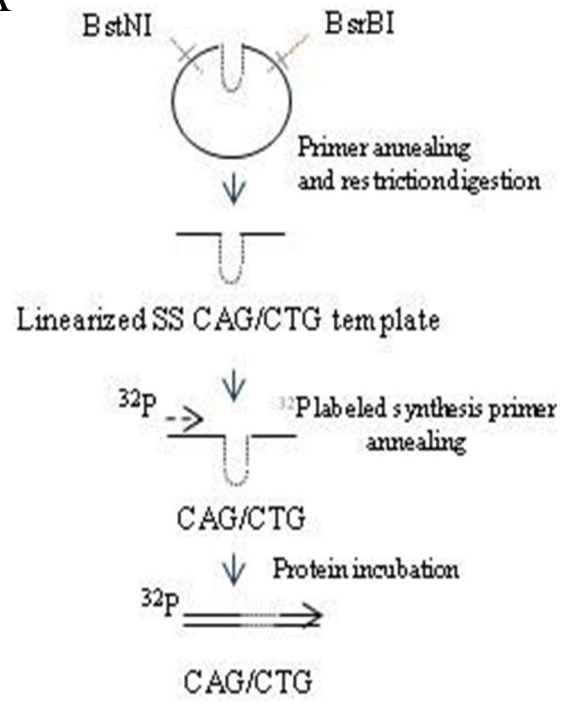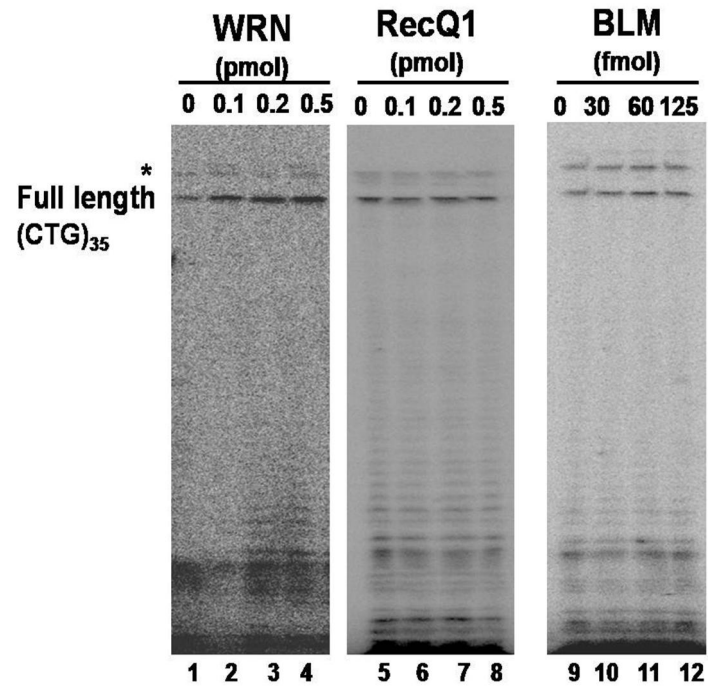
C
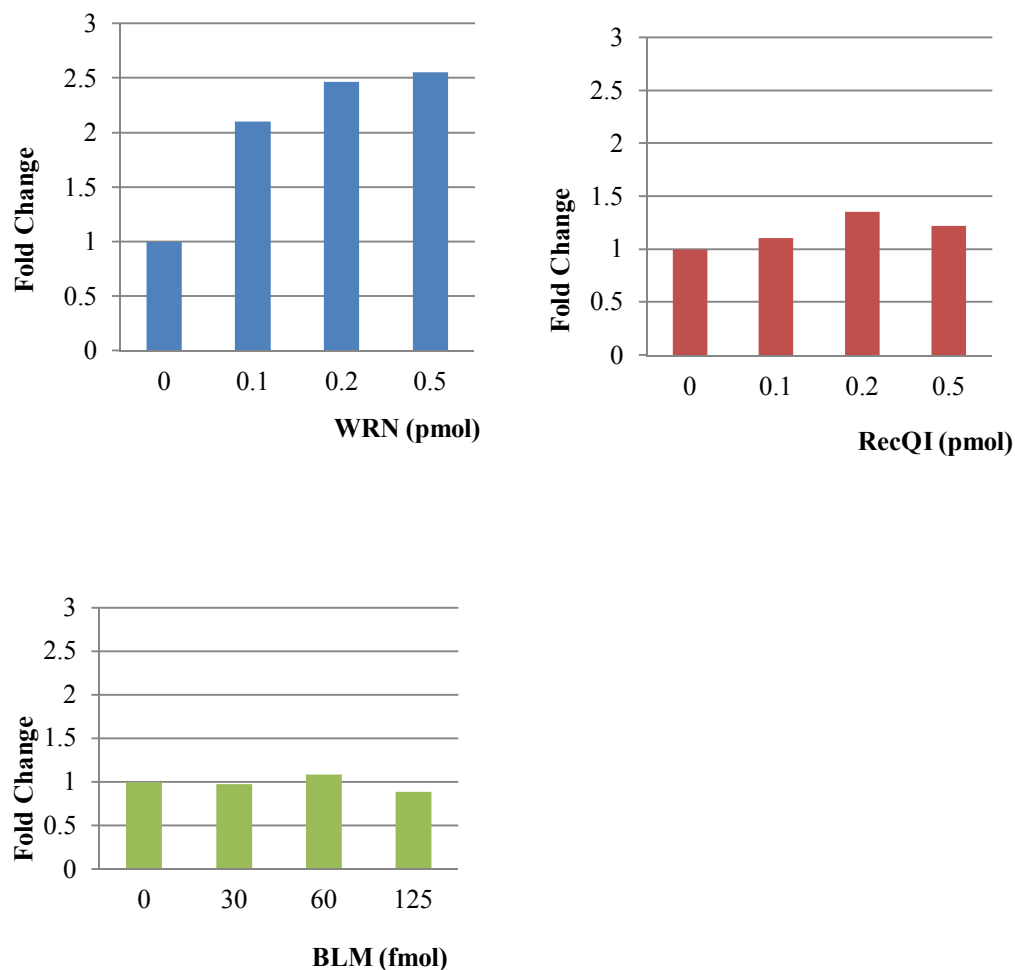
Figure 4.4 WRN stimulated (CTG)$_{35}$ synthesis. A) Primer extension assay. Detailed method is mentioned in Section 2.5. Briefly, 5'-^{32}P labeled primed substrate was incubated with pol δ (12 ng), with or without DNA helicases, at 37 $^{\circ}$C for 20 min. Reaction products were resolved on a 6% 8 M urea denaturing polyacrylamide gel followed by phosphorimager detection. B) Only WRN can enhance DNA synthesis on (CTG)$_{35}$ templates. Full length products are indicated by arrows. Faint bands longer than the full length products (labeled with an asterisk) also appeared in the synthesis. The longer DNA products were confirmed to be non-specific products (refer to Fig. 4.5). C) Quantification of the (CTG)$_{35}$ synthesis results. Quantification was done with Kodak MI SE.

