



DNA polymerase θ promotes CAG•CTG repeat expansions in Huntington's disease *via* insertion sequences of its catalytic domain

Received for publication, May 12, 2021, and in revised form, August 23, 2021. Published, Papers in Press, August 30, 2021.

<https://doi.org/10.1016/j.jbc.2021.101144>

Kara Y. Chan^{1,2}, Xueying Li¹, Janice Ortega¹, Liya Gu¹, and Guo-Min Li^{1,2,*}

From the ¹Department of Radiation Oncology, University of Texas Southwestern Medical Center, Dallas, Texas, USA; ²Department of Toxicology and Cancer Biology, University of Kentucky College of Medicine, Lexington, Kentucky, USA

Edited by Patrick Sung

Huntington's disease (HD), a neurodegenerative disease characterized by progressive dementia, psychiatric problems, and chorea, is known to be caused by CAG repeat expansions in the HD gene *HTT*. However, the mechanism of this pathology is not fully understood. The translesion DNA polymerase θ (Pol θ) carries a large insertion sequence in its catalytic domain, which has been shown to allow DNA loop-outs in the primer strand. As a result of high levels of oxidative DNA damage in neural cells and Pol θ 's subsequent involvement in base excision repair of oxidative DNA damage, we hypothesized that Pol θ contributes to CAG repeat expansion while repairing oxidative damage within *HTT*. Here, we performed Pol θ -catalyzed *in vitro* DNA synthesis using various CAG•CTG repeat DNA substrates that are similar to base excision repair intermediates. We show that Pol θ efficiently extends (CAG)_n•(CTG)_n hairpin primers, resulting in hairpin retention and repeat expansion. Pol θ also triggers repeat expansions to pass the threshold for HD when the DNA template contains 35 repeats upward. Strikingly, Pol θ depleted of the catalytic insertion fails to induce repeat expansions regardless of primers and templates used, indicating that the insertion sequence is responsible for Pol θ 's error-causing activity. In addition, the level of chromatin-bound Pol θ in HD cells is significantly higher than in non-HD cells and exactly correlates with the degree of CAG repeat expansion, implying Pol θ 's involvement in triplet repeat instability. Therefore, we have identified Pol θ as a potent factor that promotes CAG•CTG repeat expansions in HD and other neurodegenerative disorders.

The expansion of trinucleotide repeats (TNRs) is associated with a number of neurodegenerative disorders, including Huntington's disease (HD) and myotonic dystrophy (1, 2). Once the expansion exceeds a certain threshold, for example, 35 CAG repeats in HD, it inactivates the expression and/or

alters the function of the affected genes, which leads to disease onset. Despite extensive studies, how the TNR expands is still not fully understood. Elucidating the mechanisms of TNR expansion will significantly impact the therapeutic approaches for diseases caused by this process (3, 4).

Several DNA metabolic pathways, including DNA replication and repair, have been implicated in TNR instability (1, 2, 5–10). A common feature associated with all DNA metabolic reactions is DNA breaks or free DNA ends, which can induce the formation of DNA hairpins within TNRs *via* strand slippage (5, 11, 12). DNA hairpin formation results in TNR expansion (1, 2, 10–12) when these hairpins are not removed (13, 14). Because cell division does not occur in the human brain, CAG/CTG repeat expansions in HD and other neurodegenerative disorders may not be related to DNA replication, but to DNA repair.

Reactive oxygen species (ROS) are a major source of DNA damage in the human brain, and guanine, which is enriched in CAG/CTG repeats, is the favored target of ROS (15). The repair of oxidative adduct 8-oxo-guanine (8-oxoG) by the 8-oxo-guanine DNA glycosylase 1 (OGG1) has been implicated in CAG/CTG repeat expansions, as depleting *Ogg1* abolishes age-dependent CAG repeat expansion in HD mouse models (16). OGG1-initiated base excision repair (OGG1-BER) involves lesion recognition, flipping the lesion from the DNA double helix into the base-binding pocket (active site) of OGG1, and site-specific changes in the DNA structure (17, 18). As a bifunctional DNA glycosylase, OGG1 excises the oxidized guanine base to generate an abasic (AP) site by using its glycosylase activity and cleaves the phosphodiester bond 3' to the AP site by using its AP lyase activity. AP endonuclease 1 cuts the phosphodiester backbone immediately 5' to the AP site. These cleavages result in a one-nucleotide gap, which is filled in by DNA polymerase β (Pol β), followed by DNA ligase III-catalyzed ligation (19). However, the OGG1 AP lyase activity is ~500-fold less efficient than its glycosylase activity (20). In this case, cleavage of the AP site by AP endonuclease 1 leaves a 3'-hydroxyl and 5'-deoxyribose phosphate (5'-dRP) terminus. The latter can be removed by Pol β 's 5'-dRP lyase activity (21). Alternatively, flap-endonuclease 1, which is required for long-patch BER, can remove the 5'-dRP to

* For correspondence: Guo-Min Li, guo-min.li@utsouthwestern.edu.

Present address for Kara Y. Chan: Twist Bioscience, 681 Gateway Blvd, South San Francisco, California 94080, USA.

Present address for Janice Ortega: Autonomous Therapeutics, Inc, 1530E Jefferson Street, Maryland 20852, USA.

Polymerase θ promotes CAG•CTG repeat expansion

produce an ssDNA gap with 1 to 4 nucleotides, depending on the DNA polymerase involved in long-patch BER (19). Regardless, the ssDNA gap generated by BER promotes hairpin formation within CAG repeats *via* strand slippage (16, 22). Subsequently, Pol β fills the DNA gap by efficiently utilizing a CAG hairpin structure as a primer for DNA synthesis, which leads to CAG repeat expansion (23). Consistently, there is a close association between age-dependent somatic CAG repeat expansion and oxidative DNA damage in HD mouse models (24, 25), and both OGG1 and Pol β have been shown to promote CAG repeat expansions (16, 22, 23, 26, 27).

