

Trinucleotide repeat deletion via a unique hairpin bypass by DNA polymerase β and alternate flap cleavage by flap endonuclease 1

Meng Xu, Jonathan Gabison and Yuan Liu*

Department of Chemistry and Biochemistry, Florida International University, 11200 SW 8th Street, Miami, FL 33199, USA

Received August 10, 2012; Revised October 30, 2012; Accepted November 18, 2012

ABSTRACT

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

INTRODUCTION

Trinucleotide repeats (TNRs) are highly polymorphic in the human genome and confer their instability via expansions or deletions (1,2). TNR expansions are known to be the cause of more than 40 neurodegenerative diseases, including Huntington's disease, myotonic dystrophy and fragile X syndrome among others (2–4). Similarly, TNR deletions are also closely related to human disorders. Epidemiological studies have demonstrated that CAG repeat deletions in the androgen receptor are correlated with human ovarian (5) and prostate cancer (6,7).

TNR expansion has been studied extensively in bacteria (8), yeast (1), mouse (9), human (10,11) and human cell extracts (12). It was initially shown that TNR expansions occur during DNA replication of dividing somatic cells and recombination of germ cells (3,11,13,14). However, later studies have demonstrated that expansions can also be mediated by DNA repair in dividing and non-dividing somatic cells (3,15,16). It is proposed that the formation of non-B form DNA secondary structures during DNA replication, recombination and repair contributes to expansions (11,17). In the human genome, typical non-B form structures arising from (CAG)_n/(CTG)_n, (CTG)_n/(CAG)_n, (CGG)_n/(CCG)_n and (GAA)_n/(TTC)_n sequence contexts include hairpins, tetraplexes and triplexes (11,18). These DNA secondary structures can cause DNA polymerase pausing and DNA slippage (19,20), replication fork stalling and collapse (14,18,21), inhibition of flap cleavage by flap endonuclease 1 (FEN1) (22–24), trapping of mismatch repair proteins (25,26) and disruption of coordination between DNA repair proteins (27). All of these events can compromise cellular repair mechanisms for removing the non-B form structures, thereby allowing them to be integrated into the human genome for TNR expansions. Hence, TNR expansion appears to be a consequence of the interactions between DNA secondary structures and DNA replication, repair and recombination. Compared with the other DNA metabolic pathways, DNA repair plays a more versatile role in modulating TNR stability because it can repair the non-B form DNA structures generated from both DNA damages and other DNA metabolic pathways.

Several DNA repair pathways have been shown to modulate TNR expansion during both DNA replication and damage repair. The repair pathways that are initiated during DNA replication include mismatch repair (28–32) and double-stranded DNA break repair, along with DNA recombination (1,17,33–38). A mismatch repair protein complex, MSH2/MSH3 can bind and stabilize hairpin structures, thereby promoting TNR expansion (25,26). Double-stranded break repair processes the intermediates

*To whom correspondence should be addressed. Tel: +1 305 348 3628; Fax: +1 305 348 3772; Email: yualiu@fiu.edu

resulting from TNR-induced replication blockage and fork stalling (14,18,39), therefore facilitates large expansions. A specific DNA repair pathway can be initiated directly by DNA damage in the context of TNR to modulate repeat stability (15,27). For example, 8-oxoguanine (8-oxoG) in TNRs initiates DNA base excision repair (BER) through 8-oxoG DNA glycosylase (OGG1), which plays an essential role in mediating age-dependent somatic CAG or CGG repeat expansions in a Huntington's disease and fragile X syndrome mouse models (15,40,41). The DNA glycosylase removes the oxidized base lesion leaving an abasic site (27), which is subsequently incised by apurinic/apyrimidinic (AP) endonuclease 1 (APE1). This results in single-stranded DNA (ssDNA) breakage, which leads to DNA slippage and the formation of hairpins and a multi-nucleotide gap (27). This further causes the disruption of functional coordination between DNA polymerase β (pol β) and FEN1 leading to an inefficient pol β multi-nucleotide gap-filling synthesis and FEN1 alternate cleavage, i.e. the cleavage of a short 5'-flap attached to the 5'-end of a hairpin. This subsequently results in production of a ligatable nick for ligating a hairpin with newly synthesized repeats leading to repeat expansion (27). Thus, inefficient BER leads to TNR expansion (42). A multi-nucleotide gap generated during BER could also allow the formation of a hairpin on the template strand that may mediate TNR deletion induced by DNA damages.

TNR deletions can be induced in human lymphoblasts (43), bacterial cells (44) and mouse sperm (45), by DNA-damaging agents including ethyl methanesulfonate, mitomycin C, ethylnitrosourea, ultraviolet radiation, cisplatin and ionizing radiation (43,45,46). Because the DNA-damaging agents can induce alkylating DNA damage, DNA strand breaks, DNA cross-links that are usually subjected to BER and nucleotide excision repair, it suggests that BER and NER are actively involved in mediating TNR deletion. The transcription-coupled NER was found to be associated with TNR deletion in mammalian cells (47–49) and in *Drosophila* (50), thereby supporting this notion. NER proteins, Xeroderma pigmentosum group A-complementing protein (48), Xeroderma pigmentosum group G-complementing protein and Cockayne syndrome B protein are also known to promote TNR deletion (47,49,50). These findings suggest an important role of DNA damage and repair in modulating TNR deletion.

Although significant progress has been made in exploring TNR deletion in recent years, our knowledge about this process remains limited. This results in poor understanding of the mechanisms by which deletions arise. Understanding the molecular basis of TNR deletions and how chemicals may induce deletions through DNA damage and repair is critically important for the development of new methods for prevention and treatment of human diseases caused by both TNR deletions and expansions (51). Interestingly, in addition to its role in promoting CAG repeat expansion (15,27), oxidative DNA damage was also shown to induce repeat deletions in bacteria (44). This suggests that oxidized base lesions and BER play a dual role in causing both expansions and

deletions. This concept has been further supported by *in vitro* evidence that pol β slippage is involved in CAG repeat deletion (52). However, it remains unknown if or how DNA base lesions and BER cause deletions. To explore the potential role of BER in TNR deletions, we examined CAG repeat deletions during BER and identified the roles of major BER enzymes in the steps that could lead to TNR deletions. Surprisingly, we discovered that repair of a DNA base lesion in the context of (CAG)₂₀ and (CAG)₂₅ repeats led to repeat deletions with the concomitant loss of 1–21 triplet repeats per repair event, depending upon the length of repeats. The deletions were mediated by pol β hairpin bypass in coordination with FEN1 alternate flap cleavage. As a result, we propose a hypothetical model to describe a unique mechanism underlying TNR deletions involving a base lesion-induced loss of TNR through pol β hairpin bypass and FEN1 alternate flap cleavage. Our studies provide the first evidence that TNR deletion can be initiated by DNA base lesions and mediated by BER.

MATERIALS AND METHODS

Materials

DNA oligonucleotides were from Integrated DNA Technologies (IDT Inc., Coralville, IA, USA). The radionucleotides [γ -³²P] ATP (6000 mCi/mmol) and Cordycepin 5'-triphosphate 3'-[α -³²P] (5000 mCi/mmol) were from PerkinElmer Inc. (Boston, MA, USA). Micro Bio-Spin 6 chromatography columns were purchased from Bio-Rad (Hercules, CA, USA). Deoxynucleoside 5'-triphosphates (dNTPs) and terminal deoxynucleotidyl transferase were from Fermentas (Glen Burnie, MD, USA). T4 polynucleotide kinase was from USB Corp. (Cleveland, OH, USA). Mung Bean Nuclease was from Epicenter (Madison, WI, USA). The original TA cloning kit was from Invitrogen/Life Technologies Corp. (Grand Island, NY, USA). The Big Dye Terminator version 3.1 Cycle Sequencing Kit was from Applied Biosystems/Life Technologies Corp (Grand Island, NY, USA). Polyclonal and monoclonal antibodies of FEN1 and tubulin and non-specific IgG were from Abcam (Boston, MA, USA). All standard chemical reagents were purchased from Sigma-Aldrich (St. Louis, MO, USA) and Thermo Fisher Scientific (Pittsburgh, PA, USA). Purified *Escherichia coli* uracil-DNA glycosylase (UDG), APE1, pol β , FEN1, DNA ligase I (LIG I) and a monoclonal pol β antibody were generous gifts from Dr Samuel Wilson at the National Institute of Environmental Health Sciences/National Institutes of Health, Research Triangle Park, NC, USA.

Oligonucleotide substrates

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

deoxyuridine residue. A substrate with a native abasic site was generated by incubating an oligonucleotide containing a deoxyuridine residue, with *E. coli* UDG (2.5 nM) at 37°C for 15 min. Substrates were constructed by annealing the damaged strands to their template strands at a molar ratio of 1:1.5. Substrates with a perfect 15-nt base-paired hairpin or a preformed (CTG)₇ or a (CAG)₇ or a (CTG)₂₅ hairpin in their template strands were constructed to mimic the repair intermediates containing a template hairpin with or without a 5'-downstream flap. The downstream primers contained a 5'-phosphate-THF alone or a 5'-phosphorylated THF with a 21-nt-dT flap or a (CAG)₇ or a (CTG)₇ or a (CAG)₁₂ or a (CAG)₂₅ flap/hairpin. Upstream and downstream primers were annealed to their templates at a molar ratio of 1:1.5. Substrates were radiolabeled at the 5'-end of the damaged strand or that of the upstream primer or at the 3'- or 5'-end of the template strand for measuring various types of enzymatic activities. The sequences of oligonucleotide substrates are listed in Supplementary Table S1. DNA size markers that correspond to the cleavage products of Mung Bean Nuclease and the products of pol β nucleotide insertions were synthesized and purified by polyacrylamide gel electrophoresis (PAGE). A DNA fragment that contained (CAG)₂₀ without any DNA base lesions was used as a marker for DNA fragment analysis.