Initially, we conducted the primer extension on a circular DNA as previously described (110,111). Since WRN translocates on single-stranded DNA in a 3'to 5' direction (112), WRN can bind to the upstream single-stranded region on the template and dissociate the primer:template complex. Therefore, removing the upstream single-stranded region can allow us to have more accurate evaluation of the helicase activity. With the limited restriction sites available, BstNI, a restriction enzyme with low specificity was used to remove the upstream single-stranded DNA region. It is noted that a DNA product longer than the template appeared in all the reactions (labeled with an asterisk in the figures). Southern blotting using a $(CTG)_{10}$ probe was conducted to confirm that the longer DNA did not contain any CAG repeats (Fig. 4.5). The longer DNA product only appeared in the presence of BstNI (5'cutting). It is still unclear how the upper band was formed. It may be due to the non-specific BstNI cutting generates a double-stranded primer and template complex.
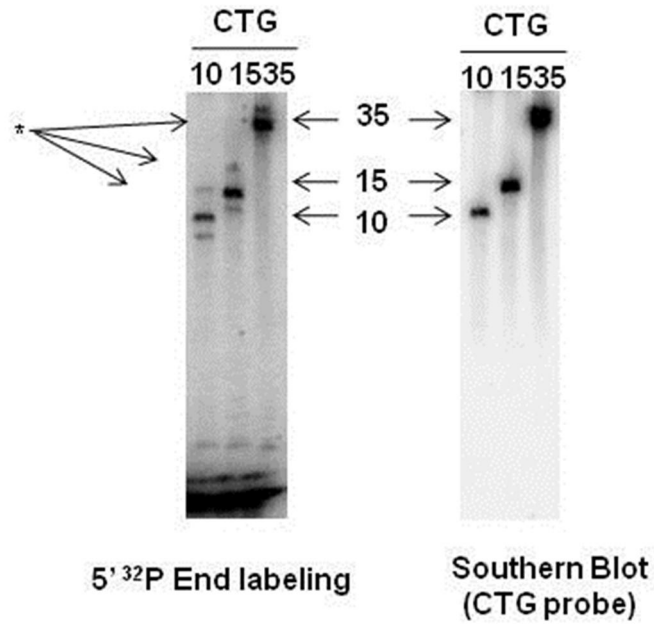
Figure 4.5 Identification of the longer product in the primer extension assay. Left panel: The primer extension assay was conducted with 5'-^{32}P end labeled primer. The longer products (upper bands) are indicated by arrows and an asterisk. Right panel, same primer extension was conducted using unlabeled primer. Results were visualized by Southern blotting using a radiolabeled $(CTG)_{10}$ probe. The longer product visualized in the reaction using the labeled primer cannot be detected in the Southern blot, indicating it does not contain any CAG repeats.
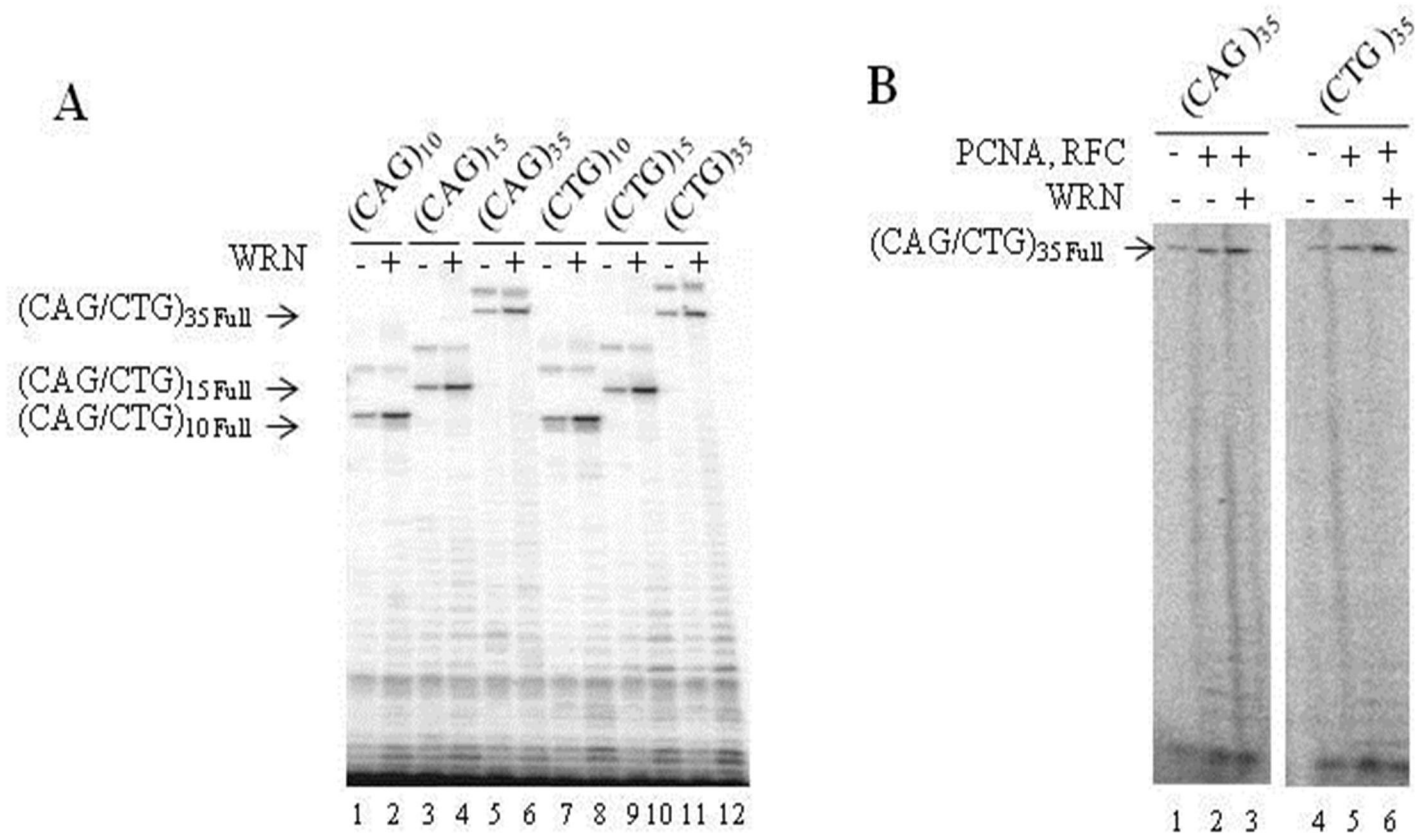
Figure 4.6 WRN stimulates DNA synthesis on different CAG/CTG repeats even in the presence of accessory factors. A) WRN stimulates DNA synthesis on templates containing different repeat length of CAG and CTG. B) WRN stimulate $(CAG)_{35}$ and $(CTG)_{35}$ synthesis even in the presence of accessory factors. Reaction was conducted in the presence of 2 pmol PCNA and 110 fmol RFC without 0.35 pmol WRN on both $(CAG)_{35}$ and $(CTG)_{35}$ templates. Addition of WRN to the reactions containing PCNA and RFC further enhanced the synthesis of DNA on both templates.

4.2.3 WRN helicase activity is responsible for stimulating CAG/CTG repeat synthesis

WRN possesses both a 3'to 5' helicase activity and a 3'to 5' exonuclease activity (113). To determine if one or both activities are required for stimulating CAG/CTG repeat synthesis, we conducted the primer extension assay in the presence of non-hydrolysable ATP (adenylyl-imidodiphosphate, AMP-PNP). WRN is capable of utilizing ATP, dATP, CTP and dCTP as an energy source for helicase activity (112), so it is impossible to conduct the primer extension assay in the absence of an energy source for WRN. Therefore, a tenfold excess of AMP-PNP was added in the primer extension assay to outcompete dATP and dCTP binding to WRN. In the presence of a high AMP-PNP concentration, WRN is incapable of stimulating DNA synthesis on $(CAG)_{35}$ or $(CTG)_{35}$ templates (Fig. 4.7A, lanes 3 and 6), implying the helicase activity is responsible for the induction of CAG/CTG synthesis.

To further confirm the role of the helicase activity in the stimulation of synthesis, WRN helicase mutant (K577M) was used in the assay. Previous studies have shown that a point mutation in the ATPase domain of WRN (K577M) abrogates the helicase but not the exonuclease activity (113,114). Similar to the AMP-PNP result, the WRN helicase mutant K577M could not stimulate the synthesis on $(CAG)_{35}$ or $(CTG)_{35}$ templates (Fig. 4.7B, lanes 3 and 6), suggesting WRN stimulates pol δ-catalyzed DNA synthesis on $(CAG)_{35}$ and $(CTG)_{35}$ templates by disrupting CAG/CTG hairpins.
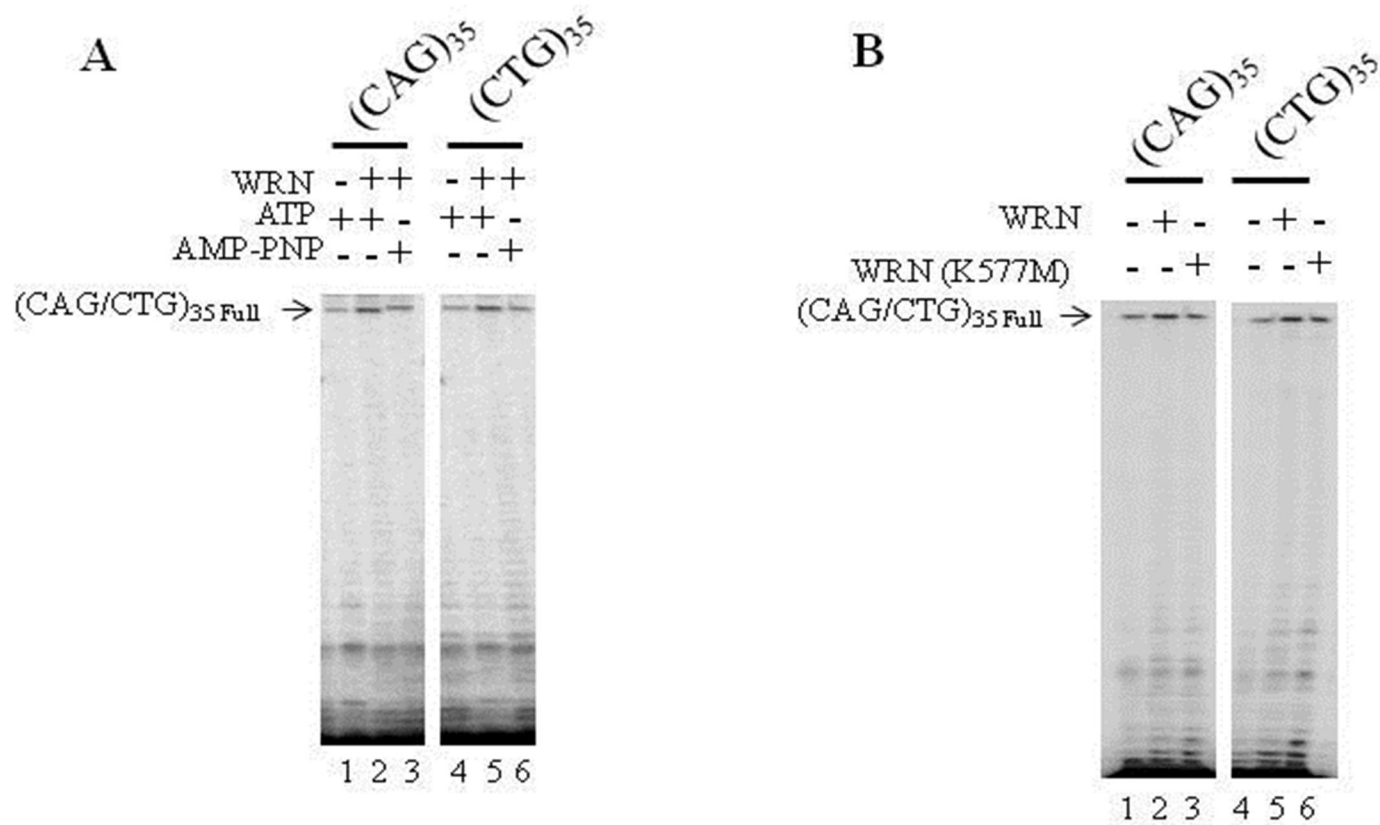
Figure 4.7 WRN helicase activity is responsible for stimulating CAG/CTG synthesis. A) The non-hydrolyzable ATP analog, AMP-PNP, inhibits the stimulatory effect of WRN on synthesis of $(CAG)_{35}$ and $(CTG)_{35}$ templates. B) WRN helicase mutant K577M failed to stimulate DNA synthesis of $(CAG)_{35}$ and $(CTG)_{35}$ templates.