In addition to Pol β , DNA polymerase θ (Pol θ), a low-fidelity family A DNA polymerase (28), participates in BER of oxidative DNA damage (29). Pol θ is known to add single nucleotides to homopolymeric runs at a high rate (30). This is probably related to Pol θ 's large insertion loops (Fig. 1A), particularly insertion 2 in the polymerase thumb domain (31). Recent structural studies have revealed that the presence of insertion 2 creates a big flexible cavity in the DNA-binding surface (32, 33). Strikingly, Pol θ can extend a loop-out-containing primer (32), an activity also associated with Pol β (34). Therefore, we hypothesized that Pol θ is a major polymerase that promotes CAG/CTG repeat expansions during DNA synthesis.

We performed Pol θ -catalyzed DNA synthesis using a series of BER intermediates containing various numbers of CAG/CTG repeats with or without a hairpin. Like Pol β , Pol θ can

effectively extend hairpin primers to stabilize hairpin structures, and it induces large expansions when copying a DNA template that contains CAG/CTG repeats. However, a Pol θ mutant that lacks insertion 2 does not induce CAG/CTG repeat expansions regardless of the primers and templates used, which suggests that insertion 2 is responsible for Pol θ 's error-prone activity. We also found that Pol θ 's chromatin level is significantly higher in HD cells than in non-HD cells and closely correlates with the degree of CAG repeat expansion, which implies the involvement of Pol θ in CAG repeat expansion. Therefore, this study has identified Pol θ as a potent factor that promotes the CAG/CTG repeat expansions that cause HD and other neurodegenerative diseases.

Results

Pol θ extends CAG/CTG hairpin primers *in vitro*

To determine whether Pol θ can extend a CAG or CTG hairpin primer, we conducted *in vitro* DNA synthesis using a limited amount of HeLa nuclear extracts supplemented with a purified 90-kD catalytic Pol θ polypeptide (Fig. S1) (32), a (CAG)₅ or (CTG)₅ hairpin primer that anneals to ssM13mp18(CTG)₁₀ or ssM13mp18(CAG)₁₀ (Fig. 1B), respectively, as described previously (23). The reaction products were fractionated by denaturing PAGE, followed by Southern blot analysis using a probe that specifically

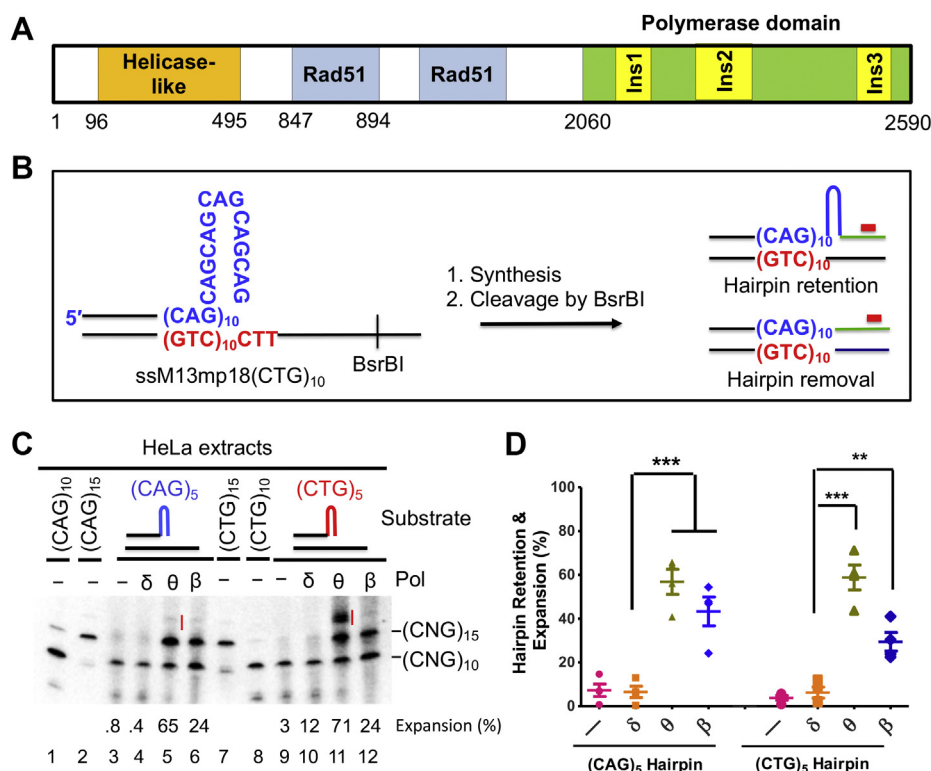


Figure 1. Pol θ promotes CAG/CTG repeat expansion. A, diagram of Pol θ functional domains. The polymerase domain contains three insertion loop motifs. B, illustration of hairpin primer extension-Southern hybridization assay. Red bars represent ³²P-labeled probe. C, hairpin primer extension assay shows Pol θ 's ability to use a CAG or CTG hairpin as a primer for DNA synthesis, leading to hairpin retention and repeat expansions beyond the hairpin. D, quantification and comparison of hairpin retention and repeat expansions by individual polymerases, as indicated. Data are from four independent determinations. ***p < 0.001. Pol θ , DNA polymerase θ ; Pol $\theta\Delta i2$, insertion 2-depleted Pol θ .

recognizes the downstream sequence near the BsrBI site of the newly synthesized strand (Fig. 1B, red bar). Thus, whether the (CAG)₅ or (CTG)₅ hairpin is removed or retained can be readily determined based on its mobility during gel electrophoresis (Fig. 1B).