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

Cell extracts were made as described previously (53). Briefly, pol β null (pol $\beta^{-/-}$) and wild-type pol β (pol $\beta^{+/+}$) mouse embryonic fibroblast (MEF) cell lines were grown to near confluence. Cells were harvested and lysed and cell lysates were centrifuged at 12 000 rpm for 30 min to obtain whole-cell extracts that were subsequently dialyzed into BER reaction buffer containing 50 mM Tris-HCl, pH 7.5, 50 mM KCl, 0.1 mM EDTA, 0.1 mg/ml bovine serum albumin and 0.01% Nonidet P-40. FEN1 deficient cell extracts were prepared by incubating wild-type MEF cell extracts (100 μ g) with 5 μ g FEN1 polyclonal antibodies at 4°C overnight. Protein-antibody complexes were precipitated with protein A agarose beads (Pierce-Thermo Scientific, Rockford, IL, USA) for FEN1 depletion. Control cell extracts were prepared under the same conditions by incubating cell extracts with 5 μ g IgG. Levels of FEN1, pol β and tubulin (loading control) were determined by immunoblotting. BER was performed by incubating 60 μ g cell extracts with the (CAG)₂₀-THF or a (CAG)₂₀-uracil-containing substrate (25 nM). A reaction (25 μ l) was assembled in BER reaction buffer with 5 mM MgCl₂, 2 mM ATP and 50 μ M dNTPs. BER reactions were also performed by incubating substrates with the pol $\beta^{-/-}$ cell extracts supplied with increasing concentrations of purified pol β . Substrates were pre-incubated with 50 nM APE1 at 37°C for 30 min to generate ssDNA break intermediates. The reaction mixture was assembled on ice and incubated at 37°C for 30 min. The reactions were terminated by transfer to 95°C for 5 min and the reaction mixture was subsequently subjected to protease K digestion at 55°C for 30 min. DNA was precipitated and

dissolved into buffer containing 95% formamide and 2 mM EDTA. Substrates and repair products were separated by 15% urea-denaturing PAGE and detected using a Pharos FX Plus PhosphorImager from Bio-Rad. Repair products were excised from gels and eluted at room temperature overnight with TE buffer (10 mM Tris-HCl, pH 7.5 and 1 mM EDTA) and then were precipitated and dissolved in TE buffer. Twenty nanograms of the repaired DNA was used for subsequent DNA fragment analysis. All substrates were ³²P-radiolabeled at the 5'-end of the strand that contained a deoxyuridine or a THF residue.

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

In vitro BER of a native or modified sugar in the context of (CAG)₂₀ repeats was performed with purified UDG, APE1, pol β , FEN1, LIG I and (CAG)₂₀-containing substrates with a deoxyuridine or a THF residue. The 20- μ l reaction was reconstituted with the indicated concentrations of substrates and BER enzymes in BER reaction buffer that contained 5 mM MgCl₂, 2 mM ATP and 50 μ M dNTPs. Reaction mixtures were assembled on ice and incubated at 37°C for 15 min. Reactions were then terminated by transfer to 95°C for 10 min in buffer containing 95% formamide and 10 mM EDTA. Substrates and products were separated by 15% urea-denaturing PAGE and detected by PhosphorImager. Repair products were isolated for size analysis. Substrates were ³²P-radiolabeled at the 5'-end of the strand with a deoxyuridine or a THF residue. BER mediated by pol β hairpin bypass was reconstituted by incubating purified pol β , FEN1 and LIG I with 25 nM substrates that contained a hairpin in the template strands with or without a downstream flap/hairpin. Reactions were performed in BER reaction buffer with 5 mM MgCl₂ and 50 μ M dNTPs. To isolate repair products, the template or upstream strands of the substrates were biotinylated at the 3'- or 5'-end. Repair products were incubated with Pierce avidin agarose beads (Pierce-Thermo Scientific, Rockford, IL, USA) in binding buffer that contained 0.1 M phosphate, 0.15 M NaCl, pH 7.2 and 1% Nondet P-40 at 4°C for 2 h with rotation. The agarose beads were centrifuged at 5000 rpm for 1 min and were washed with the binding buffer three times. The repaired strands were dissociated and separated from their template strands with 0.15 M NaOH and centrifugation at 5000 rpm for 2 min. Repaired strands were then precipitated and dissolved in TE buffer for subsequent amplification by polymerase chain reaction (PCR), cloning and DNA sequencing (Supplementary Figure S1).

Enzymatic activity assays

Pol β DNA synthesis was measured using substrate that contained (CAG)₂₀ repeats with a THF alone or substrates that contained a template hairpin with or without a downstream 5'-THF-flap/hairpin. The DNA synthesis was measured at 37°C in a 20- μ l reaction mixture that contained BER reaction buffer with 5 mM MgCl₂ and 50 μ M dNTPs. FEN1 cleavage activity was examined in the same buffer with 5 mM MgCl₂ and 50 μ M dNTPs in

the absence or presence of increasing concentrations of pol β at 37°C for 15 min. Substrates and products were separated by 15% urea-denaturing PAGE and detected by a PhosphorImager.

Probing of hairpin structures by Mung Bean Nuclease digestion

The formation of hairpin structures by CAG and CTG repeats was probed by incubating 2 U Mung Bean Nuclease with 200 nM substrates that contained (CAG)₂₀ or (CTG)₂₀ repeats with a THF embedded in the first CAG or CTG or in a random DNA sequence. Hairpins present in the substrates containing a preformed template (CTG)₇ and (CTG)₂₅ hairpin were also probed by the enzyme digestion. Substrates were pre-incubated with 10 nM APE1 at 37°C for 30 min to generate a ssDNA breakage intermediate. The 10- μ l reaction mixture was assembled by incubating 2 U of Mung Bean Nuclease with 200 nM substrates in reaction buffer containing 30 mM sodium acetate (pH 4.6), 50 mM NaCl, 1 mM zinc acetate and 0.01% Triton X-100. The reaction was incubated at 37°C for 1, 2, 3, 5 and 8 min and subsequently subjected to protease K digestion at 55°C for 30 min to remove Mung Bean Nuclease. Substrates and products were separated by 15% urea-denaturing PAGE and detected by a PhosphorImager.

Sizing of repair products with CAG repeats by DNA fragment analysis

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

Cloning and sequencing of repair products

The repair products resulting from pol β hairpin bypass were amplified by PCR under the following conditions: 95°C for 5 min, 1 cycle; 95°C for 1 min, 54°C for 30 s and 72°C for 1 min, 25 cycles; 72°C for 10 min, 1 cycle. PCR products were then ligated into the pCR 2.1-TOPO vector (Invitrogen/Life Technologies Corp., Grand Island, NY, USA). Recombinant plasmids that contained repair products were then transformed into SURE 2 Supercompetent Cells (Stratagene, Santa Clara, CA, USA). The plasmids were isolated and purified by the Miniprep Kit from Qiagen (Valencia, CA, USA) and were subjected to DNA sequencing. PCR sequencing reactions were performed with the Big Dye Terminator version 3.1 Cycle Sequencing Kit from Applied

Biosystems/Life Technologies under the following conditions: 94°C for 15 s, 50°C for 10 s and 60°C for 4 min, 35 cycles. The sequences of repair products were determined using ABI Genetic Analyzer 3130 version 1 (Applied Biosystems/Life Technologies Corp) by the Florida International University DNA Sequencing Core Facility and analyzed with Mac Vector version 12.0.2 software (MacVector Inc., Cary, NC, USA).

RESULTS

Pol β promotes CAG repeat deletion during BER

Pol β , a central component of BER, removes a native deoxyribose phosphate residue and fills in a single-nucleotide gap to fulfill the efficient single nucleotide BER (42). Pol β can also perform multi-nucleotide gap-filling synthesis and strand displacement synthesis that leads to the relatively inefficient long-patch BER (42). Pol β multi-nucleotide gap-filling synthesis was found to promote CAG repeat expansion by producing extra repeat units (27), whereas its slippage synthesis is suggested to be involved in repeat deletion. Thus the enzyme plays a pivotal role in modulating TNR instability. To further determine how pol β could be involved in TNR deletion, we initially studied deletions in the cell extract-based BER of a THF embedded at the first CAG of a (CAG)₂₀ repeat-containing substrate. Experiments were performed using pol $\beta^{-/-}$ (Figure 1A, panel b), pol $\beta^{+/+}$ cell extracts (Figure 1A, panel c) or pol $\beta^{-/-}$ cell extracts complemented with 10 nM purified pol β (Figure 1A, panel d). The results revealed that significant amounts of deletion products formed with pol $\beta^{+/+}$ cell extracts and pol $\beta^{-/-}$ cell extracts complemented with purified pol β (Figure 1A, panels c and d). However, only a small amount of deletion product was produced by pol $\beta^{-/-}$ cell extracts (Figure 1A, panel b). The amounts of deletion products increased by 3- to 4-fold by the presence of pol β (Figure 1A, panel e). The repair mainly resulted in (CAG)₁₉- and (CAG)₁₈-repeat containing products with one or two repeat deletion (Figure 1A, panels c and d). To further confirm the role of pol β in CAG repeat deletion, we determined the deletions in a reconstituted BER by incubating the (CAG)₂₀-containing substrate with purified APE1, FEN1, LIG I and increasing concentrations of pol β . Consistent with the results from the pol β cell extracts, the reconstituted BER also resulted in deletion of one or two CAG repeats with the amounts of the products slightly increased with increasing concentrations of pol β (Figure 1B, panels b, c and d). Thus, the results indicate that pol β promotes CAG repeat deletion. To exclude the possibility that deletion products could result from PCR artifacts due to Taq polymerase DNA slippage, a (CAG)₂₀ repeat-containing marker without any base lesions was PCR-amplified and subjected to sizing analysis under the same conditions as the ones used for repair products. The results showed that little deletion product was from the amplified marker (Figure 1A and B, panel a), indicating that repeat deletions were from BER rather than from PCR artifacts.