4.2.4 WRN unwinds CTG hairpins with its helicase activity

To determine whether WRN is capable of unwinding CTG hairpins, we designed a helicase assay for detecting CTG hairpin unwinding (Fig. 4.8A). In our assay, the labeled $(CTG)_{35}$ was heated and slowly cooled to room temperature to allow intra-strand hairpin formation. When CTG forms an intra-strand hairpin, it cannot pair up with a complementary unlabeled CAG oligonucleotide. However, when the CTG hairpin is unwound, it can anneal with the unlabeled CAG oligonucleotide to form a double-stranded DNA. Since single-stranded DNA migrates differently than double-stranded DNA on native polyacrylamide gel electrophoresis, hairpin unwinding activity can be measured by the mobility difference. In the absence of WRN, a small amount of double stranded product was seen at 30 min. Since addition of high concentration of cold CAG could also result in the mobility shift even in the absence of any protein, this product may be formed by base pairing of CAG disruption of CTG hairpin formation. In the presence of wild-type WRN, unwinding of the hairpin was evident by 5 min (Fig. 4.8B, lane 6) and higher amount of hairpin was unwound in 30 min (Fig. 4.8B, lane 8). However, this unwinding was not observed in the absence of ATP, which is required for WRN helicase activity (Fig. 4.8B, lanes 9-12). In addition, the WRN helicase mutant K577M failed to unwind the CTG hairpin, further confirming that WRN unwinds CTG hairpins using its helicase activity. Therefore, our data suggests that WRN helicase promotes DNA hairpin repair by unwinding the hairpin structure so that it can be effectively used as a template for DNA synthesis.

Our helicase assay revealed WRN cannot 100% unwind $(CTG)_{35}$ (Fig. 4.8B, lane 8). With the fact that WRN has limited unwinding activity (on average unwinding less

than 40 base pairs) in the absence of replication protein A (RPA) (115), it is possible that WRN showed weaker hairpin unwinding activity on long CTG hairpin. In order to prove that the low helicase activity is not caused by the artifact of the assay, we repeated the helicase assay with a shorter CTG repeats (15 repeats). As expected, WRN could unwind almost all of the $(CTG)_{15}$ hairpins, confirming the assay's specificity (Fig. 4.8C).
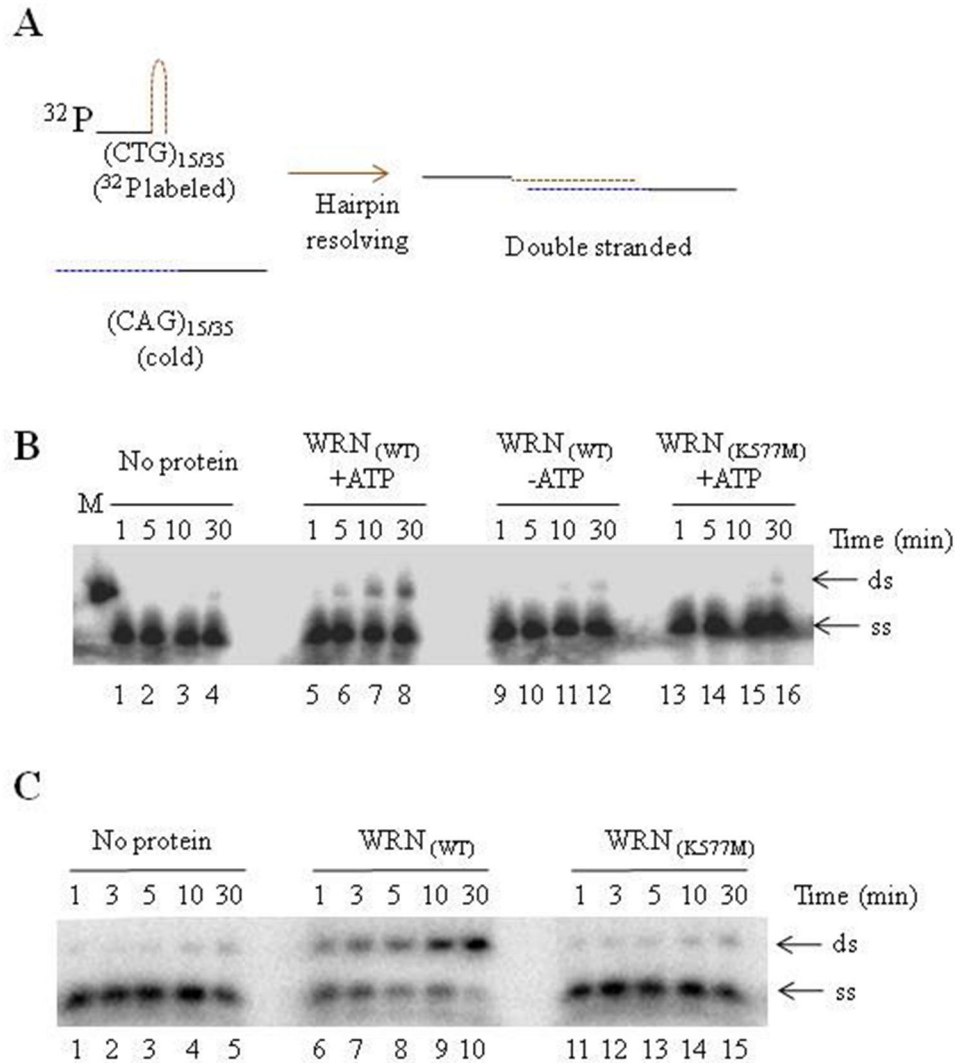
Figure 4.8 WRN unwinds CTG hairpins in a helicase activity. A) Experimental design of the helicase assay. 5'-^{32}P end labeled oligonucleotide containing (CTG)$_{35}$ was heated and slowly cooled down to room temperature to allow the intra-strand hairpin formation. When the CTG hairpin is unwound by helicase, it can hybridize with the unlabeled (CAG)$_{35}$ containing oligonucleotide causing a mobility shift. Native polyacrylamide gel (6%) electrophoresis was conducted to separate the single and double stranded DNA (indicated by arrows in B). A phosphorimager was used to detect the signal. B) Wild type WRN requires ATP to unwind the (CTG)$_{35}$ hairpin (lanes 5-8). Mutant WRN (K577M), which lacks helicase activity was unable to unwind the hairpin. The CAG/ CTG repeat region is indicated as dotted line. C) WRN is able to completely unwind (CTG)$_{15}$ hairpin. M: marker for double-stranded DNA, ss: single-stranded DNA, ds: double-stranded DNA.

4.3 Discussion

RecQ helicases are involved in many DNA metabolic pathways, such as DNA replication, recombination and repair (109). Five RecQ helicases have been found in humans, Werner Syndrome Protein (WRN), Bloom Syndrome Protein (BLM), RecQ1, RecQ4 and RecQ5. RecQ helicases migrate in a 3'to 5' direction on single-stranded DNA, using ATP as an energy source (68). When RecQ helicases migrate to the DNA duplex region, they dissociate the double-stranded DNA. Therefore, RecQ helicases are defined as 3'to 5' helicases because of the direction they travel relative to the DNA strand on which they are bound. RecQ helicases help maintaining genomic stability during DNA replication, recombination, transcription and repair. They can prevent the replication fork collapse by resolving the non-B DNA structures during DNA synthesis or and inhibiting recombination (116). Mutation of RecQ helicases can result in extensive DNA deletions, predisposition to cancer and pre-mature aging (117).

It is reported that CAG/CTG repeats form a stable hairpin *in vitro* and *in vivo* (16,118). Since CAG/CTG hairpins can hinder polymerase progression, it may trigger DNA damage and error prone repair or synthesis, leading to CAG/CTG repeat instability (60). If the hairpin is not resolved on the template strand, it induces repeat contraction. We show here that WRN plays an important role in CTG hairpin repair by unwinding the hairpins and preventing the repeat contraction. This finding is consistent with previous studies showing that depletion of the SGS1 helicase activity in yeast results in CTG repeat contraction (77).

In our primer extension assay, there is no strong pausing site in the absence of WRN. This could occur if the $(CTG)_{35}$ repeat does not form a uniform structure. It is still

83

unclear whether long CTG repeats can form multiple small hairpin structures or a single large hairpin (40). RecQ1 and BLM could not stimulate DNA synthesis on the $(CTG)_{35}$ substrate, suggesting $(CTG)_{35}$ repeats may cause greater hindrance to DNA synthesis.WRN has been reported to have a similar substrate specificity as BLM (116), but our results showed that only WRN could enhance pol δ synthesis on $(CTG)_{35}$, suggesting WRN has a higher substrate specificity and/or ability to resolve CAG/CTG hairpins.

Our helicase assay is based on the annealing of two single-stranded DNA. Since WRN also processes strand pairing activity of the single-stranded DNA (119), one may argue that the results shown in the helicase assay were caused by the strand pairing activity instead of helicase activity. As mentioned in Machwe et al. (119), strand pairing can still occur in the absence of ATP. In our case, the final product (double-strand DNA) that results when the hairpin is unwound was only seen in the presence of ATP, which eliminates the possibility of strand pairing activity. Furthermore, WRN helicase mutant (K577M), which cannot hydrolyze ATP and therefore lacks helicase activity, was unable to generate any double stranded products, indicating that our helicase assay is specific for assessing the unwinding activity of helicases for CAG/CTG hairpins. The final double-stranded product from the helicase assay contains a 5' overhang, so it cannot be unwound by the 3' to 5' helicase activity of WRN.