Consistent with our previous observations (23), incubating the (CAG)₅ hairpin substrate with HeLa nuclear extracts supplemented with the proofreading-active Pol δ , which removes the hairpin structure, generated a major product that migrates similarly to the (CAG)₁₀-containing DNA fragment (Fig. 1C, lane 4). Supplementing the reaction with Pol β yielded two major products: a hairpin-retained (CAG)₁₅ product and a hairpin-removed (CAG)₁₀ product (Fig. 1C, lane 6). When purified Pol θ replaced Pol β or Pol δ in the synthesis reaction, most yielded products were the same as those observed in the Pol β reaction (Fig. 1C, lane 5). This suggests that, like Pol β , Pol θ can effectively use a CAG hairpin as a primer for DNA synthesis. In addition, a slowly migrated minor product (product I) was also observed in both the Pol θ - and the Pol β -containing reactions (Fig. 1C, lanes 5 and 6), which indicates that both polymerases promote CAG repeat expansions beyond the (CAG)₅ hairpin size.

We then tested Pol θ 's ability to extend a (CTG)₅ hairpin primer by using the same *in vitro* DNA synthesis assay. The results revealed that, in addition to the hairpin-retained product, that is, the band migrating at the same place as the (CTG)₁₅ band, a slowly migrated product merged as the major one, whose size is similar to that of product I generated by Pol β and Pol θ during (CAG)₅-primed DNA synthesis. This product is much more abundant in the Pol θ -catalyzed reaction than in the Pol β -catalyzed one (compare product I between lanes 11 and 12), which suggests that Pol θ is more error-prone than Pol β when synthesizing (CTG)_n-primed DNA. Quantitative analysis revealed that the expanded species (hairpin retained and expanded) account for more than 60% of the synthesized products (Fig. 1, C and D). Taken together, these data suggest that Pol θ can effectively synthesize hairpin-primed DNA to promote CAG/CTG repeat expansions.

Pol θ 's insertion 2 is responsible for CAG/CTG repeat expansion

Pol θ 's large insertion 2 is responsible for the polymerase's error-prone nature during DNA synthesis (30, 32, 35). To explore the impact of insertion 2 on the CAG/CTG expansion by Pol θ that we observed, we generated an insertion 2-deleted Pol θ mutant, as described (35). We examined the mutant Pol θ for its ability to synthesize (CAG)₅- or (CTG)₅-primed DNA. Remarkably, no product I was generated by the insertion 2-deleted Pol θ (Pol $\theta\Delta i2$), regardless of whether a (CAG)₅ or (CTG)₅ hairpin primer was used (Fig. 2A, lanes 6 and 13). In addition, the amount of the hairpin-retained product, that is, (CAG)₁₅ or (CTG)₁₅, was much lower in the reactions that contained Pol $\theta\Delta i2$ (Fig. 2A). Pol $\theta\Delta i2$ only generated 8.9% of the (CAG)₅ hairpin-retained and 14.9% of (CTG)₅ hairpin-retained products produced by Pol θ (Fig. 2A, compare lane 5 with lane 6, and lane 12 with lane 13). Quantitative analysis shows that this reduction in the

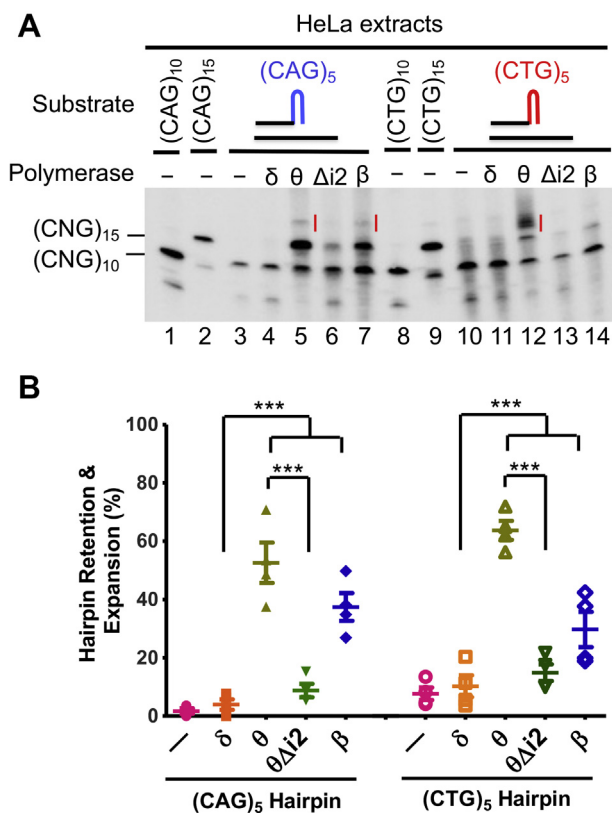


Figure 2. Insertion 2 of Pol θ is responsible for error-prone DNA synthesis. A, hairpin primer extension assay showing less error-prone DNA synthesis with Pol $\theta\Delta i2$ than with Pol θ . B, quantification and comparison of error-prone DNA synthesis by individual DNA polymerases, as indicated. Data are from four independent determinations. *** $p < 0.001$. Pol θ , DNA polymerase θ ; Pol $\theta\Delta i2$, insertion 2-depleted Pol θ .

hairpin-retained and expanded products in Pol $\theta\Delta i2$ -catalyzed reactions is highly significant (Fig. 2B). These results suggest that insertion 2 is indeed responsible for Pol θ -mediated CAG/CTG repeat expansions.