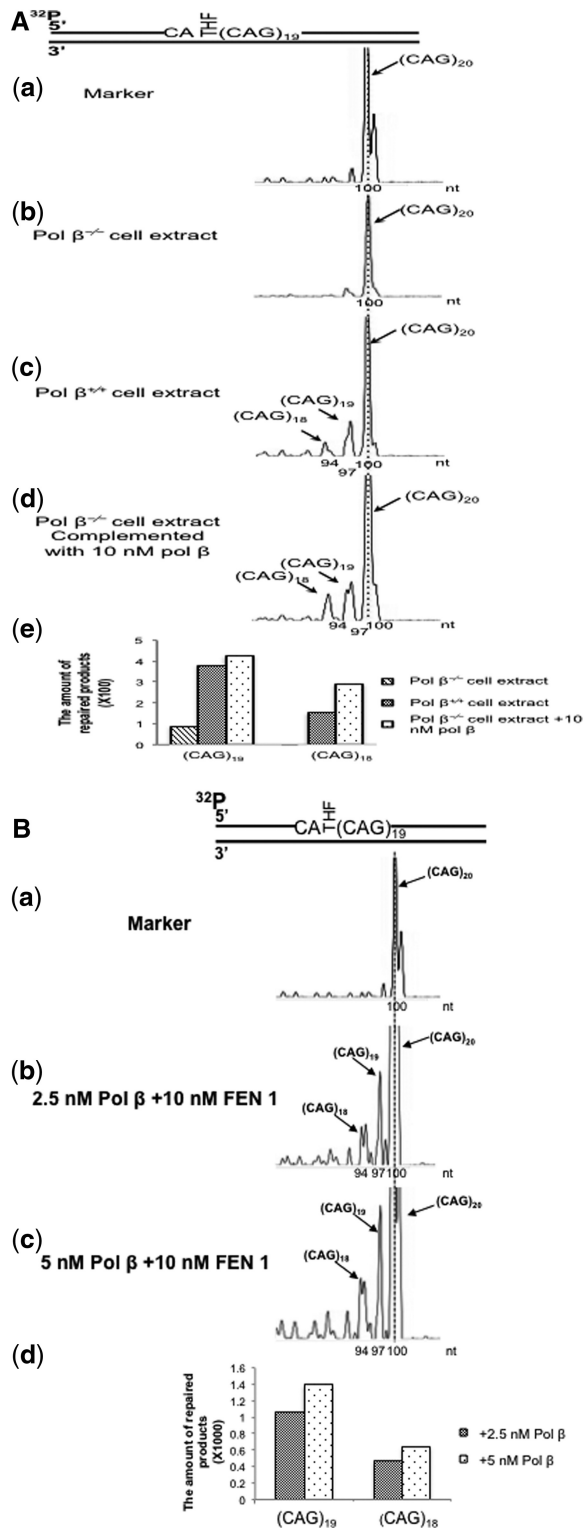


Figure 1. CAG repeat deletion during BER. (A) Repair of an abasic site in the context of CAG repeats was examined with a (CAG)₂₀-containing substrate with a THF embedded in the first CAG. BER reactions were reconstituted with cell extracts from pol β^{+/+}, pol β^{-/-} MEFs and pol β^{-/-} MEF cell extracts complemented with 10 nM of purified wild-type pol β. Gel-isolated repair products were amplified by PCR with 6-carboxyfluorescein-labeled primers. PCR products were separated by capillary electrophoresis and their size was determined by DNA fragment analysis. The products are illustrated as different individual peaks. The height of a peak represents the abundance of

Formation of CAG and CTG repeat hairpin structures during BER

It is proposed that TNR deletion involves the formation of hairpin structures on the template DNA strand and hairpin bypass by DNA polymerases during DNA replication (11,54). Thus, it is conceivable that hairpins may also form on the template strands during BER leading to pol β hairpin bypass DNA synthesis. To test this possibility, we initially determined the formation of hairpin structures in the damaged and template strands of the substrates containing (CAG)₂₀ or (CTG)₂₀ with a THF residue, using Mung Bean Nuclease. The nuclease preferentially makes cleavages at a single-strand loop region and the sites with mismatched base pairs in the stem region of a hairpin. A specific cleavage pattern of Mung Bean Nuclease on CAG or CTG repeats indicates the formation of a hairpin with a specific size. The enzyme cleavage on the template strand of the (CAG)₂₀-repeat substrate resulted in products with 20, 22, 27, 31, 34 and 37 nt, respectively (Figure 2A, left panel). The cleavage pattern indicated that a (CTG)₇ hairpin formed next to the 5'-side 20 nt-random sequence of the template strand. The hairpin consisted of a one-CTG loop and a six-CTG repeat stem as indicated (Figure 2A). The enzyme cleavage on the damaged strand mainly generated products with 25, 29, 32, 34 and 36 nt, respectively (Figure 2A, right panel), indicating the formation of a (CAG)₇ repeat hairpin next to the 3'-side 20 nt-random sequence of the damaged strand. The hairpin was composed of a (CAG)₃-repeat loop and a stem with (CAG)₄ repeats (Figure 2A). The nuclease cleavage on the template strand resulted in products with 25, 29, 32, 34, 37 and 40 nt (Figure 2B), respectively. The cleavage pattern demonstrated the formation of a (CAG)₈ repeat hairpin next to the 5'-side 20-nt random sequence of the template strand. The hairpin contained a (CAG)₄-repeat loop region and a stem region with (CAG)₄ repeats (Figure 2B). The nuclease digestion of a substrate containing random sequence failed to generate any cleavage products (Supplementary Figure S2A). The results indicated that the formation of hairpins on both the damaged and template strands during BER is repeat sequence-dependent.

Pol β bypasses a template hairpin effectively, thereby facilitating CAG repeat deletions during BER

Since TNR deletion is proposed to be mediated by DNA polymerase bypass of a template hairpin generated during DNA replication (54), a similar mechanism could also be adopted by pol β to mediate TNR deletion during BER. To test this hypothesis, pol β DNA synthesis was

Figure 1. Continued

repair products. The sizes of repair products are illustrated in nucleotides and in number of repeat units. (B) BER reactions were reconstituted by incubating 25 nM substrate radiolabeled at the 5'-end of its damaged strand with pol β (2.5 and 5 nM), 50 nM APE1, 10 nM FEN1 and 5 nM LIG I along with 50 μM dNTPs under the conditions described in the 'In vitro BER in mouse embryonic fibroblast cell extracts' of 'Experimental Procedures'.