Previous studies pointed out that that WRN unwinding capacity is enhanced if the unwound product is prevented from reannealing (120). As for the primer synthesis assay, the pol δ polymerization can prevent reannealing of the unwound hairpin. Therefore, even though WRN did not show high $(CTG)_{35}$ unwinding activity in the helicase assay, it

is still possible that WRN can have higher unwinding activity when it cooperates with pol δ on $(CTG)_{35}$ synthesis by its helicase activity. In addition, with the fact that pol δ physically interacts with WRN (121), we cannot eliminate the possibility that pol δ binding enhances the unwinding capacity of WRN. Based on our results, we propose that when it comes to the replication/resynthesis blockage due to the CAG/CTG hairpin, WRN is recruited via interacting physically with pol δ, and unwinds the hairpin structure, allowing pol δ to resume polymerization reaction (Fig. 4.9)
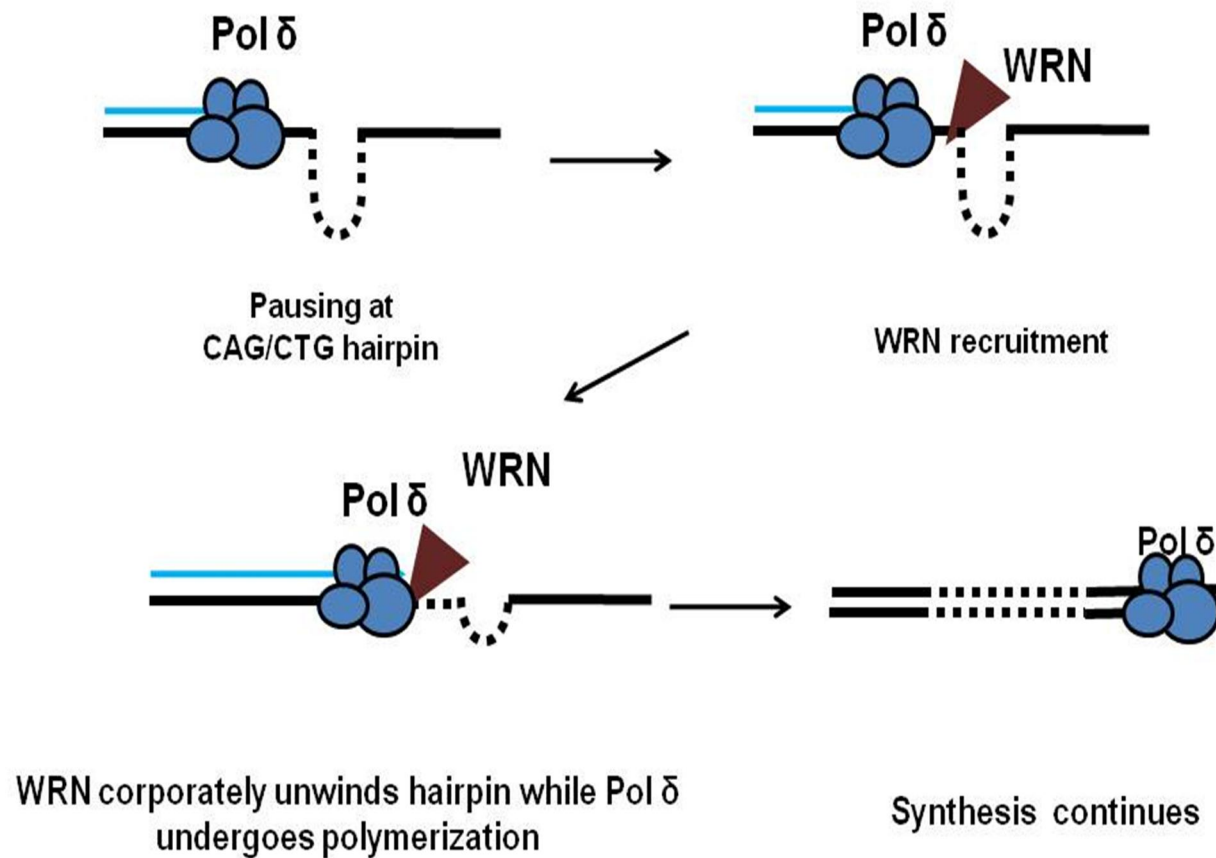
Figure 4.9 Proposed model for WRN stimulate CAG/CTG hairpin synthesis. Proposed model of WRN in maintaining CAG/CTG repeat stability. CAG/CTG hairpins hinder the pol δ progression. WRN is then recruited to the CAG/CTG hairpin through protein-protein interaction with pol δ. WRN resolves the CAG/CTG hairpin while pol δ is responsible for polymerization.

There is no clinical report linking WRN deficiency and CAG/CTG instability. However, one study found that the WRN knockout mice exhibited a deletion of telomere due to a deficiency of lagging strand synthesis (76). Since the telomere deletion may cause more serious problem than CAG/CTG repeat deletion, it is likely that WRN knockout mice do not live long enough to develop CAG/CTG repeat instability.

Our *in vitro* hairpin repair assay reveals that WRN stimulated hairpin repair when the hairpin was formed in the template strand (CTG slip-in). Therefore, inhibiting the CTG slip-in hairpin repair, repeat contraction will be resulted. It is hypothesized such contraction can be applied to the therapeutic treatment of repeat expansion associated diseases by reducing the expanded repeat to the sub-threshold level (4). With the limited knowledge of the proteins involved in the hairpin repair, our discovery of WRN in CTG unwinding can allow us to screen for the interacting proteins specifically responsible for promoting repeat contraction.

CHAPTER 5 Summary and Future Directions

5.1 Summary of findings

CAG/CTG hairpin formation and DNA synthesis are two key factors for repeat instability, suggesting both DNA polymerases and helicases play important roles in repeat instability. However, little is known regarding the proteins involved in tolerating or resolving CAG/CTG hairpins. Therefore, by screening some known polymerases and helicases, we identified pol β as a contributor to promote limited DNA synthesis after the hairpin, while its interaction with pol δ can have synergetic response to repeat expansion *in vitro*. Surprisingly, we discovered that pol δ can employ its proofreading activity to repair the CTG hairpin.

From the *in vitro* hairpin repair assay using the circular hairpin substrate, we identified WRN could stimulate the hairpin repair by resolving the hairpin. With the fact that CAG/CTG repeat instability associated disorders are still incurable, our studies can provide several potential therapeutic targets for treating and/or preventing CAG/CTG repeat associated disorders.

5.1.1 Pol β promotes hairpin retained

Hairpins caused by 3' slippage of DNA are hypothesized to be one of the intermediates triggering CAG/CTG expansion in cells (97). DNA synthesis without removing the hairpin is believed to convert the hairpin intermediate into the expanded product in the newly synthesized strand (27,97). Nonetheless, how 3' slipped hairpin

promotes repeat expansion is unknown. Our data reveal that pol β performs the limited

DNA synthesis at the hairpin tip, followed by the extension step by pol δ. In addition,

since pol δ can perform the extension step when there are two perfectly matched

nucleotides after the hairpin, it is suggested that pol β has to fill in some extra nucleotides

after the hairpin in order to promote repeat expansion. However, pol δ fails to perform the

extension step when there is a mismatch 5 bases away from after the hairpin, suggesting

the location of the mismatch can determine whether or not pol δ can repair the CTG

hairpin. As pol β is an error-prone enzyme and on occasion will misincorporate a

nucleotide after the hairpin, not all the hairpin substrate synthesized by pol β can be

converted to the extension product. Hence, this may act as another mechanism to prevent

CAG/CTG repeat expansion.

5.1.2 Pol δ removes hairpins by its exonuclease activity

Pol δ 3'to 5' exonuclease activity is important for maintaining genomic stability.

Our present studies demonstrate that pol δ proofreading activity can remove CTG

hairpins. Defects in pol δ proofreading activity contribute to elevated mutation

frequencies in yeast (122) and mice (123). Since there are no studies reporting that HD

patients have a higher mutation susceptibility, this suggests that pol δ's proofreading

activity should be proficient in these patients. If this is true, then how can the DNA of

HD patients undergo repeat expansion even though they possess pol δ with proficient

proofreading activity?

Limitations on the hairpin removal by pol δ may explain the repeat expansion

seen in HD patient cells. Our studies reveal that pol δ can only remove short CTG

hairpins (less than 5 CTG repeats) and this activity only works at the hairpin tip.

Therefore, CAG/CTG repeat expansion can still occur when the hairpin cannot be removed by pol δ. In short, our data suggests that pol δ may not play a significant role in hairpin removal due to these limitations.

5.1.3 WRN promotes CAG/CTG hairpin synthesis by its helicase activity

Disruption of the CAG/CTG hairpin structure is an alternative way to prevent repeat instability. Our studies reveal that WRN can enhance the $(CTG)_{35}$ hairpin repair by unwinding the hairpin and making it available for DNA synthesis. Since WRN is the homolog of yeast SGS1, our results support the findings in yeast that the depletion of SGS1 results in CTG repeat deletion during lagging strand synthesis (77).