Pol θ -mediated CAG/CTG repeat expansion depends on the repeat length

We next examined Pol θ 's ability to synthesize CAG/CTG repeats by conducting *in vitro* DNA synthesis using a defined DNA synthesis system (Fig. 3A) that contains purified Pol θ , proliferating cellular nuclear antigen (PCNA) and replication factor C (RFC), a 5' ³²P-labeled nonhairpin primer, and M13mp18 derivatives with various lengths of CAG/CTG repeats in the presence or absence of replication protein A, an ssDNA-binding protein that protects ssDNA from nuclease attack and secondary structure formation (36). We observed no expanded DNA products in reactions with high-fidelity polymerases Pol δ and T7, regardless of the DNA templates used, that is, (CAG)₂₀, (CTG)₂₀, (CAG)₂₅, (CTG)₂₅, (CAG)₃₅, (CTG)₃₅, or random DNA sequences (Fig. 3, B–D). Pol $\theta\Delta i2$ -generated products were essentially the same as those produced by Pol δ or T7 polymerase in all reactions (Fig. 3, B–D), which further suggests that insertion 2 is mutagenic.

Polymerase θ promotes CAG•CTG repeat expansion

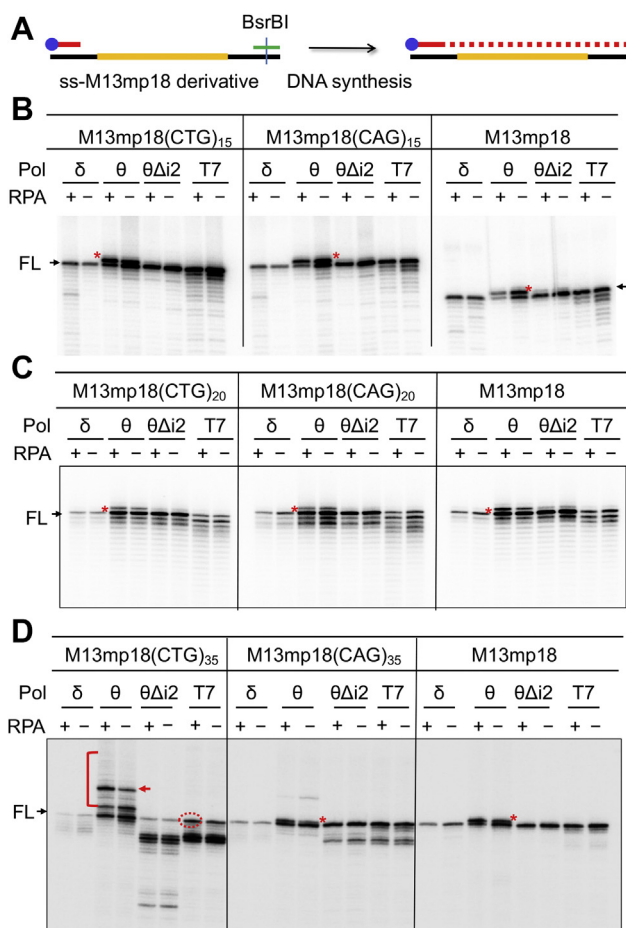


Figure 3. Pol θ 's error-prone activity depends on the repeat length. *A*, diagram of *in vitro* DNA synthesis. ssM13mp18 derivatives with various CAG/CTG repeats (yellow line) were annealed with an oligonucleotide (green line) containing BsrBI restriction recognition sequence and digested with BsrBI before being used as a template for DNA synthesis. Blue sphere represents 32 P-labeling. *B–D*, primer extension assays using various CAG and CTG repeat numbers as templates, as indicated. ssDNA of M13mp18 was used as a nonrepeat control. In panel *B*, the same primer was used in all reactions, which is why the products in all M13mp18 reactions are smaller than those in repeat-containing template reactions; in panels *C* and *D*, the primer used in all M13mp18 reactions was adjusted to make the same length products as in repeat-containing template reactions. The DNA bands labeled with a red arrow or a circle were sequenced. Bands with a red asterisk indicate the Pol θ TdT-generated products. Pol θ , DNA polymerase θ ; RPA, replication protein A; TdT, terminal deoxynucleotidyl transferase.

However, Pol θ -catalyzed products differed from those produced by Pol δ , T7, and Pol $\theta\Delta i2$. First, in addition to full-length products, Pol θ generated a product that is one nucleotide larger than the full-length band in essentially all reactions (see red asterisks in Fig. 3). This one-nucleotide-larger product appears to be unrelated to CAG/CTG repeats, as also seen in reactions with the M13mp18 template. This product is probably derived from Pol θ 's terminal deoxynucleotidyl transferase (TdT) activity (37), which catalyzes the incorporation of single deoxynucleotides into the 3'-OH terminus of ssDNA or dsDNA. Second, when M13mp18-(CAG) $_{35}$ or M13mp18-(CTG) $_{35}$ was used as a template for DNA synthesis, we detected several expanded products in Pol θ -catalyzed reactions (Fig. 3D), especially for the M13mp18(CTG) $_{35}$ template (Fig. 3D, red

bracket). This is probably because a CTG hairpin is easier to form and more stable than a CAG hairpin (11, 23, 38).