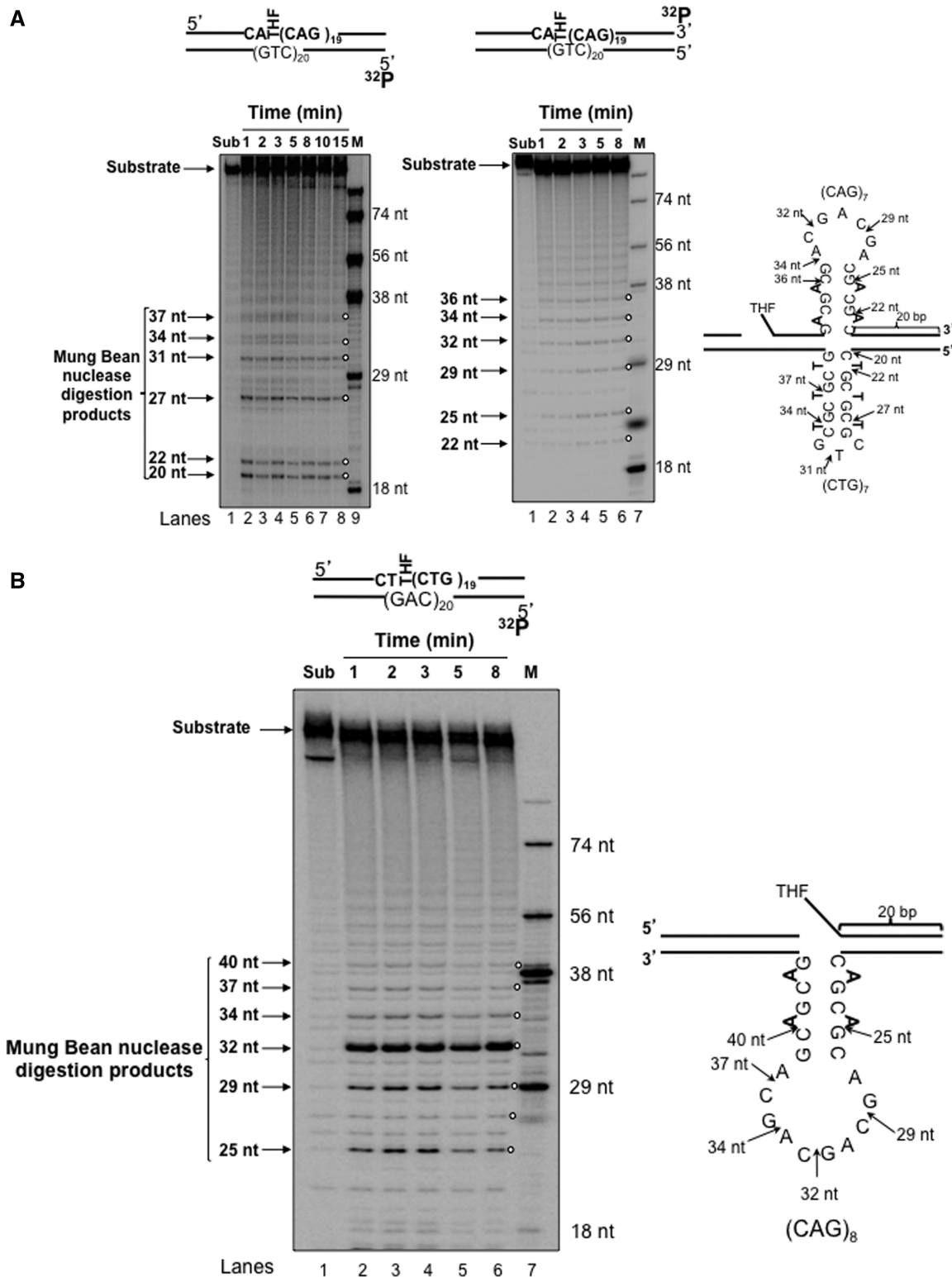


Figure 2. Formation of trinucleotide repeat hairpins during BER. (A) Formation of a template hairpin on the (CAG)₂₀ repeat substrate with a THF residue was probed by Mung Bean Nuclease digestion (left panel). The substrate was radiolabeled at the 5'-end of its template strand and was precut by 10 nM APE1. Subsequently, the substrate was incubated with 2 U of Mung Bean Nuclease at 1-, 2-, 3-, 5-, 8-, 10- and 15-min time intervals (lanes 2–8). Lane 1 represents the undigested substrate. Lane 9 represents synthesized size markers (M) with 18, 29, 38, 56 and 74 nucleotides, respectively. Hairpins formed on the damaged strand were probed by the enzyme digestion of the substrate radiolabeled at the 3'-end of the strand (right panel, lanes 2–6). Lane 1 represents the undigested substrate. Lane 7 represents size markers (M) with 18, 29, 38, 56 and 74 nucleotides, respectively. (B) The template strand that contained (CTG)₂₀ repeats was labeled at its 5'-end. The substrate was incubated with 2 U of Mung Bean Nuclease at 1-, 2-, 3-, 5- and 8-min time intervals, respectively (lanes 2–6). Lane 1 represents the undigested substrate. Lane 7 represents markers (M). For all the experiments, 200 nM of substrate was used. Arrows indicate the major digestion products. The substrates are illustrated schematically above the gel. A hairpin deduced by a specific nuclease cleavage pattern and the nuclease digestion sites is illustrated schematically beside the gel.

examined with substrates containing a template hairpin consisting of a 6-nt loop and a 15-nt perfect base-paired stem (Figure 3A and B, lanes 2–4) or a template hairpin with (CAG)₇ (Figure 3A and B, lanes 6–8) or (CTG)₇ (Figure 3A and B, lanes 11–13) or (CTG)₂₅ (Figure 3A and B, lanes 16–18). These substrates mimic the BER intermediates containing a small or large template hairpin with or without a 5'-downstream flap that result from APE1 5'-incision and DNA slippage. The formation of the hairpins on both the template and damaged strands of these substrates was verified by Mung Bean Nuclease digestion (Supplementary Figure S2B–S2D). Substrates were radiolabeled at the 5'-end of the upstream primers for characterizing any effects of downstream flaps/hairpins on pol β primer extension for its hairpin bypass. With all the substrates, pol β exhibited efficient multi-nucleotide synthesis (illustrated schematically by red lines) indicating that it readily bypassed the template hairpins with varying sizes (Figure 3A and B, lanes 2, 6, 11 and 16). For both (CAG)₇ and (CTG)₇ template hairpins, pol β inserted five to seven repeat units (Figure 3A, lanes 6, 7, 11 and 12), indicating that the enzyme passed through almost the entire length of the hairpins. In contrast, pol β only inserted 3–10 repeats in bypassing (CTG)₂₅ hairpin (Figure 3A and B, lanes 16 and 17), suggesting that the enzyme can skip over the hairpin by passing through its portion. Interestingly, the presence of FEN1 did not affect pol β DNA synthesis on the substrates without a downstream flap, but inhibited the DNA synthesis on the substrates with a 5'-flap (compare lanes 2 and 3, 6 and 7, 11 and 12, and 16 and 17 of Figure 3B). These results suggested that FEN1 suppressed pol β hairpin bypass synthesis, possibly by competing with the polymerase to bind a flap. To determine whether pol β hairpin bypass could result in repeat deletions, BER reactions were reconstituted using the same set of template hairpin-containing substrates and purified pol β, FEN1 and LIG I. The results showed that pol β bypass of the perfect base-paired hairpin and the (CTG)₂₅ hairpin led to repair products with deletions of 4–39 and 3–63 nt (1–21 repeats), respectively, which were identified by DNA sequencing (Figure 6A and B). Consistent with the observation, pol β bypass of a (CAG)₂₅ hairpin also inserted 1–6 repeats and resulted in 3–60 nt (1–20 repeats) (Supplementary Figure S3). Surprisingly, pol β bypass of (CTG)₇ and (CAG)₇ hairpins failed to cause any repeat deletion (Figure 6A). This indicated that it passed through only a portion of the large hairpins and the entire small hairpins. Thus, pol β can bypass a large hairpin directly leading to repeat unit loss, but can faithfully copy almost all the repeats that constitute a small hairpin. However, other repair steps that can shorten the repeats may also contribute to repeat deletion. During BER, this could result from FEN1 flap cleavage activity.

FEN1 promotes CAG repeat deletion during BER

FEN1 has been found to play controversial roles in modulating TNR stability in replication and BER. FEN1 cleavage on a 5'-TNR-flap during DNA replication

is crucial for preventing TNR expansion (24,55,56). However, FEN1 alternate cleavage of a 5'-short flap attached to a hairpin during long-patch BER, promotes CAG repeat expansion (27). FEN1 flap cleavage may also be involved in repeat deletion during BER by shortening repeat length. To test this possibility, we initially examined FEN1 cleavage during BER of a THF embedded in (CAG)₂₀ repeats. We found that with increasing concentrations of pol β (2.5, 5 and 10 nM), FEN1 (5, 10 and 25 nM) cleaved up to four CAG repeats in the presence of pol β (Figure 4A, lanes 4–12). Pol β alone failed to give any cleavage products (Figure 4A, lane 3) indicating the products were FEN1 specific. To further determine if the cleavage products may result from FEN1 cleavage in the presence of a template hairpin, we examined its cleavage on a 5'-THF-(CAG)₇ or 5'-THF-(CTG)₇ or 5'-THF-(CAG)₂₅ or 5'-THF-(CTG)₂₅ flap/hairpin in the presence of a template (CAG)₇ or (CTG)₇ or (CTG)₂₅ or (CAG)₂₅ hairpin (Figure 4B and Supplementary Figure S3C). We found that FEN1 removed up to five or six CAG or CTG repeats from the (CAG)₇ or (CTG)₇ flap/hairpin in the presence of pol β (5–25 nM) (Figure 4B, lanes 1–6) which was similar to its cleavage during BER. Interestingly, FEN1 removed up to seven to eight repeats from the (CAG)₂₅ or (CTG)₂₅ flap/hairpin (Figure 4B, lanes 7–9 and Supplementary Figure S3C) suggesting that a large hairpin allowed FEN1 to remove more repeats for larger deletion. To further determine if FEN1 cleavage is essential for repeat deletion, we examined repeat deletion in the absence and presence of FEN1 with a substrate containing a deoxyuridine that substituted the first guanine of (CAG)₂₀ repeats (Figure 5). We reasoned that removal of the base lesion by UDG would create an abasic site that results in a native 5'-sugar phosphate and a single-nucleotide gap, and these can be directly removed by pol β deoxyribose phosphate lyase and filled by pol β gap-filling synthesis without the need of FEN1 cleavage. Thus, no deletion products should be detected because the repeat length cannot be shortened by pol β. However, if FEN1 is included in the repair, it could remove short CAG repeats from the 5'-end of a downstream hairpin, leading to shortening of repeats and deletions. In support of this idea, we identified a small amount of deletion product in the absence of FEN1 (Figure 5, panel b), but significant amounts of deletion products in the presence of 10 nM FEN1 (Figure 5, panel c). Quantitative analysis demonstrated that FEN1 increased the deletion by ~3-fold (Figure 5, panel d). In contrast, reduction of FEN1 level in pol β^{+/+} MEF cell extracts by ~5-fold through immuno-depletion (Supplementary Figure S4) decreased the amount of deletion products only by ~2-fold (Figure 5, panels f–i). The immuno-depletion of FEN1 did not affect pol β level in the cell extracts (Supplementary Figure S4) indicating that reduction of deletion products was specifically from FEN1 deficiency. Thus, FEN1 is not essential for, but promotes TNR deletion. Our results indicate that FEN1 plays an important role in facilitating CAG repeat deletion by cooperating with pol β hairpin bypass synthesis during long-patch BER.