5.2 Further Studies

5.2.1 Identifying proteins or inhibitors involved in hairpin removal

Our novel *in vitro* hairpin repair assay not only confirms the role of endonucleases in hairpin repair, but also reveals several characteristics of the hairpin removal. First, the incision and excision mechanisms for DNA removal are not overlapping. By screening for different nuclear extracts for deficiency in hairpin repair, we identified one of the HD cell lines is deficient in incision (Fig. 5.1). Second, cells may rely more on incision when there are perfectly base paired nucleotides at the 3' end of the hairpin. Third, incision is more dominant when the hairpin is large. Since incision is more structure-specific and responsible for repairing a larger hairpin, identifying the endonuclease involved in the CAG/CTG DNA removal may allow us to understand how the large hairpin is repaired. Several cell lines are defective in endonuclease activity. Therefore, by using protein fractionation and complementation assays, it is likely that the endonuclease or inhibitors can be identified.
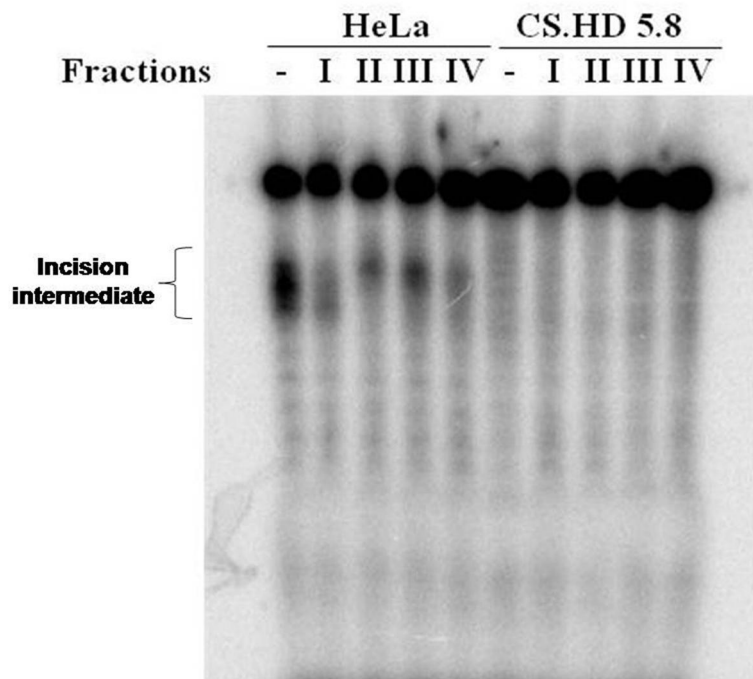
Figure 5.1 A cell line that is deficient in incision while it is proficient in excision. 5'-^{32}P labeled Substrate II (mentioned in Fig. 3.1) was incubated with HeLa nuclear extract and HD patient nuclear extract. HD nuclear extract showed deficient in incision. The smaller band showed in HeLa nuclear extracts is believed to be the hairpin removal intermediate, suggestion incision occurs in HeLa nuclear extract. Addition of Q-sepharose fraction could not restore the incision intermediate, suggesting inhibitor may be present in the HD patient cells. Reaction was conducted as the hairpin primer extension assay except in the presence of aphidicolin but absence of dNTP. Detailed method was mentioned in section 2.7.

5.2.2 Are other polymerases involved in CAG/CTG repeat instability?

Our results indicate that pol β is involved in promoting CTG expansion (and possibility contraction). In our experiments, we screened a number of different polymerases that were available to us. We cannot rule out the possibility that other error-prone polymerases that we did not test could be involved in the hairpin mediated expansion. In addition, by modifying our substrate as shown in Fig. 5.2, we can also screen polymerases to see if they can specifically promote repeat contraction when CAG/CTG hairpins are present.
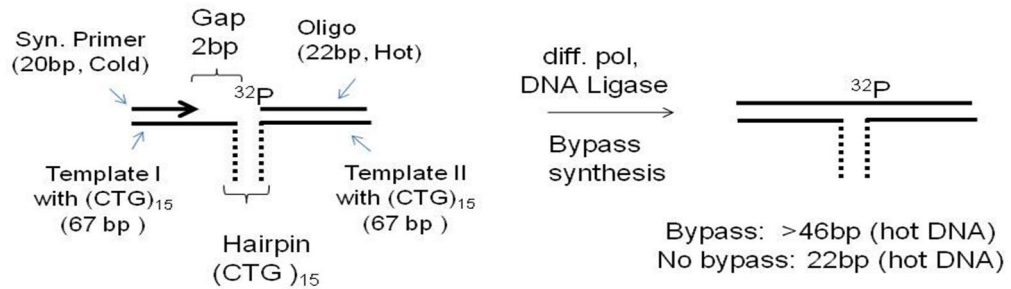
Figure 5.2 Proposed hairpin substrate for screening polymerase promoting repeat contraction. The hairpin gap substrate is created by annealing with 4 different oligonucleotides so as to simulate hairpin formation on the template during lagging strand synthesis. Template I and Template II contains 22 bp random sequence with 15 CTG repeats. Synthesis primer (Syn. Primer) is non-labeled with 20 bp complementary with Template I. Oligo is the 5-^{32}P oligonucleotide complementary to random sequence to Template II. After annealing, a hairpin substrate with 2 bp gap will be formed. After incubating with different polymerases and DNA ligase, if there is a hairpin bypass on the template strand, it will generate more than 46 bp radiolabeled DNA. However, if there is no bypass synthesis, radiolabeled substrate with 22 bp is expected to be found.

5.2.3 Are other helicases involved in CAG/CTG repeat instability?

There are numerous DNA helicases with different substrate specificities and activities. It has been shown that a yeast helicase, SRS2, demonstrates exceptionally high specificity and activity for unwinding CAG/CTG hairpins (37). Our results clearly demonstrate that WRN can indeed enhance synthesis through unwinding a CTG hairpin. However, patients with WRN deficiency do not demonstrate CAG/CTG repeat instability. It is likely that other helicases or proteins can be the backup to WRN and resolve CAG/CTG hairpins. Since human FBH1 is proposed to be the human orthologue of SRS2, functional assays using CAG/CTG repeats can be carried out to test a possible role of FBH-1 in repeat instability (72).

APPENDIX

Abbreviation

8-oxo-G	8-oxo-7,8-dihydro-2'-deoxyguanosine
A	Adenine
ATP	Adenosine Triphosphate
BER	Base excision repair
BLM	Bloom syndrome protein
C	Cytosine
CSA, CSB	Cockayne syndrome proteins A and B
DNA-PKcs	DNA protein kinase, catalytic subunit
DTT	Dithiothreitol
dNTP	Deoxynucleotide Triphosphate
dsDNS	Double strand DNA
EDTA	Ethylenediamine tretracetic acid
FBH1	F-Box helicase 1
G	Guanine
GGR	Global genome repair
HAP	Hydroxylapatite
HD	Huntington's disease
HPR	Hairpin repair
MMR	Mismatch repair
MutSα	Complex of MSH2 and MSH 6
MutSβ	Complex of MSH2 and MSH 3
NER	Nucleotide excision repair
NFDM	Non-fat dry milk

NP-40	Noniodent P-40
OGG1	8-oxoguanine DNA glycosylase
Pol	Polymerase
PAGE	Polyacrylamide gel electrophoresis
PCNA	Proliferating cell number antigen
RFC	Replication factor C
RPA	Replication protein A
SCA7	Spinocerebellar ataxia type 7
SDSA	Synthesis dependent strand annealing
SDS	Sodium Dodecyl Sulfate
ssDNA	Single-stranded DNA
SRS2	Suppressor of RAD Six Screen Mutant 2
T4PNK	T4 polynucleotide kinase
T	Thymine
TCR	transcription-couple repair
TNR	trinucleotide repeat
TLS	Translesion DNA synthesis
UV	Ultraviolet
WRN	Werner Syndrome Protein
XPC	Xeroderma pigmentosum complementation group C

REFERENCES

1.   de la Chapelle, A. (2003) Microsatellite instability. *N Engl J Med*, 349, 209-210.

2.   Rolfsmeier, M.L., Dixon, M.J., Pessoa-Brandao, L., Pelletier, R., Miret, J.J. and Lahue, R.S. (2001) Cis-elements governing trinucleotide repeat instability in Saccharomyces cerevisiae. *Genetics*, 157, 1569-1579.

3.   Pearson, C.E., Nichol Edamura, K. and Cleary, J.D. (2005) Repeat instability: mechanisms of dynamic mutations. *Nat Rev Genet*, 6, 729-742.

4.   Lopez Castel, A., Cleary, J.D. and Pearson, C.E. (2010) Repeat instability as the basis for human diseases and as a potential target for therapy. *Nat Rev Mol Cell Biol*, 11, 165-170.

5.   Dalby, A.R. (2009) A comparative proteomic analysis of the simple amino acid repeat distributions in Plasmodia reveals lineage specific amino acid selection. *PLoS One*, 4, e6231.

6.   Sinden, R.R. (1999) Biological implications of the DNA structures associated with disease-causing triplet repeats. *Am J Hum Genet*, 64, 346-353.

7.   Pearson, C.E. and Sinden, R.R. (1998) Trinucleotide repeat DNA structures: dynamic mutations from dynamic DNA. *Curr Opin Struct Biol*, 8, 321-330.

8.   Brouwer, J.R., Willemsen, R. and Oostra, B.A. (2009) Microsatellite repeat instability and neurological disease. *Bioessays*, 31, 71-83.

9.   Kang, S., Jaworski, A., Ohshima, K. and Wells, R.D. (1995) Expansion and deletion of CTG repeats from human disease genes are determined by the direction of replication in E. coli. *Nat Genet*, 10, 213-218.

10.   Maurer, D.J., O'Callaghan, B.L. and Livingston, D.M. (1996) Orientation dependence of trinucleotide CAG repeat instability in Saccharomyces cerevisiae. *Mol Cell Biol*, 16, 6617-6622.

11.   Cleary, J.D., Nichol, K., Wang, Y.H. and Pearson, C.E. (2002) Evidence of cis-acting factors in replication-mediated trinucleotide repeat instability in primate cells. *Nat Genet*, 31, 37-46.

12.   Panigrahi, G.B., Cleary, J.D. and Pearson, C.E. (2002) In vitro (CTG)*(CAG) expansions and deletions by human cell extracts. *J Biol Chem*, 277, 13926-13934.