To determine the nature of the expansion, we recovered a major Pol θ -expanded band (Fig. 3D, red arrow) and a T7-generated full-length band (Fig. 3D, red circle) from the gel. The DNA samples were PCR-amplified, cloned into a vector, and transfected into a bacterial strain. Plasmid DNA samples were isolated from six clones derived from the T7-catalyzed product and 11 clones derived from the Pol θ -catalyzed band before DNA-sequencing analysis. The results reveal that in six T7-derived clones, four of them showed original 35 repeats (Fig. S2A), and two exhibited 33 repeats, but in 11 Pol θ -expanded clones, one clone contained 40 repeats (Fig. S2B), seven showed 35 repeats, and three demonstrated 34 repeats. These results are consistent with a previous study, which shows that while the majority of (CTG) $_{180}$ clones undergo contractions after replicating for 100 generations, ~20% of these clones still remain the full-length repeats (39). We therefore conclude that the 40 repeats represent the original length of the Pol θ -expanded band, that is, an expansion of five repeats by Pol θ during DNA synthesis (Fig. S2C). We did not observe this repeat length-dependent expansion by Pol θ in reactions with Pol $\theta\Delta i2$ (Fig. 3D), which suggests that insertion 2 is also responsible for Pol θ -induced repeat length-dependent expansion. However, the mechanism by which Pol θ induces repeat expansions in a repeat length-dependent manner is unclear. Notably, DNA substrates used in this assay were gap-containing molecules, and the primer contained no repeat sequences and no hairpin. Thus, Pol θ used an error-free primer to initiate DNA synthesis. When it fills a short repeat gap, for example, 15 or 20 repeats, Pol θ may have sufficient processivity to quickly finish synthesizing the repeats and reach a nonrepeat template, which provides less of a chance for Pol θ to make errors. When Pol θ encounters a template with a large repeat number, such as 35 repeats, which is challenging for all DNA polymerases, it may have to pause a couple of times, and this may allow a repeat hairpin to form, leading to repeat expansion. Future studies are required to test these possibilities.

Mn $^{2+}$ stimulates Pol θ 's error-prone activity for repeat expansion

Divalent metal ion Mn $^{2+}$ alters the structural flexibility of polymerase active sites in favor of error-prone synthesis (40–43). To further determine whether the catalytic flexibility of the Pol θ active site is responsible for the observed CAG/CTG repeat instability, we performed an *in vitro* DNA synthesis assay using oligonucleotide templates that contained 20 or 35 CTG repeats in the presence of Mg $^{2+}$ or Mn $^{2+}$. Although Pol θ 's TdT activity could only add one extra nucleotide when the (CTG) $_{20}$ template was used for synthesis in the presence of Mg $^{2+}$ (Fig. 4A, lane 12, red asterisk), it incorporated multiple nucleotides in the presence of Mn $^{2+}$ (Fig. 4A, lanes 15 and 16, red bracket). We found that Mn $^{2+}$ could also induce Pol δ to generate a product that is one nucleotide (1-nt) bigger than the full-length product (Fig. 4A, lanes 7 and 8). We observed the

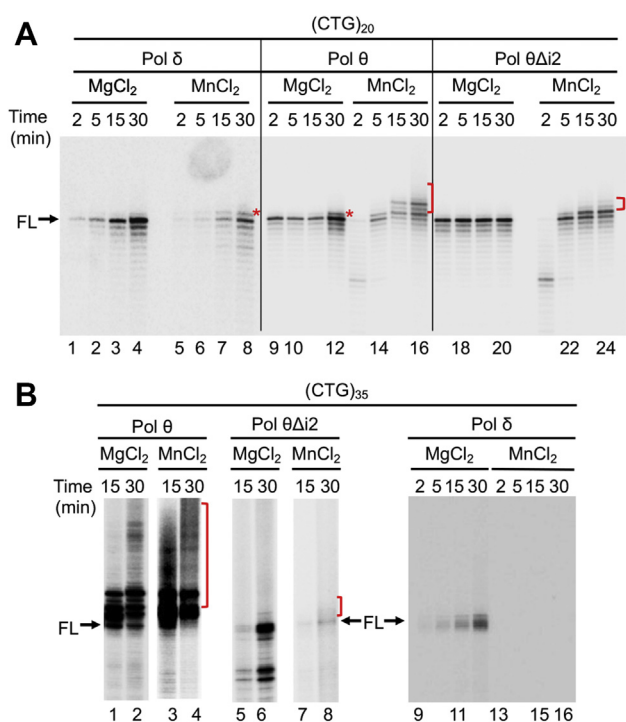


Figure 4. Mn²⁺ augments Pol θ 's repeat expansion capability. *A*, *in vitro* DNA synthesis assay using a synthesized oligonucleotide template containing 20 CTG repeats, which shows that Mn²⁺, but not Mg²⁺, triggers Pol θ to conduct repeat expansions. *B*, *in vitro* DNA synthesis assay showing that Mn²⁺ stimulates Pol θ , but not Pol $\theta\Delta i2$ and Pol δ , to carry out large repeat expansions when using an oligonucleotide containing 35 CTG repeats as a template. Pol θ , DNA polymerase θ ; Pol $\theta\Delta i2$, insertion 2–depleted Pol θ .

same product in Pol $\theta\Delta i2$ -catalyzed reactions in the presence of Mn²⁺ (Fig. 4A, lanes 23 and 24). This product is probably derived from the 1-nt insertion during DNA synthesis, rather than from a TdT activity, particularly for reactions catalyzed by Pol δ , which does not have a TdT activity. Strikingly, when the repeat length of the template was 35, Mn²⁺ further enhanced Pol θ 's error-prone activity, as we found more expanded products (Fig. 4B, red bracket) in reactions with Mn²⁺ than in those with Mg²⁺ (Fig. 4B, compare lane 1 with lane 3, and lane 2 with lane 4). These results indeed suggest that a more flexible active site induced by Mn²⁺ further enhances Pol θ 's error-prone capability. However, Pol $\theta\Delta i2$ generated very few expanded products in the presence of Mn²⁺ (Fig. 4B, lanes 7 and 8), again suggesting that insertion 2 accounts for repeat expansion caused by Pol θ .