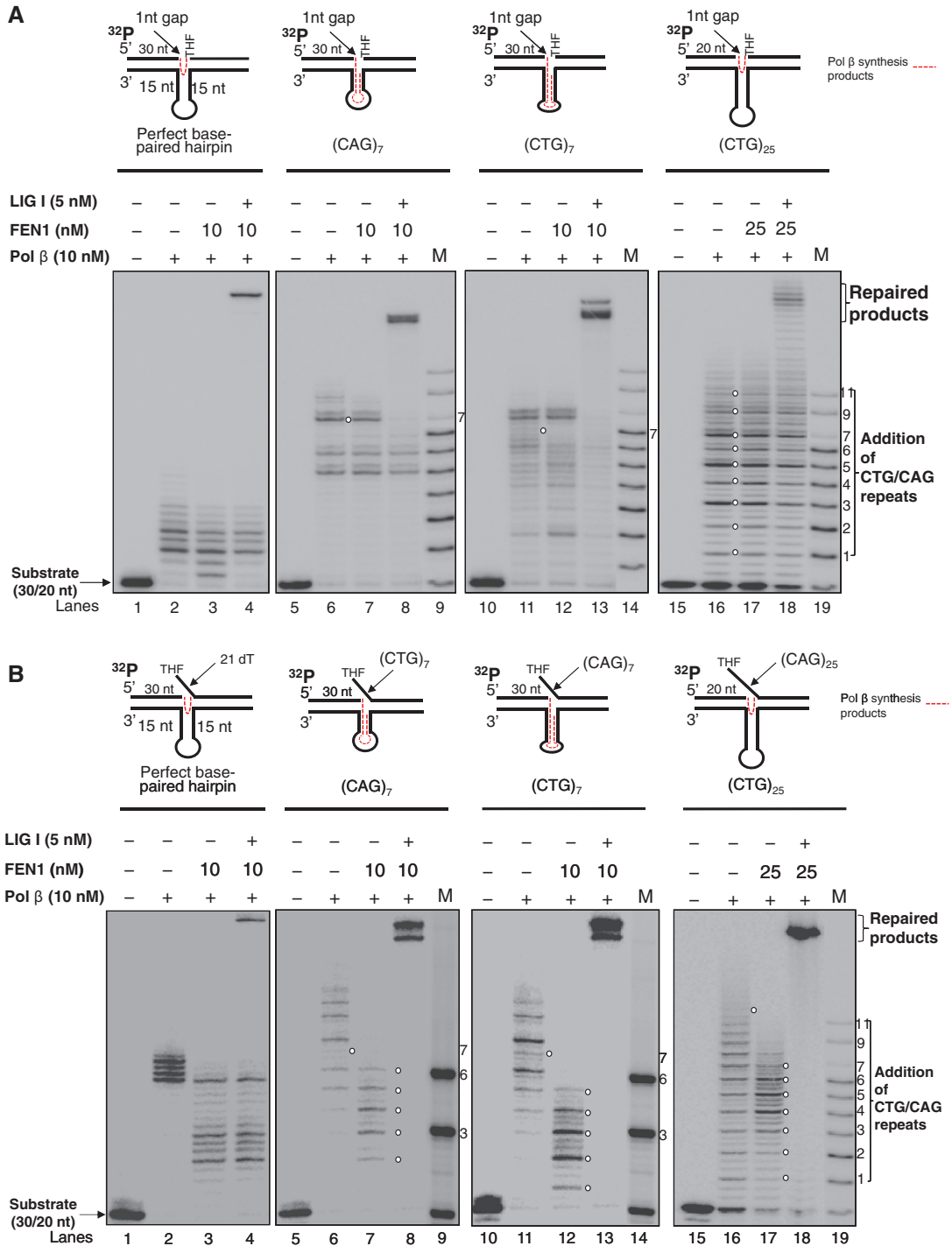


Figure 3. Hairpin bypass synthesis of pol β. (A) Pol β hairpin bypass synthesis and BER were examined in the context of a template (CAG)₇ or a (CTG)₇ hairpin or a (CTG)₂₅ hairpin or a template hairpin with a 6-nt loop and 15-nt perfect base pairs at its stem region. The substrates contained THF residues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

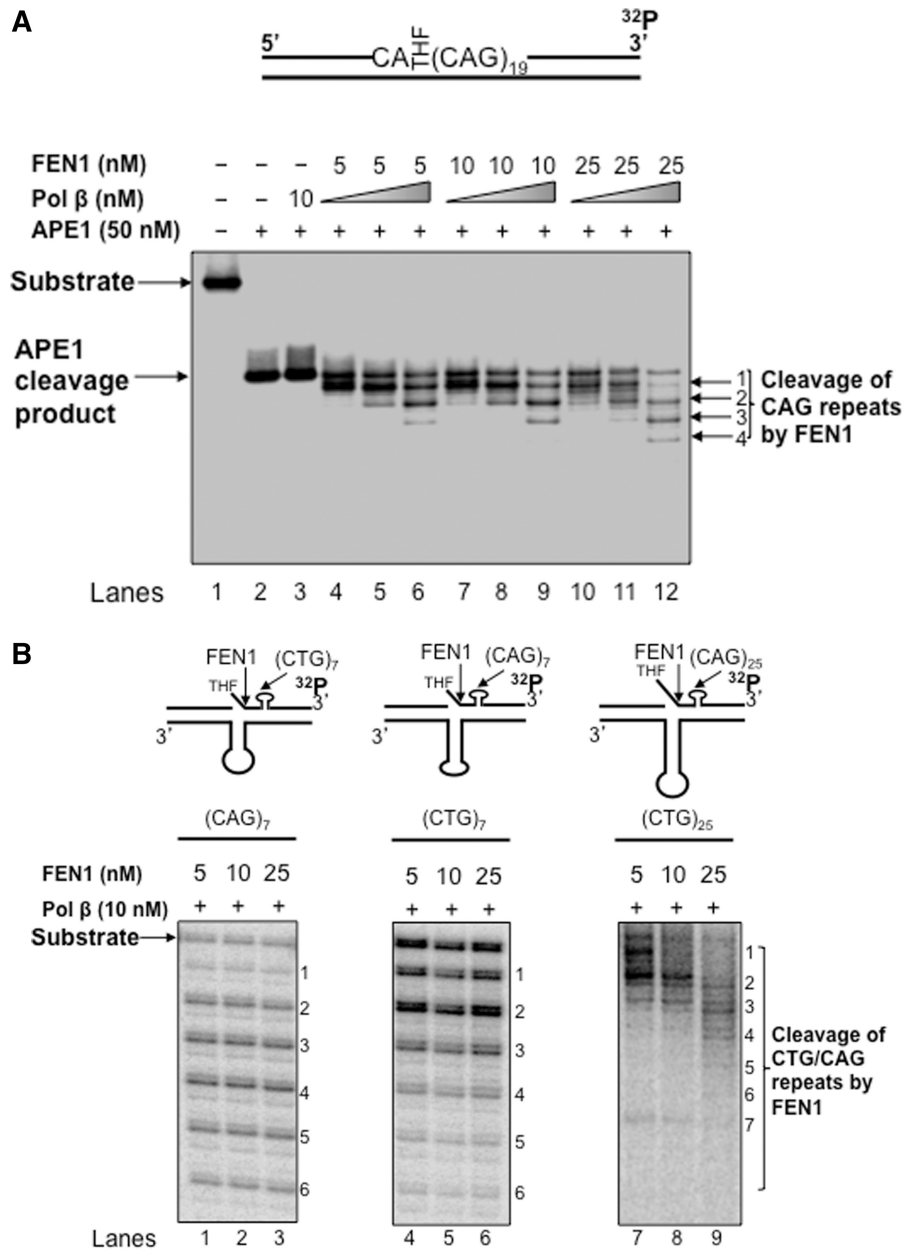


Figure 4. FEN1 cleavage of CAG repeats during BER. (A) FEN1 cleavage activity on a downstream 5'-THF during BER was examined with a substrate that contained a THF in the context of (CAG)₂₀ repeats. Lane 1 corresponds to substrate only. Lane 2 corresponds to reaction mixture with 50 nM APE1. Lane 3 corresponds to reaction mixture with 10 nM pol β in the presence of 50 nM APE1. Lanes 4–12 correspond to reaction mixtures with 50 nM APE1 and increasing concentrations of pol β in the presence of various concentrations of FEN1 as indicated. (B) FEN1 flap cleavage was examined in the context of a template hairpin with a substrate containing a downstream 5'-THF-(CAG)₇ or 5'-THF-(CTG)₇ or 5'-THF-(CAG)₂₅ hairpin/flap. Lanes 1–9 correspond to reaction mixtures with 25 nM substrate and 10 nM pol β in the presence of various concentrations of FEN1 (5, 10 and 25 nM). DNA substrate (25 nM) was used in all the reactions and FEN1 alternate flap cleavage sites are illustrated on the substrates shown above the gels.