13.   Dere, R., Napierala, M., Ranum, L.P. and Wells, R.D. (2004) Hairpin structure-forming propensity of the (CCTG.CAGG) tetranucleotide repeats contributes to the genetic instability associated with myotonic dystrophy type 2. *J Biol Chem*, 279, 41715-41726.

14.   Cleary, J.D. and Pearson, C.E. (2005) Replication fork dynamics and dynamic mutations: the fork-shift model of repeat instability. *Trends Genet*, 21, 272-280.

15.   Hartenstine, M.J., Goodman, M.F. and Petruska, J. (2000) Base stacking and even/odd behavior of hairpin loops in DNA triplet repeat slippage and expansion with DNA polymerase. *J Biol Chem*, 275, 18382-18390.

16.   Pearson, C.E., Tam, M., Wang, Y.H., Montgomery, S.E., Dar, A.C., Cleary, J.D. and Nichol, K. (2002) Slipped-strand DNAs formed by long (CAG)*(CTG) repeats: slipped-out repeats and slip-out junctions. *Nucleic Acids Res*, 30, 4534-4547.

17.   Liu, G., Chen, X., Bissler, J.J., Sinden, R.R. and Leffak, M. (2010) Replication-dependent instability at (CTG) x (CAG) repeat hairpins in human cells. *Nat Chem Biol*, 6, 652-659.

18.   Yang, J. and Freudenreich, C.H. (2007) Haploinsufficiency of yeast FEN1 causes instability of expanded CAG/CTG tracts in a length-dependent manner. *Gene*, 393, 110-115.

19.   Liu, Y., Prasad, R., Beard, W.A., Hou, E.W., Horton, J.K., McMurray, C.T. and Wilson, S.H. (2009) Coordination between polymerase beta and FEN1 can modulate CAG repeat expansion. *J Biol Chem*, 284, 28352-28366.

20.   Mirkin, S.M. (2007) Expandable DNA repeats and human disease. *Nature*, 447, 932-940.

21.   Wells, R.D., Dere, R., Hebert, M.L., Napierala, M. and Son, L.S. (2005) Advances in mechanisms of genetic instability related to hereditary neurological diseases. *Nucleic Acids Res*, 33, 3785-3798.

22.   Kang, S., Ohshima, K., Shimizu, M., Amirhaeri, S. and Wells, R.D. (1995) Pausing of DNA synthesis in vitro at specific loci in CTG and CGG triplet repeats from human hereditary disease genes. *J Biol Chem*, 270, 27014-27021.

23.   Krasilnikova, M.M. and Mirkin, S.M. (2004) Replication stalling at Friedreich's ataxia (GAA)n repeats in vivo. *Mol Cell Biol*, 24, 2286-2295.

24.   Lee, K.Y. and Myung, K. (2008) PCNA modifications for regulation of post-replication repair pathways. *Mol Cells*, 26, 5-11.

25.   Michel, B., Grompone, G., Florès, M.J. and Bidnenko, V. (2004) Multiple pathways process stalled replication forks. *Proc Natl Acad Sci U S A*, 101, 12783-12788.

26.   Dixon, M.J. and Lahue, R.S. (2002) Examining the potential role of DNA polymerases eta and zeta in triplet repeat instability in yeast. *DNA Repair (Amst)*, 1, 763-770.

27.   McMurray, C.T. (2010) Mechanisms of trinucleotide repeat instability during human development. *Nat Rev Genet*, 11, 786-799.

28.   Lin, Y., Hubert, L., Jr. and Wilson, J.H. (2009) Transcription destabilizes triplet repeats. *Mol Carcinog*, 48, 350-361.

29.   Nouspikel, T. and Hanawalt, P.C. (2000) Terminally differentiated human neurons repair transcribed genes but display attenuated global DNA repair and modulation of repair gene expression. *Mol Cell Biol*, 20, 1562-1570.

30.   Fishel, M.L., Vasko, M.R. and Kelley, M.R. (2007) DNA repair in neurons: so if they don't divide what's to repair? *Mutat Res*, 614, 24-36.

31.   Li, G.M. (2008) Mechanisms and functions of DNA mismatch repair. *Cell Res*, 18, 85-98.

32.   Jaworski, A., Rosche, W.A., Gellibolian, R., Kang, S., Shimizu, M., Bowater, R.P., Sinden, R.R. and Wells, R.D. (1995) Mismatch repair in Escherichia coli enhances instability of (CTG)n triplet repeats from human hereditary diseases. *Proc Natl Acad Sci U S A*, 92, 11019-11023.

33.   Manley, K., Shirley, T.L., Flaherty, L. and Messer, A. (1999) Msh2 deficiency prevents in vivo somatic instability of the CAG repeat in Huntington disease transgenic mice. *Nat Genet*, 23, 471-473.

34.   Owen, B.A., Yang, Z., Lai, M., Gajec, M., Badger, J.D., 2nd, Hayes, J.J., Edelmann, W., Kucherlapati, R., Wilson, T.M. and McMurray, C.T. (2005) (CAG)(n)-hairpin DNA binds to Msh2-Msh3 and changes properties of mismatch recognition. *Nat Struct Mol Biol*, 12, 663-670.

35.   Tian, L., Gu, L. and Li, G.M. (2009) Distinct nucleotide binding/hydrolysis properties and molar ratio of MutSalpha and MutSbeta determine their differential mismatch binding activities. *J Biol Chem*, 284, 11557-11562.

36.   Tian, L., Hou, C., Tian, K., Holcomb, N.C., Gu, L. and Li, G.M. (2009) Mismatch recognition protein MutSbeta does not hijack (CAG)n hairpin repair in vitro. *J Biol Chem*, 284, 20452-20456.

37.   Dhar, A. and Lahue, R.S. (2008) Rapid unwinding of triplet repeat hairpins by Srs2 helicase of Saccharomyces cerevisiae. *Nucleic Acids Res*, 36, 3366-3373.

38.   Kramer, P.R., Pearson, C.E. and Sinden, R.R. (1996) Stability of triplet repeats of myotonic dystrophy and fragile X loci in human mutator mismatch repair cell lines. *Hum Genet*, 98, 151-157.

39.   Panigrahi, G.B., Lau, R., Montgomery, S.E., Leonard, M.R. and Pearson, C.E. (2005) Slipped (CTG)*(CAG) repeats can be correctly repaired, escape repair or undergo error-prone repair. *Nat Struct Mol Biol*, 12, 654-662.

40.   Panigrahi, G.B., Slean, M.M., Simard, J.P., Gileadi, O. and Pearson, C.E. (2010) Isolated short CTG/CAG DNA slip-outs are repaired efficiently by hMutSbeta, but clustered slip-outs are poorly repaired. *Proc Natl Acad Sci U S A*, 107, 12593-12598.

41.   Smith, G.K., Jie, J., Fox, G.E. and Gao, X. (1995) DNA CTG triplet repeats involved in dynamic mutations of neurologically related gene sequences form stable duplexes. *Nucleic Acids Res*, 23, 4303-4311.

42.   Hanawalt, P.C., Ford, J.M. and Lloyd, D.R. (2003) Functional characterization of global genomic DNA repair and its implications for cancer. *Mutat Res*, 544, 107-114.

43.   Parniewski, P., Bacolla, A., Jaworski, A. and Wells, R.D. (1999) Nucleotide excision repair affects the stability of long transcribed (CTG*CAG) tracts in an orientation-dependent manner in Escherichia coli. *Nucleic Acids Res*, 27, 616-623.

44.   Dragileva, E., Hendricks, A., Teed, A., Gillis, T., Lopez, E.T., Friedberg, E.C., Kucherlapati, R., Edelmann, W., Lunetta, K.L., MacDonald, M.E. *et al.* (2009) Intergenerational and striatal CAG repeat instability in Huntington's disease knock-in mice involve different DNA repair genes. *Neurobiol Dis*, 33, 37-47.

45.   Lin, Y., Dion, V. and Wilson, J.H. (2006) Transcription promotes contraction of CAG repeat tracts in human cells. *Nat Struct Mol Biol*, 13, 179-180.

46.   Krokan, H.E., Nilsen, H., Skorpen, F., Otterlei, M. and Slupphaug, G. (2000) Base excision repair of DNA in mammalian cells. *FEBS Lett*, 476, 73-77.

47.   Kovtun, I.V., Liu, Y., Bjoras, M., Klungland, A., Wilson, S.H. and McMurray, C.T. (2007) OGG1 initiates age-dependent CAG trinucleotide expansion in somatic cells. *Nature*, 447, 447-452.

48.   Noseworthy, M.D. and Bray, T.M. (1998) Effect of oxidative stress on brain damage detected by MRI and in vivo 31P-NMR. *Free Radic Biol Med*, 24, 942-951.

49.   Beard, W.A., Batra, V.K. and Wilson, S.H. (2010) DNA polymerase structure-based insight on the mutagenic properties of 8-oxoguanine. *Mutat Res*, 703, 18-23.

50.   Goula, A.V., Berquist, B.R., Wilson, D.M., 3rd, Wheeler, V.C., Trottier, Y. and Merienne, K. (2009) Stoichiometry of base excision repair proteins correlates with increased somatic CAG instability in striatum over cerebellum in Huntington's disease transgenic mice. *PLoS Genet*, 5, e1000749.

51.   Floyd, R.A. and Hensley, K. (2002) Oxidative stress in brain aging. Implications for therapeutics of neurodegenerative diseases. *Neurobiol Aging*, 23, 795-807.

52.   Cortes-Ledesma, F. and Aguilera, A. (2006) Double-strand breaks arising by replication through a nick are repaired by cohesin-dependent sister-chromatid exchange. *EMBO Rep*, 7, 919-926.

53.   Wyman, C. and Kanaar, R. (2006) DNA double-strand break repair: all's well that ends well. *Annu Rev Genet*, 40, 363-383.