We also noted that Pol $\theta\Delta i2$ generated fewer expanded products in replicating a (CTG)₃₅ template (Fig. 4B, lanes 7 and 8) than replicating a (CTG)₂₀ template (Fig. 4A, lanes 23 and 24). We believe that this is related to the repeat length in the DNA template and the relatively high fidelity of Pol $\theta\Delta i2$. Depleting insertion 2 from Pol θ apparently converted an error-prone Pol θ into a high-fidelity Pol $\theta\Delta i2$, as judged by the fact that Pol $\theta\Delta i2$ -generated products are very similar to those produced by Pol δ (Figs. 3 and 4). Our data presented here support the notion that polymerases with no insertion sequence in their catalytic domains, for example, Pol δ and Pol $\theta\Delta i2$, have difficulty synthesizing long triplet repeats, and

that Mn²⁺ adds more difficulty for Pol δ and Pol $\theta\Delta i2$, but not for Pol θ , in synthesizing long triplet repeats (Figs. 3 and 4A). To confirm this, we examined Pol δ 's ability to replicate the (CTG)₃₅ template in the presence of Mg²⁺ or Mn²⁺. The results show that, although Pol δ could generate a limited amount of full-length product in the presence of Mg²⁺ (Fig. 4B, lanes 9–12), it failed to produce any products in the presence of Mn²⁺ (Fig. 4B, lanes 13–16). Taken together, these results suggest that large repeat expansions can only be carried out by polymerases with a large insertion in their catalytic domain, such as Pol θ , and that Mn²⁺ further enhances the error-prone property of Pol θ , but not of polymerases with no insertion in their active sites, such as Pol δ and Pol $\theta\Delta i2$.

Pol θ is highly chromatin enriched in cell lines derived from patients with HD

The data presented above strongly suggest that Pol θ could be an important driving factor for CAG/CTG repeat expansion in HD. To explore this possibility, we set out to determine Pol θ 's expression and chromatin binding in cells derived from individuals with and without HD. We obtained seven fibroblast cell lines derived from patients with HD (GM04208, GM04212, GM04210, GM04220, GM04230, GM21756, and GM09197) and two non-HD cell lines (GM04204 and GM02153) from the Coriell Institute for Medical Research. The first five HD lines were from the same family, with GM04204 being an unaffected family member (44). To ensure the cell lines' HD status, we amplified their *HTT* exon 1 sequence that contained CAG repeats by PCR and analyzed the resulting products in a polyacrylamide gel (Fig. 5A). We then calculated the CAG repeat numbers based on the migration distances of known CAG repeat numbers in HD cell lines, as the CAG repeat number (y) is a linear function of the logarithm of its migration distance (x) in gel electrophoresis: $y = -193.37x + 191.44$. The known CAG repeat numbers are 18/17 for GM04204, 45/19 for GM04230, 44/21 for GM04208, 70/15 for GM21756, 180/19 for GM09197, and 32/16 for GM02153 (Table S1). This analysis indeed confirmed the cell lines' HD origin, with GM04212, GM04210, and GM04220 carrying an allele of at least 40 CAG repeats (Fig. 5B and Table S1). We noted that the 180-CAG repeat allele of GM09197 was not detectable in Figure 5A. This is probably because the large repeat number was hard to amplify, as only a small amount of the 70-CAG repeat product was obtained (Fig. 5A).

We then determined Pol θ expression in these cell lines by Western blotting. In general, all cells expressed abundant Pol θ (Fig. 5C, upper panel), but Pol θ levels in whole-cell lysates (WCLs) in several HD cells were significantly higher than in non-HD control GM04204 (Fig. 5D). Strikingly, all HD cells exhibited a chromatin-bound Pol θ level significantly higher than that in non-HD control GM04204 (Fig. 5C, lower panel, and Fig. 5E), which suggests that Pol θ is efficiently recruited to chromatin in HD cells. For example, although GM04204 (Fig. 5C, lane 1) displayed an overall Pol θ level slightly lower than that of GM04210 (Fig. 5C, lane 4) and GM04220 (Fig. 5C,