DISCUSSION

In this study, for the first time, we discovered that CAG repeats can be deleted during BER (Figures 1, 5 and 6). We provided direct evidence that repeat deletions are generated in both cell extract-based and reconstituted BER (Figure 1) and that this process is accomplished by pol β hairpin bypass synthesis in coordination with FEN1 alternate cleavage. We demonstrated that during BER,

hairpin structures are generated on the damaged and template strands of (CAG)₂₀ repeats (Figure 2). These hairpins govern TNR deletion by cooperatively interacting with pol β hairpin bypass and FEN1 alternate flap cleavage. We discovered that pol β readily bypasses a small portion of a large template hairpin, directly resulting in loss of large size CAG repeats (Figure 3A and B, lanes 16 and 17; Figure 6); however, pol β can pass through almost the entire length of a small template hairpin

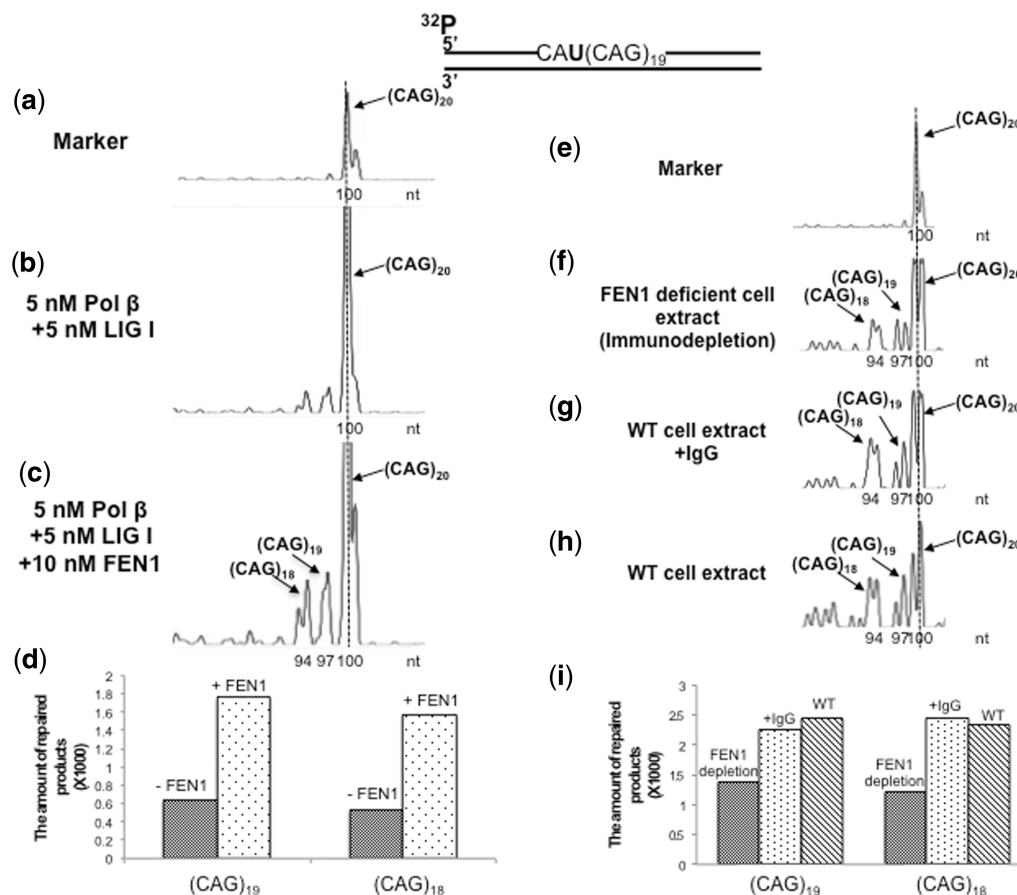


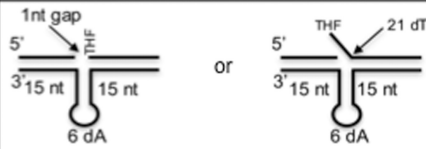
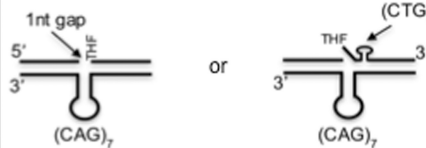
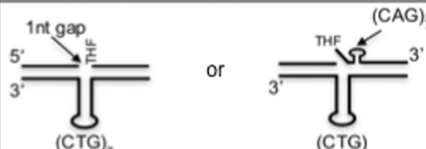
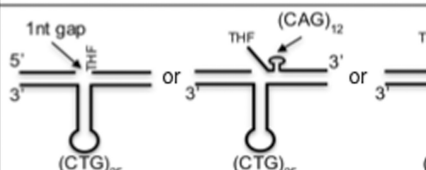
Figure 5. FEN1 promotes limited CAG repeat deletion. Repair of a deoxyridine in the context of (CAG)₂₀ repeats was examined with an oligonucleotide substrate containing a deoxyridine residue that substituted the guanine in the first CAG of (CAG)₂₀ repeats. BER reactions were reconstituted with 5 nM pol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

causing the loss of only one or two repeats (Figure 3A and B, lanes 7, 8, 11 and 12; Figure 6). With the unique hairpin bypass of pol β, repeat deletions will have to be ultimately fulfilled through FEN1 cleavage. In supporting this notion, we found that FEN1 alternate flap cleavage removed up to four CAG repeats in the presence of a small template hairpin (Figure 4A), while pol β hairpin bypass only inserted three or four repeats in the presence of FEN1 (Figure 3B). Thus, FEN1 removes more repeats than those inserted by pol β leaving a gap that led to the formation of a loop containing one or two repeats on the template strand for small deletions. Consistent with this idea, we observed that FEN1 effectively cleaved six to seven repeat units from a pre-formed (CTG)₇ or (CAG)₇ flap/hairpins (Figure 4B) and significantly stimulated small repeat deletions (Figure 5). With longer repeats [i.e. (CAG)₂₅ or (CTG)₂₅] that can form a large template hairpin, pol β skips over the hairpin by passing through a small portion of the hairpin. This allows it to perform a strand-displacement synthesis to displace a longer repeat flap for FEN1 cleavage, thereby causing larger sizes of deletions (Figures 3B, 4B and 6). Our data support a model for TNR deletion during BER

of oxidative DNA base lesions where ssDNA breakage is induced by OGG1 and APE1 (Figure 7). This in turn leads to DNA slippage and the formation of a 5'-downstream hairpin, template hairpin and multi-nucleotide gap. For short repeats, a small template hairpin allows pol β to pass through the large portion of the hairpin. FEN1 removes a short 5'-flap associated with a small downstream hairpin, leaving a gap. This further allows the formation of a loop with one or two repeats on the template leading to a ligatable nick, which upon nick sealing affords a repeat deletion (Figure 7, sub-pathway on the left). For a long repeat tract, ssDNA breakage leads to the formation of a large template hairpin that allows pol β to bypass only one small portion of the hairpin. The polymerase then skips over the hairpin to perform a strand displacement synthesis generating a long repeat flap that is removed by FEN1 leading to a large repeat deletion (Figure 7, right sub-pathway).

In this study, we also provide the first evidence that CAG repeat deletion during BER exhibits a length dependency with a relatively large deletion induced by long repeats that is similar to its expansion (1). This suggests that TNR deletion and expansion share the same

A DNA Sequencing Results of Repair Products

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

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

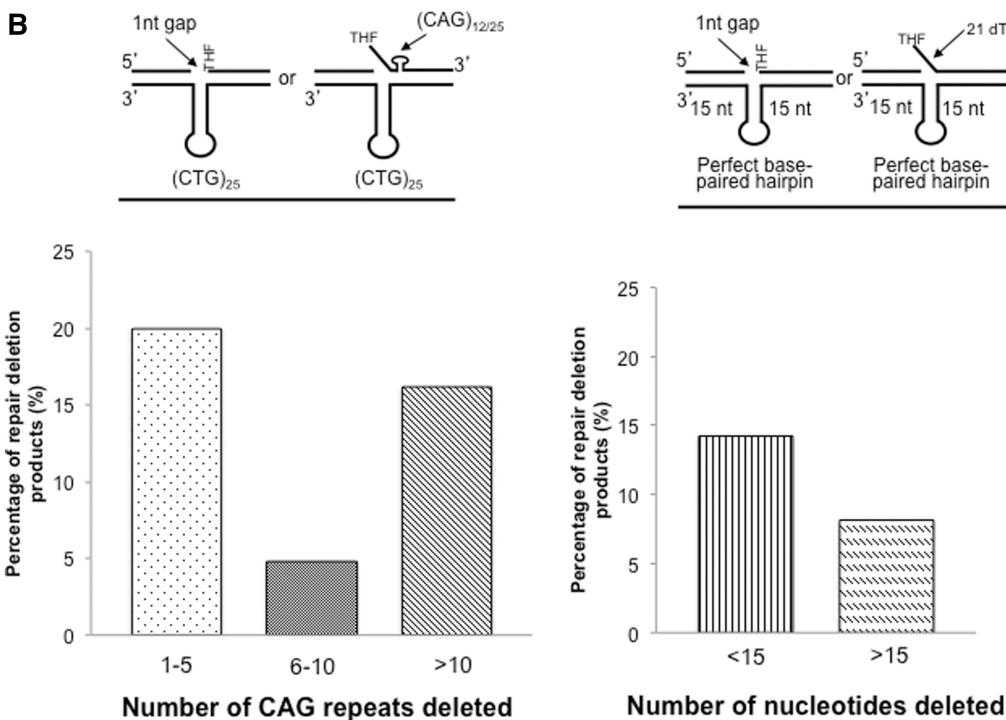


Figure 6. Spectrums of CAG repeat deletions resulting from BER. The products resulting from BER in the context of a small or large template hairpin were isolated with avidin beads, amplified by PCR and cloned into TA vector for DNA sequencing. (A) The number of nucleotides deleted and the percentage of deletions were listed. (B) The spectrums of repeat deletions from the substrate with a (CTG)₂₅ repeat-containing template hairpin and a perfect base-paired template stem-loop structure were illustrated.