54.   Meservy, J.L., Sargent, R.G., Iyer, R.R., Chan, F., McKenzie, G.J., Wells, R.D. and Wilson, J.H. (2003) Long CTG tracts from the myotonic dystrophy gene induce deletions and rearrangements during recombination at the APRT locus in CHO cells. *Mol Cell Biol*, 23, 3152-3162.

55.   Jankowski, C., Nasar, F. and Nag, D.K. (2000) Meiotic instability of CAG repeat tracts occurs by double-strand break repair in yeast. *Proc Natl Acad Sci U S A*, 97, 2134-2139.

56.   Hebert, M.L., Spitz, L.A. and Wells, R.D. (2004) DNA double-strand breaks induce deletion of CTG.CAG repeats in an orientation-dependent manner in Escherichia coli. *J Mol Biol*, 336, 655-672.

57.   Freudenreich, C.H., Kantrow, S.M. and Zakian, V.A. (1998) Expansion and length-dependent fragility of CTG repeats in yeast. *Science*, 279, 853-856.

58.   O'Driscoll, M. and Jeggo, P.A. (2006) The role of double-strand break repair - insights from human genetics. *Nat Rev Genet*, 7, 45-54.

59.   Savouret, C., Brisson, E., Essers, J., Kanaar, R., Pastink, A., te Riele, H., Junien, C. and Gourdon, G. (2003) CTG repeat instability and size variation timing in DNA repair-deficient mice. *EMBO J*, 22, 2264-2273.

60.   Mittelman, D., Moye, C., Morton, J., Sykoudis, K., Lin, Y., Carroll, D. and Wilson, J.H. (2009) Zinc-finger directed double-strand breaks within CAG repeat tracts promote repeat instability in human cells. *Proc Natl Acad Sci U S A*, 106, 9607-9612.

61.   Hashem, V.I., Rosche, W.A. and Sinden, R.R. (2004) Genetic recombination destabilizes (CTG)n.(CAG)n repeats in E. coli. *Mutat Res*, 554, 95-109.

62.   Paques, F. and Haber, J.E. (1999) Multiple pathways of recombination induced by double-strand breaks in Saccharomyces cerevisiae. *Microbiol Mol Biol Rev*, 63, 349-404.

63.   Mangiarini, L., Sathasivam, K., Mahal, A., Mott, R., Seller, M. and Bates, G.P. (1997) Instability of highly expanded CAG repeats in mice transgenic for the Huntington's disease mutation. *Nat Genet*, 15, 197-200.

64.   Bowater, R.P., Jaworski, A., Larson, J.E., Parniewski, P. and Wells, R.D. (1997) Transcription increases the deletion frequency of long CTG.CAG triplet repeats from plasmids in Escherichia coli. *Nucleic Acids Res*, 25, 2861-2868.

65.   Svejstrup, J.Q. (2002) Mechanisms of transcription-coupled DNA repair. *Nat Rev Mol Cell Biol*, 3, 21-29.

66.   Lin, Y., Leng, M., Wan, M. and Wilson, J.H. (2010) Convergent transcription through a long CAG tract destabilizes repeats and induces apoptosis. *Mol Cell Biol*, 30, 4435-4451.

67.   Hou, C., Chan, N.L., Gu, L. and Li, G.M. (2009) Incision-dependent and error-free repair of (CAG)(n)/(CTG)(n) hairpins in human cell extracts. *Nat Struct Mol Biol*, 16, 869-875.

68.   Bernstein, K.A., Gangloff, S. and Rothstein, R. (2010) The RecQ DNA helicases in DNA repair. *Annu Rev Genet*, 44, 393-417.

69.   Bhattacharyya, S. and Lahue, R.S. (2004) Saccharomyces cerevisiae Srs2 DNA helicase selectively blocks expansions of trinucleotide repeats. *Mol Cell Biol*, 24, 7324-7330.

70.   Bhattacharyya, S. and Lahue, R.S. (2005) Srs2 helicase of Saccharomyces cerevisiae selectively unwinds triplet repeat DNA. *J Biol Chem*, 280, 33311-33317.

71.   Huang, M.E., de Calignon, A., Nicolas, A. and Galibert, F. (2000) POL32, a subunit of the Saccharomyces cerevisiae DNA polymerase delta, defines a link between DNA replication and the mutagenic bypass repair pathway. *Curr Genet*, 38, 178-187.

72.   Chiolo, I., Saponaro, M., Baryshnikova, A., Kim, J.H., Seo, Y.S. and Liberi, G. (2007) The human F-Box DNA helicase FBH1 faces Saccharomyces cerevisiae Srs2 and postreplication repair pathway roles. *Mol Cell Biol*, 27, 7439-7450.

73.   Fry, M. and Loeb, L.A. (1999) Human werner syndrome DNA helicase unwinds tetrahelical structures of the fragile X syndrome repeat sequence d(CGG)n. *J Biol Chem*, 274, 12797-12802.

74.   Kamath-Loeb, A.S., Loeb, L.A., Johansson, E., Burgers, P.M. and Fry, M. (2001) Interactions between the Werner syndrome helicase and DNA polymerase delta specifically facilitate copying of tetraplex and hairpin structures of the d(CGG)n trinucleotide repeat sequence. *J Biol Chem*, 276, 16439-16446.

75.   Fukuchi, K., Martin, G.M. and Monnat, R.J., Jr. (1989) Mutator phenotype of Werner syndrome is characterized by extensive deletions. *Proc Natl Acad Sci U S A*, 86, 5893-5897.

76.   Crabbe, L., Verdun, R.E., Haggblom, C.I. and Karlseder, J. (2004) Defective telomere lagging strand synthesis in cells lacking WRN helicase activity. *Science*, 306, 1951-1953.

77.   Kerrest, A., Anand, R.P., Sundararajan, R., Bermejo, R., Liberi, G., Dujon, B., Freudenreich, C.H. and Richard, G.F. (2009) SRS2 and SGS1 prevent chromosomal breaks and stabilize triplet repeats by restraining recombination. *Nat Struct Mol Biol*, 16, 159-167.

78.   Ramilo, C., Gu, L., Guo, S., Zhang, X., Patrick, S.M., Turchi, J.J. and Li, G.M. (2002) Partial Reconstitution of Human DNA Mismatch Repair In Vitro: Characterization of the Role of Human Replication Protein A. *Mol Cell Biol*, 22, 2037-2046.

79.   Podust, V.N., Chang, L.S., Ott, R., Dianov, G.L. and Fanning, E. (2002) Reconstitution of human DNA polymerase delta using recombinant baculoviruses: the p12 subunit potentiates DNA polymerizing activity of the four-subunit enzyme. *J Biol Chem*, 277, 3894-3901.

80.   Cai, J., Uhlmann, F., Gibbs, E., Flores-Rozas, H., Lee, C.G., Phillips, B., Finkelstein, J., Yao, N., O'Donnell, M. and Hurwitz, J. (1996) Reconstitution of human replication factor C from its five subunits in baculovirus-infected insect cells. *Proc Natl Acad Sci U S A*, 93, 12896-12901.

81.   Fien, K. and Stillman, B. (1992) Identification of replication factor C from Saccharomyces cerevisiae: a component of the leading-strand DNA replication complex. *Mol Cell Biol*, 12, 155-163.

82.   Holmes, S.E., Wentzell, J.S., Seixas, A.I., Callahan, C., Silveira, I., Ross, C.A. and Margolis, R.L. (2006) A novel trinucleotide repeat expansion at chromosome 3q26.2 identified by a CAG/CTG repeat expansion detection array. *Hum Genet*, 120, 193-200.

83.   Nakamori, M., Pearson, C.E. and Thornton, C.A. (2011) Bidirectional transcription stimulates expansion and contraction of expanded (CTG)*(CAG) repeats. *Hum Mol Genet*, 20, 580-588.

84.   Li, X.J. and Li, S. (2010) Proteasomal dysfunction in aging and Huntington disease. *Neurobiol Dis*.

85.   Mirkin, E.V. and Mirkin, S.M. (2007) Replication fork stalling at natural impediments. *Microbiol Mol Biol Rev*, 71, 13-35.

86.   Gacy, A.M., Goellner, G., Juranic, N., Macura, S. and McMurray, C.T. (1995) Trinucleotide repeats that expand in human disease form hairpin structures in vitro. *Cell*, 81, 533-540.

87.   Moore, H., Greenwell, P.W., Liu, C.P., Arnheim, N. and Petes, T.D. (1999) Triplet repeats form secondary structures that escape DNA repair in yeast. *Proc Natl Acad Sci U S A*, 96, 1504-1509.

88.   Petruska, J., Arnheim, N. and Goodman, M.F. (1996) Stability of intrastrand hairpin structures formed by the CAG/CTG class of DNA triplet repeats associated with neurological diseases. *Nucleic Acids Res*, 24, 1992-1998.

89.   Hou, C., Zhang, T., Tian, L., Huang, J., Gu, L. and Li, G.M. (2011) The Role of XPG in Processing (CAG)n/(CTG)n DNA Hairpins. *Cell Biosci*, 1, 11.

90.   Petruska, J., Hartenstine, M.J. and Goodman, M.F. (1998) Analysis of strand slippage in DNA polymerase expansions of CAG/CTG triplet repeats associated with neurodegenerative disease. *J Biol Chem*, 273, 5204-5210.

91.   Pelletier, R., Farrell, B.T., Miret, J.J. and Lahue, R.S. (2005) Mechanistic features of CAG*CTG repeat contractions in cultured cells revealed by a novel genetic assay. *Nucleic Acids Res*, 33, 5667-5676.