Polymerase θ promotes CAG•CTG repeat expansion

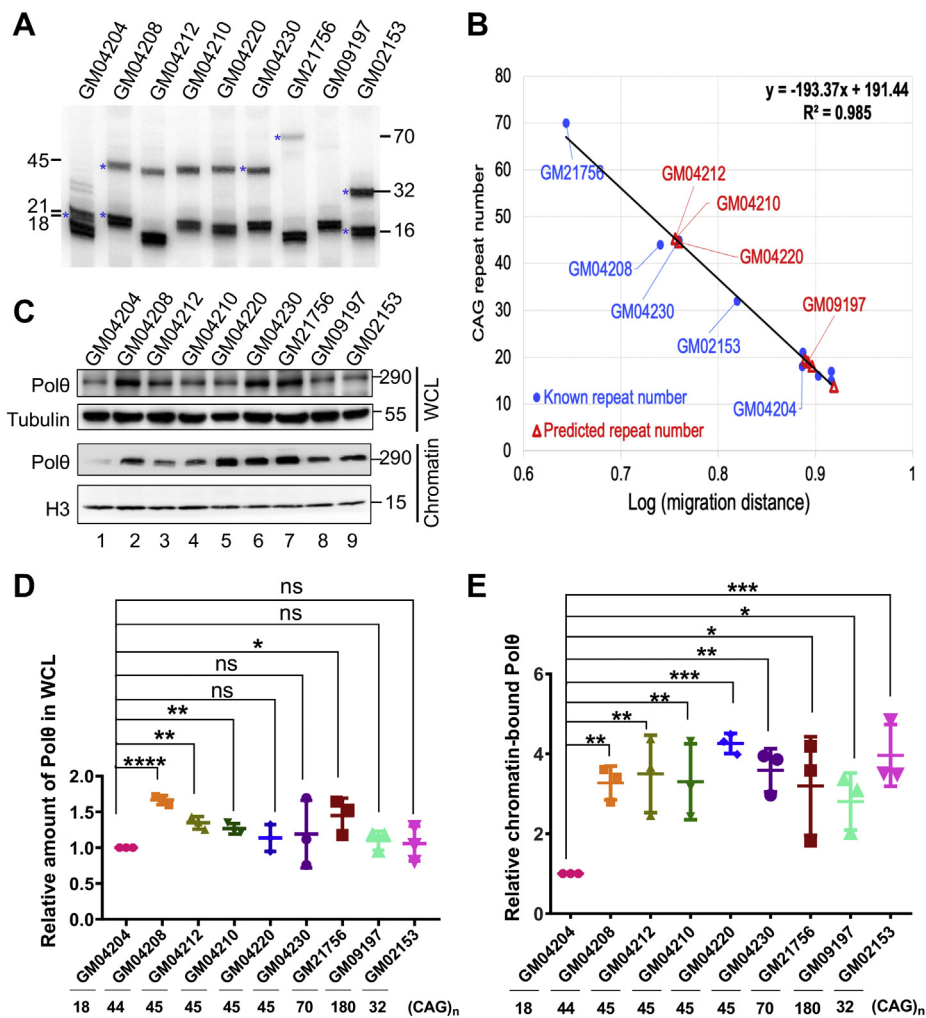


Figure 5. Pol θ 's chromatin levels correlate with *HTT* CAG repeats. *A*, fractionation of the PCR products of *HTT* CAG repeat numbers of HD and non-HD cell lines by denaturing PAGE. PCR products were visualized by Southern blot analysis using a ^{32}P -labeled (CTG)₅ probe. *B*, determination of CAG repeat numbers by using a unique linear function ($y = -193.37x + 191.44$) between the CAG repeat number (y) and the logarithmic value of its migration distance. Cells lines with known or calculated CAG repeat numbers are shown in blue and red, respectively. *C*, determination of Pol θ levels in whole-cell lysates (WCLs) and chromatin by Western blotting. *D* and *E*, quantification of Pol θ levels in WCL (*D*) and chromatin (*E*) in HD and non-HD cell lines. * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$. HD, Huntington's disease; Pol θ , DNA polymerase θ .

lane 5), the former showed much less chromatin-bound Pol θ than the latter ones (Fig. 5E). These results indicate a close association between Pol θ chromatin binding and CAG repeat expansion. Interestingly, although GM02153 only carried an *HTT* with 32 CAG repeats, its chromatin-bound Pol θ level is essentially the same as in several HD cell lines (Fig. 5C). This may explain why some patients display HD symptoms although their CAG repeat number is lower than the threshold 35 (45, 46). Thus, Pol θ 's chromatin level could be a reliable hallmark for HD diagnosis.

Discussion

We show here that Pol θ promotes CAG repeat instability in multiple ways. Pol θ can extend a CAG repeat hairpin, which results in hairpin retention, and, in turn, CAG repeat expansion (Figs. 1 and 2). Pol θ also induces large expansions when it copies a DNA template that contains 35 CAG repeats. This

error-prone property of Pol θ is due to its insertion 2, as Pol $\theta\Delta i2$ cannot induce repeat expansions regardless of the primers or templates used (Fig. 3). Mn^{2+} , which induces flexibility in polymerases' active sites, further stimulates Pol θ 's error-prone activity (Fig. 4). Therefore, we have identified Pol θ as a potent factor in promoting CAG repeat expansions in HD and other neurodegenerative diseases.

Pol θ 's involvement in CAG/CTG repeat expansion appears to be well justified. First, Pol θ participates in BER (29), a DNA repair pathway implicated in CAG repeat instability (16). The brains of patients with HD are characterized by high levels of ROS (47), which can induce 8-oxoG adducts in guanine-rich CAG repeats (15). Repairing 8-oxoG requires the OGG1 glycosylase, which, together with the AP endonuclease, generates a small gap in the damaged DNA strand. This gap is then filled by Pol θ or Pol β because both enzymes are required to repair oxidative damage (29). Because CAG repeats are prone to hairpin formation in the presence of a strand break/gap, a

hairpin formed within the repeats can be utilized as a primer for DNA synthesis by Pol θ (this study) or Pol β (23, 34), which leads to hairpin retention and large repeat expansion. In addition, Pol θ possesses a large insertion 2 in the thumb domain (31), which creates a large flexible cavity in the DNA-binding surface (32, 33). This cavity facilitates misalignments between the primer and template during DNA synthesis (32), which leads to the addition of nucleotides to homopolymeric runs (30, 32). We showed that Pol θ uses its large catalytic domain for CAG/CTG hairpin-priming synthesis (Figs. 1 and 2) and for large repeat expansions without a pre-existing hairpin structure in a repeat length-dependent manner (Fig. 3). Depleting insertion 2 makes Pol θ a relatively high-fidelity polymerase, as Pol $\theta\Delta i2$ no longer synthesizes error-containing DNA (Figs. 3 and 4). Finally, Pol θ is highly chromatin-enriched in HD cells (Fig. 5). This gives Pol θ the advantage of participating in DNA metabolic reactions, including the repair of oxidative DNA damage, so it can stimulate CAG repeat expansions in HD cells. However, how Pol θ is preferentially recruited to chromatin in HD cells remains to be investigated.