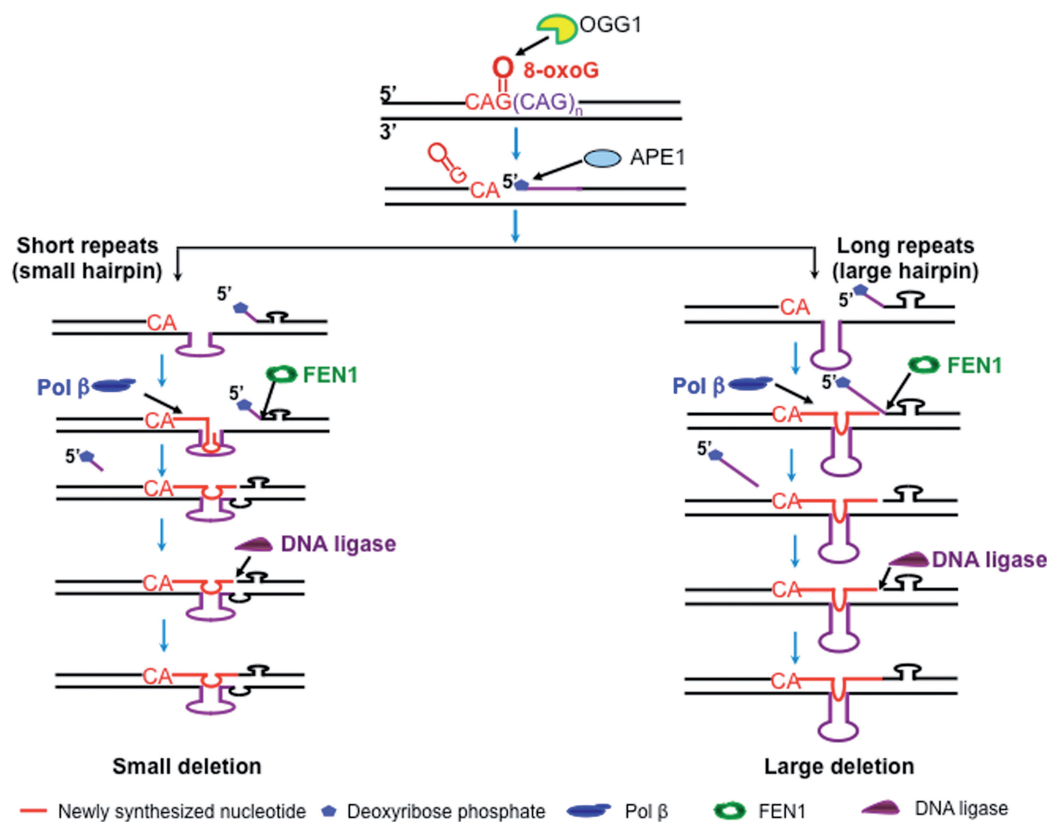


Figure 7. Models for CAG repeat deletion during BER. An oxidized DNA base lesion in the context of CAG repeats is removed by OGG1 leaving an abasic site. APE1 incises the 5'-end of the abasic site resulting in ssDNA breaks. This allows the downstream strand to dissociate from its template forming various sizes of hairpins associated with short CAG repeat flaps. This results in multi-nucleotide gaps that further lead to the formation of a template hairpin. For short repeats, a small template hairpin allows pol β to pass through almost its entire length (left panel). FEN1 then removes one or two repeat units leaving a gap with one or two repeats. This then allows the template to loop out forming a ligatable nick that leads to small repeat deletion (left panel). For long repeats, a large template hairpin can form. This allows pol β to pass through one portion of the hairpin and displace downstream repeats for FEN1 to remove a large flap leading to large repeat deletion (right panel).

underlying molecular basis, i.e. the formation of hairpin structures. Our results indicate that the formation of a series of hairpin structures with varying sizes and stability dependent upon their length appear to govern TNR deletion. Thus, it is of interest to understand the length dependency of TNR instability mediated by BER, as this will further provide a basis for therapeutic purposes of TNR-related human diseases by manipulating the length of TNR through DNA repair.

Our observation of pol β passing through one part or an entire length of a template hairpin suggests that the polymerase can use its strand-displacement synthesis to pass through a hairpin. In this manner, pol β may destabilize hairpin structures, thereby preventing repeat deletion. It is possible that BER cofactors and the proteins from other repair pathways that can modulate pol β strand-displacement synthesis may help prevent or facilitate TNR deletion. For example, BER cofactor HMGB1 (57) that stimulates pol β strand-displacement synthesis may inhibit TNR deletion. However, proteins that can bind and stabilize hairpin structures such as MSH2/MSH3 (25,26) may inhibit pol β strand-displacement within a template hairpin and promote pol β skip-over of the hairpin facilitating repeat deletion. In contrast, DNA helicases that can destabilize hairpins, such as Bloom

syndrome protein and Werner syndrome protein (58–60), may help pol β to pass through the entire length of a hairpin, thus preventing repeat deletion. The mechanisms by which BER proteins modulate TNR deletion through cooperating with the proteins from other repair pathways need to be further elucidated.

Cellular oxidative stress may induce single or multiple DNA base lesions at other locations, such as in the middle of repeat tracks or at the 3'-side of repeats, the locations of base lesions could result in the formation of different sizes of hairpins at specific sites of repeat tracks. This may modulate pol β hairpin bypass and FEN1 cleavage with different efficiency. It is also conceivable that multiple rounds of base lesion repair in the context of TNR could lead to the formation and accumulation of multiple template hairpins that could be processed by BER enzymes in a manner different from the one used for processing a single hairpin, resulting in large size deletion. This is implicated by a recent report showing that an abasic lesion located at various locations of CAG/CTG repeats is repaired by BER with different efficiency (61). The effects of the numbers and positions of DNA base lesions on TNR instability need to be further investigated.

Our study reveals that pol β hairpin bypass synthesis cooperates with FEN1 cleavage of a 5'-hairpin causing

deletion of CAG repeats. Our previous work demonstrated that a robust pol β synthesis generates extra CAG repeats that can be ligated with a 5'-downstream hairpin shortened partially by FEN1 alternate flap cleavage promoting CAG repeat expansion (27). These findings suggest that repeat number synthesized by pol β and that cleaved by FEN1 governs the outcome of TNR instability as to expansion or deletion during BER. We suggest that expansion will occur if pol β synthesizes more repeats than those cleaved by FEN1, whereas deletion will occur if FEN1 removes more repeats than pol β synthesizes.

In summary, in this study, we identified an important role of oxidative DNA damage and BER in CAG repeat deletion for the first time. We provide the first evidence that a base lesion-induced ssDNA breakage can lead to the formation of a hairpin in both the template and damaged strands of TNR tracks and that pol β employs its strand-displacement synthesis to pass through a portion or the entire length of the template hairpin based upon the size of hairpin structures. In contrast, FEN1 uses its alternate cleavage to process the hairpin on the damaged strand. Our results reveal that the cooperation between pol β hairpin bypass and FEN1 alternate cleavage results in small and large deletions. Our work further suggests that TNR deletion during DNA base damage and repair is governed by the interactions between TNR hairpins, pol β unique hairpin bypass and FEN1 alternate cleavage.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Table 1 and Supplementary Figures 1–4.

ACKNOWLEDGEMENTS

We thank Samuel H. Wilson, Laboratory of Structural Biology, National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health (NIH) for generously providing purified BER enzymes and vectors for expressing BER enzymes. We also thank Samuel H. Wilson and Leroy Worth at NIEHS/NIH for critical readings of our manuscript. We thank Zhongliang Jiang and Shyama Ramjagsingh for technical assistance.

FUNDING

National Institutes of Health [ES017476 to Y.L.]. Funding for open access charge: NIH.

Conflict of interest statement. None declared.

REFERENCES

- 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.
- Paulson, H.L. and Fischbeck, K.H. (1996) Trinucleotide repeats in neurogenetic disorders. *Annu. Rev. Neurosci.*, **19**, 79–107.
- McMurray, C.T. (2010) Mechanisms of trinucleotide repeat instability during human development. *Nat. Rev. Genet.*, **11**, 786–799.
- Pearson, C.E., Nichol Edamura, K. and Cleary, J.D. (2005) Repeat instability: mechanisms of dynamic mutations. *Nat. Rev. Genet.*, **6**, 729–742.
- Schildkraut, J.M., Murphy, S.K., Palmieri, R.T., Iversen, E., Moorman, P.G., Huang, Z., Halabi, S., Calingaert, B., Gusberg, A., Marks, J.R. *et al.* (2007) Trinucleotide repeat polymorphisms in the androgen receptor gene and risk of ovarian cancer. *Cancer Epidemiol. Biomarkers Prev.*, **16**, 473–480.
- Nelson, K.A. and Witte, J.S. (2002) Androgen receptor CAG repeats and prostate cancer. *Am. J. Epidemiol.*, **155**, 883–890.
- Lieberman, A.P. and Robins, D.M. (2008) The androgen receptor's CAG/glutamine tract in mouse models of neurological disease and cancer. *J. Alzheimers Dis.*, **14**, 247–255.
- 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.
- 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.
- Kremer, B., Almqvist, E., Theilmann, J., Spence, N., Telenius, H., Goldberg, Y.P. and Hayden, M.R. (1995) Sex-dependent mechanisms for expansions and contractions of the CAG repeat on affected Huntington disease chromosomes. *Am. J. Hum. Genet.*, **57**, 343–350.
- Wells, R.D. (1996) Molecular basis of genetic instability of triplet repeats. *J. Biol. Chem.*, **271**, 2875–2878.
- Panigrahi, G.B., Cleary, J.D. and Pearson, C.E. (2002) In vitro (CTG)_n(CAG) expansions and deletions by human cell extracts. *J. Biol. Chem.*, **277**, 13926–13934.
- McMurray, C.T. (1999) DNA secondary structure: a common and causative factor for expansion in human disease. *Proc. Natl Acad. Sci. USA*, **96**, 1823–1825.
- 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.
- Kovtun, I.V., Liu, Y., Bjaras, 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.
- Kennedy, L., Evans, E., Chen, C.M., Craven, L., Detloff, P.J., Ennis, M. and Shelbourne, P.F. (2003) Dramatic tissue-specific mutation length increases are an early molecular event in Huntington disease pathogenesis. *Hum. Mol. Genet.*, **12**, 3359–3367.
- Tang, W., Dominska, M., Greenwell, P.W., Harvanek, Z., Lobachev, K.S., Kim, H.M., Narayanan, V., Mirkin, S.M. and Petes, T.D. (2011) Friedreich's ataxia (GAA)_n(TTC)_n repeats strongly stimulate mitotic crossovers in *Saccharomyces cerevisiae*. *PLoS Genet.*, **7**, e1001270.
- 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.
- 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.
- 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(CG)_n trinucleotide repeat sequence. *J. Biol. Chem.*, **276**, 16439–16446.
- Mirkin, E.V. and Mirkin, S.M. (2007) Replication fork stalling at natural impediments. *Microbiol. Mol. Biol. Rev.*, **71**, 13–35.
- Spiro, C., Pelletier, R., Rolfsmeier, M.L., Dixon, M.J., Lahue, R.S., Gupta, G., Park, M.S., Chen, X., Mariappan, S.V. and McMurray, C.T. (1999) Inhibition of FEN-1 processing by DNA secondary structure at trinucleotide repeats. *Mol. Cell*, **4**, 1079–1085.