92.   Usdin, K. and Grabczyk, E. (2000) DNA repeat expansions and human disease. *Cell Mol Life Sci*, 57, 914-931.

93.   Lange, S.S., Takata, K. and Wood, R.D. DNA polymerases and cancer. *Nat Rev Cancer*, 11, 96-110.

94.   Lehmann, A.R., Niimi, A., Ogi, T., Brown, S., Sabbioneda, S., Wing, J.F., Kannouche, P.L. and Green, C.M. (2007) Translesion synthesis: Y-family polymerases and the polymerase switch. *DNA Repair (Amst)*, 6, 891-899.

95.   Fazlieva, R., Spittle, C.S., Morrissey, D., Hayashi, H., Yan, H. and Matsumoto, Y. (2009) Proofreading exonuclease activity of human DNA polymerase delta and its effects on lesion-bypass DNA synthesis. *Nucleic Acids Res*, 37, 2854-2866.

96.   de Noronha, C.M. and Mullins, J.I. (1992) Amplimers with 3'-terminal phosphorothioate linkages resist degradation by vent polymerase and reduce Taq polymerase mispriming. *PCR Methods Appl*, 2, 131-136.

97.   Daee, D.L., Mertz, T. and Lahue, R.S. (2007) Postreplication repair inhibits CAG.CTG repeat expansions in Saccharomyces cerevisiae. *Mol Cell Biol*, 27, 102-110.

98.   Eckert, K.A., Hile, S.E. and Vargo, P.L. (1997) Development and use of an in vitro HSV-tk forward mutation assay to study eukaryotic DNA polymerase processing of DNA alkyl lesions. *Nucleic Acids Res*, 25, 1450-1457.

99.   Kunkel, T.A. (1985) The mutational specificity of DNA polymerase-beta during in vitro DNA synthesis. Production of frameshift, base substitution, and deletion mutations. *J Biol Chem*, 260, 5787-5796.

100.   Chang, L.M. and Bollum, F.J. (1973) A comparison of associated enzyme activities in various deoxyribonucleic acid polymerases. *J Biol Chem*, 248, 3398-3404.

101.   Yamtich, J. and Sweasy, J.B. (2010) DNA polymerase family X: function, structure, and cellular roles. *Biochim Biophys Acta*, 1804, 1136-1150.

102.   Fernandez-Lopez, L., Pineiro, E., Marcos, R., Velazquez, A. and Surralles, J. (2004) Induction of instability of normal length trinucleotide repeats within human disease genes. *J Med Genet*, 41, e3.

103.   Shaik, N.A., Govindan, S., Kodati, V., Rao, K.P. and Hasan, Q. (2009) Polymorphic (CAG)n repeats in the androgen receptor gene: a risk marker for endometriosis and uterine leiomyomas. *Hematol Oncol Stem Cell Ther*, 2, 289-293.

104.   Rozanska, M., Sobczak, K., Jasinska, A., Napierala, M., Kaczynska, D., Czerny, A., Koziel, M., Kozlowski, P., Olejniczak, M. and Krzyzosiak, W.J. (2007) CAG and CTG repeat polymorphism in exons of human genes shows distinct features at the expandable loci. *Hum Mutat*, 28, 451-458.

105.   Gatchel, J.R. and Zoghbi, H.Y. (2005) Diseases of unstable repeat expansion: mechanisms and common principles. *Nat Rev Genet*, 6, 743-755.

106.   Kim, S.H., Pytlos, M.J. and Sinden, R.R. (2006) Replication restart: a pathway for (CTG).(CAG) repeat deletion in Escherichia coli. *Mutat Res*, 595, 5-22.

107.   Mankouri, H.W. and Hickson, I.D. (2004) Understanding the roles of RecQ helicases in the maintenance of genome integrity and suppression of tumorigenesis. *Biochem Soc Trans*, 32, 957-958.

108.   Myung, K., Datta, A., Chen, C. and Kolodner, R.D. (2001) SGS1, the Saccharomyces cerevisiae homologue of BLM and WRN, suppresses genome instability and homeologous recombination. *Nat Genet*, 27, 113-116.

109.   Bohr, V.A. (2008) Rising from the RecQ-age: the role of human RecQ helicases in genome maintenance. *Trends Biochem Sci*, 33, 609-620.

110.   Shah, S.N., Opresko, P.L., Meng, X., Lee, M.Y. and Eckert, K.A. (2010) DNA structure and the Werner protein modulate human DNA polymerase delta-dependent replication dynamics within the common fragile site FRA16D. *Nucleic Acids Res*, 38, 1149-1162.

111.   Delagoutte, E., Goellner, G.M., Guo, J., Baldacci, G. and McMurray, C.T. (2008) Single-stranded DNA-binding protein in vitro eliminates the orientation-dependent impediment to polymerase passage on CAG/CTG repeats. *J Biol Chem*, 283, 13341-13356.

112.   Shen, J.C., Gray, M.D., Oshima, J. and Loeb, L.A. (1998) Characterization of Werner syndrome protein DNA helicase activity: directionality, substrate dependence and stimulation by replication protein A. *Nucleic Acids Res*, 26, 2879-2885.

113.   Opresko, P.L., Laine, J.P., Brosh, R.M., Jr., Seidman, M.M. and Bohr, V.A. (2001) Coordinate action of the helicase and 3' to 5' exonuclease of Werner syndrome protein. *J Biol Chem*, 276, 44677-44687.

114.   Gray, M.D., Shen, J.C., Kamath-Loeb, A.S., Blank, A., Sopher, B.L., Martin, G.M., Oshima, J. and Loeb, L.A. (1997) The Werner syndrome protein is a DNA helicase. *Nat Genet*, 17, 100-103.

115.   Brosh, R.M., Jr., Orren, D.K., Nehlin, J.O., Ravn, P.H., Kenny, M.K., Machwe, A. and Bohr, V.A. (1999) Functional and physical interaction between WRN helicase and human replication protein A. *J Biol Chem*, 274, 18341-18350.

116.   Bachrati, C.Z. and Hickson, I.D. (2003) RecQ helicases: suppressors of tumorigenesis and premature aging. *Biochem J*, 374, 577-606.

117.   Furuichi, Y. (2001) Premature aging and predisposition to cancers caused by mutations in RecQ family helicases. *Annals of the New York Academy of Sciences*, 928, 121-131.

118.   Yu, A., Dill, J., Wirth, S.S., Huang, G., Lee, V.H., Haworth, I.S. and Mitas, M. (1995) The trinucleotide repeat sequence d(GTC)15 adopts a hairpin conformation. *Nucleic Acids Res*, 23, 2706-2714.

119.   Machwe, A., Xiao, L., Groden, J., Matson, S.W. and Orren, D.K. (2005) RecQ family members combine strand pairing and unwinding activities to catalyze strand exchange. *J Biol Chem*, 280, 23397-23407.

120.   Machwe, A., Lozada, E.M., Xiao, L. and Orren, D.K. (2006) Competition between the DNA unwinding and strand pairing activities of the Werner and Bloom syndrome proteins. *BMC Mol Biol*, 7, 1.

121.   Szekely, A.M., Chen, Y.H., Zhang, C., Oshima, J. and Weissman, S.M. (2000) Werner protein recruits DNA polymerase delta to the nucleolus. *Proc Natl Acad Sci U S A*, 97, 11365-11370.

122.   Perrino, F.W. and Loeb, L.A. (1990) Hydrolysis of 3'-terminal mispairs in vitro by the 3'----5' exonuclease of DNA polymerase delta permits subsequent extension by DNA polymerase alpha. *Biochemistry*, 29, 5226-5231.

123.   Goldsby, R.E., Lawrence, N.A., Hays, L.E., Olmsted, E.A., Chen, X., Singh, M. and Preston, B.D. (2001) Defective DNA polymerase-delta proofreading causes cancer susceptibility in mice. *Nat Med*, 7, 638-639

VITA

Date of Birth: 11/22/1982   Place of Birth: Hong Kong, China

Education

PhD   University of Kentucky, Toxicology   expected 2011
MSc   University of Kentucky, Nutritional Sciences   2007
BSc   The Chinese University of Hong Kong, Biochemistry   2005

Honors

1. Research Challenge Trust Fund Fellowship (2006)
2. NIEHS Predoctorial Fellowship (2007-2010)
3. University of Kentucky Graduate School Travel Award (2009)

Publications

1. Hou C, **Chan NL**, Gu L, Li GM. (2009) Incision-dependent and error-free repair of (CAG)(n)/(CTG)(n) hairpins in human cell extracts. Nat Struct Mol Biol. 16(8):869-75.
2. Glauert HP, Tharappel JC, Lu Z, Stemm D, Banerjee S, **Chan LS**, Lee EY, Lehmler HJ, Robertson LW, Spear BT.(2008) Role of Oxidative Stress in the Promoting Activities of PCBs. Environ Toxicol Pharmacol. 25(2):247-250.
3. Glauert HP, Tharappel JC, Banerjee S, **Chan NL**, Kania-Korwel I, Lehmler HJ, Lee EY, Robertson LW, Spear BT. (2008) Inhibition of the promotion of hepatocarcinogenesis by 2,2',4,4',5,5'-hexachlorobiphenyl (PCB-153) by the deletion of the p50 subunit of NF-kappa B in mice. Toxicol Appl Pharmacol. 232(2):302-8.
4. **Chan NL**, Wang H, Wang Y, Leung HY, Leung LK. (2006) Polycyclic aromatic hydrocarbon-induced CYP1B1 activity is suppressed by perillyl alcohol in MCF-7 cells. Toxicol Appl Pharmacol. 213(2):98-104.

Presentations

1. Experimental Biology meeting, San Diego, CA, April 5-9, 2008
2. US-EU Repair of Endogenous Genome Damage Conference, Galveston, TX, February 21-25, 2009
3. Environmental Health Sciences Showcase, Cincinnati, OH, September 18, 2009
4. Gordon Research Conferences on Mammalian DNA Repair, Ventura, CA, Feb 6-11, 2011