23. Henricksen, L.A., Tom, S., Liu, Y. and Bambara, R.A. (2000) Inhibition of flap endonuclease 1 by flap secondary structure and relevance to repeat sequence expansion. *J. Biol. Chem.*, **275**, 16420–16427.
24. Liu, Y. and Bambara, R. (2003) Analysis of human flap endonuclease 1 mutants reveals a mechanism to prevent triplet repeat expansion. *J. Biol. Chem.*, **278**, 13728–13739.
25. Owen, B.A., Yang, Z., Lai, M., Gajec, M., Badger, J.D. II, Hayes, J.J., Edelman, W., Kucherlapati, R., Wilson, T.M. and McMurray, C.T. (2005) (CAG)_n-hairpin DNA binds to Msh2-Msh3 and changes properties of mismatch recognition. *Nat. Struct. Mol. Biol.*, **12**, 663–670.
26. Lang, W.H., Coats, J.E., Majka, J., Hura, G.L., Lin, Y., Rasnik, I. and McMurray, C.T. (2011) Conformational trapping of mismatch recognition complex MSH2/MSH3 on repair-resistant DNA loops. *Proc. Natl Acad. Sci. USA*, **108**, E837–E844.
27. 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.
28. 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. USA*, **92**, 11019–11023.
29. Pearson, C.E., Ewel, A., Acharya, S., Fishel, R.A. and Sinden, R.R. (1997) Human MSH2 binds to trinucleotide repeat DNA structures associated with neurodegenerative diseases. *Hum. Mol. Genet.*, **6**, 1117–1123.
30. 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.
31. Kovtun, I.V. and McMurray, C.T. (2001) Trinucleotide expansion in haploid germ cells by gap repair. *Nat. Genet.*, **27**, 407–411.
32. Wheeler, V.C., Lebel, L.A., Vrbanac, V., Teed, A., Riele, H. and MacDonald, M.E. (2003) Mismatch repair gene Msh2 modifies the timing of early disease in Hdh(Q111) striatum. *Hum. Mol. Genet.*, **12**, 273–281.
33. Napierala, M., Parniewski, P., Pluciennik, A. and Wells, R.D. (2002) Long CTG/CAG repeat sequences markedly stimulate intramolecular recombination. *J. Biol. Chem.*, **277**, 34087–34100.
34. Pluciennik, A., Iyer, R.R., Napierala, M., Larson, J.E., Filutowicz, M. and Wells, R.D. (2002) Long CTG/CAG repeats from myotonic dystrophy are preferred sites for intermolecular recombination. *J. Biol. Chem.*, **277**, 34074–34086.
35. Jankowski, C. and Nag, D.K. (2002) Most meiotic CAG repeat tract-length alterations in yeast are SPO11 dependent. *Mol. Genet. Genomics*, **267**, 64–70.
36. Bishop, A.J. and Schiestl, R.H. (2000) Homologous recombination as a mechanism for genome rearrangements: environmental and genetic effects. *Hum. Mol. Genet.*, **9**, 2427–2434.
37. Kanaar, R., Hoeijmakers, J.H. and van Gent, D.C. (1998) Molecular mechanisms of DNA double strand break repair. *Trends Cell Biol.*, **8**, 483–489.
38. Sundararajan, R., Gellon, L., Zunder, R.M. and Freudenreich, C.H. (2010) Double-strand break repair pathways protect against CAG/CTG repeat expansions, contractions and repeat-mediated chromosomal fragility in *Saccharomyces cerevisiae*. *Genetics*, **184**, 65–77.
39. Chandok, G.S., Patel, M.P., Mirkin, S.M. and Krasilnikova, M.M. (2012) Effects of Friedreich's ataxia GAA repeats on DNA replication in mammalian cells. *Nucleic Acids Res.*, **40**, 3964–3974.
40. Entezam, A., Lokanga, A.R., Le, W., Hoffman, G. and Usdin, K. (2010) Potassium bromate, a potent DNA oxidizing agent, exacerbates germline repeat expansion in a fragile X premutation mouse model. *Hum. Mutat.*, **31**, 611–616.
41. Kovtun, I.V., Johnson, K.O. and McMurray, C.T. (2011) Cockayne syndrome B protein antagonizes OGG1 in modulating CAG repeat length in vivo. *Aging*, **3**, 509–514.
42. Liu, Y. and Wilson, S.H. (2012) DNA base excision repair: a mechanism of trinucleotide repeat expansion. *Trends Biochem. Sci.*, **37**, 162–172.
43. Hashem, V.I., Pytlos, M.J., Klysiak, E.A., Tsuji, K., Khajavi, M., Ashizawa, T. and Sinden, R.R. (2004) Chemotherapeutic deletion of CTG repeats in lymphoblast cells from DM1 patients. *Nucleic Acids Res.*, **32**, 6334–6346.
44. Hashem, V.I. and Sinden, R.R. (2002) Chemotherapeutically induced deletion of expanded triplet repeats. *Mutat. Res.*, **508**, 107–119.
45. Zhang, Y., Monckton, D.G., Siciliano, M.J., Connor, T.H. and Meistrich, M.L. (2002) Detection of radiation and cyclophosphamide-induced mutations in individual mouse sperm at a human expanded trinucleotide repeat locus transgene. *Mutat. Res.*, **516**, 121–138.
46. Pineiro, E., Fernandez-Lopez, L., Gamez, J., Marcos, R., Surrallés, J. and Velazquez, A. (2003) Mutagenic stress modulates the dynamics of CTG repeat instability associated with myotonic dystrophy type 1. *Nucleic Acids Res.*, **31**, 6733–6740.
47. Lin, Y. and Wilson, J.H. (2007) Transcription-induced CAG repeat contraction in human cells is mediated in part by transcription-coupled nucleotide excision repair. *Mol. Cell Biol.*, **27**, 6209–6217.
48. Hubert, L. Jr, Lin, Y., Dion, V. and Wilson, J.H. (2011) Xpa deficiency reduces CAG trinucleotide repeat instability in neuronal tissues in a mouse model of SCA1. *Hum. Mol. Genet.*, **20**, 4822–4830.
49. 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.
50. Jung, J. and Bonini, N. (2007) CREB-binding protein modulates repeat instability in a *Drosophila* model for polyQ disease. *Science*, **315**, 1857–1859.
51. 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.
52. Kroutil, L.C. and Kunkel, T.A. (1999) Deletion errors generated during replication of CAG repeats. *Nucleic Acids Res.*, **27**, 3481–3486.
53. Biade, S., Sobol, R.W., Wilson, S.H. and Matsumoto, Y. (1998) Impairment of proliferating cell nuclear antigen-dependent apurinic/aprimidinic site repair on linear DNA. *J. Biol. Chem.*, **273**, 898–902.
54. Mirkin, S.M. (2007) Expandable DNA repeats and human disease. *Nature*, **447**, 932–940.
55. Spiro, C. and McMurray, C.T. (2003) Nuclease-deficient FEN-1 blocks Rad51/BRCA1-mediated repair and causes trinucleotide repeat instability. *Mol. Cell Biol.*, **23**, 6063–6074.
56. 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.
57. Prasad, R., Liu, Y., Deterding, L.J., Poltoratsky, V.P., Kedar, P.S., Horton, J.K., Kanno, S., Asagoshi, K., Hou, E.W., Khodyreva, S.N. et al. (2007) HMGB1 is a cofactor in mammalian base excision repair. *Mol. Cell*, **27**, 829–841.
58. Bachrati, C.Z. and Hickson, I.D. (2008) RecQ helicases: guardian angels of the DNA replication fork. *Chromosoma*, **117**, 219–233.
59. Brosh, R.M. Jr, Majumdar, A., Desai, S., Hickson, I.D., Bohr, V.A. and Seidman, M.M. (2001) Unwinding of a DNA triple helix by the Werner and Bloom syndrome helicases. *J. Biol. Chem.*, **276**, 3024–3030.
60. Mohaghegh, P., Karow, J.K., Brosh, R.M. Jr, Bohr, V.A. and Hickson, I.D. (2001) The Bloom's and Werner's syndrome proteins are DNA structure-specific helicases. *Nucleic Acids Res.*, **29**, 2843–2849.
61. Goula, A.V., Pearson, C.E., Della Maria, J., Trottier, Y., Tomkinson, A.E., Wilson, D.M. III and Merienne, K. (2012) The nucleotide sequence, DNA damage location, and protein stoichiometry influence the base excision repair outcome at CAG/CTG repeats. *Biochemistry*, **51**, 3919–3932.