Reports about whether people with an *HTT* gene that contains 27 to 35 CAG repeats are at risk for developing HD are controversial. Although CAG repeats in the 27 to 35 range are highly stable (48), some patients had the repeats extend into the HD onset range (>36), and some displayed HD symptoms although their CAG repeats remained

unchanged (45, 46). This suggests that a portion of individuals with 27 to 35 intermediate CAG repeats can develop HD with or without the repeats expanding to ≥ 36 . However, biomarkers that indicate HD onset and mechanisms that extend the repeats beyond the HD threshold are unknown. Our data presented here provide possible explanations for both questions. We showed that cells derived from patients with HD had significantly higher levels of Pol θ in chromatin than cells from non-HD controls (Fig. 5). Surprisingly, we also found that, despite carrying an intermediate allele of 32 CAG repeats and being considered a non-HD control, the GM02153 cell line showed chromatin-bound Pol θ levels similar to those of HD cells, rather than those of non-HD cells (Fig. 5). This suggests that the patient from whom GM02153 was derived might exhibit clinical features of HD. Furthermore, because Pol θ carries out repeat length-dependent expansion, the polymerase can expand the intermediate (*i.e.*, 27–35) repeats to 36 and beyond. Therefore, Pol θ concentration, particularly the chromatin-bound level, could be an important marker to determine the risk of HD for individuals with intermediate *HTT* alleles.

Based on published data and the results presented here, we propose a model to elucidate the mechanism by which Pol θ promotes CAG repeat expansions in nondividing neural cells during OGG1-mediated repair of 8-oxoG (Fig. 6). Once BER generates a strand break or a small gap within the CAG repeats, a CAG hairpin structure forms *via* strand slippage.

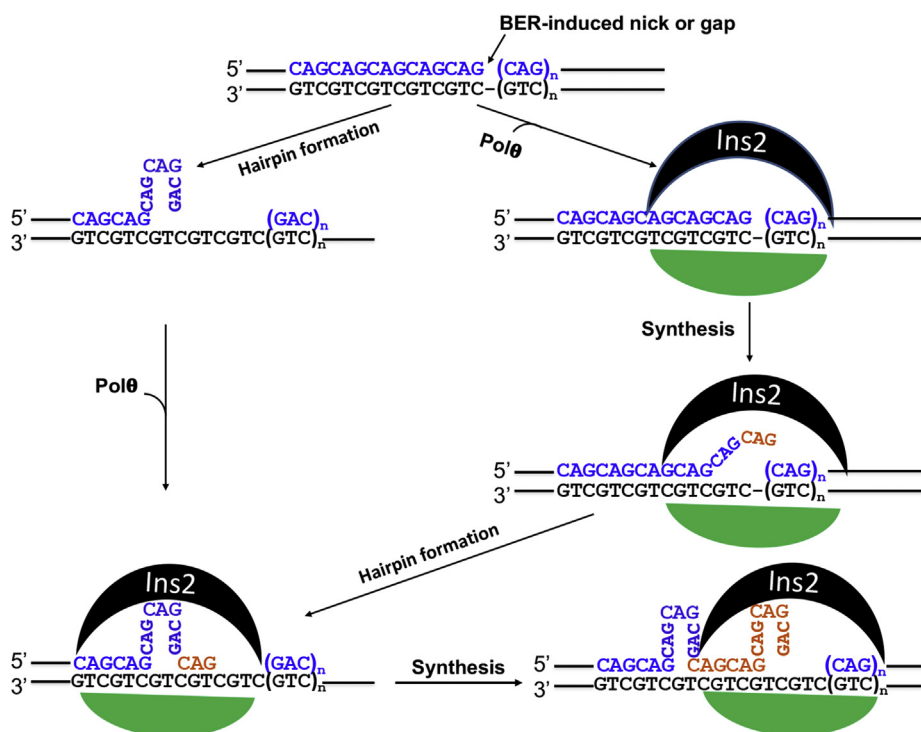


Figure 6. Model of Pol θ -catalyzed CAG repeat expansion. OGG1-mediated repair of oxidative lesions within CAG repeats generates a strand break, which triggers the formation of a hairpin *via* strand slippage. Repair DNA synthesis is initiated when Pol θ is recruited to the hairpin or nick. In the former case, Pol θ utilizes the hairpin as a primer for extension, leading to hairpin retention and repeat expansion. In the latter case, Pol θ incorporates nucleotides to the 3' end of the nick. However, the newly incorporated nucleotides are trapped into the packet of insertion 2 (Ins2) to form a loop-out or a hairpin, which can grow into a big hairpin or shift out of insertion 2. The continuous synthesis by Pol θ can trap a new hairpin formation in insertion 2, resulting in large expansion of the CAG repeats. OGG1, 8-oxo-guanine DNA glycosylase 1; Pol θ , DNA polymerase θ .

Polymerase θ promotes CAG•CTG repeat expansion

Pol θ can use the resulting hairpin as a primer for error-prone synthesis. Alternatively, Pol θ incorporates nucleotides to the 3' end of the nick/gap, and the newly incorporated nucleotides enter into the insertion 2–formed packet to form a loop-out, which can then easily convert into a (CAG)_n hairpin. The hairpin can either grow further or move out of the insertion 2 packet as the polymerase continues to synthesize. In the latter case, additional hairpins can form, which leads to variously expanded CAG repeats. However, the model, which is mainly based on our biochemical studies, requires thorough biological investigations in model organisms. For example, gene knockdown or KO of Pol θ , Pol β , and/or other DNA polymerases may provide definitive evidence of how these polymerases promotes CAG repeat instability.

Experimental procedures

Cells and cell culture

Human fibroblast cell lines GM04204, GM04210, GM04230, GM04212, GM04208, GM04220, GM21756, GM09197, and GM02153 were purchased from the Coriell Institute for Medical Research and cultured in minimum essential medium with 15% fetal bovine serum (FBS). HeLa S3 cells were cultured in RPMI 1640 with 10% FBS. Human embryonic kidney (HEK) 293T cells were cultured in Dulbecco's modified Eagle's medium supplied with 10% FBS, and HEK 293GnTI⁻ cells were grown in suspension in FreeStyle 293 Expression Medium (Gibco) supplemented with 1% FBS. All human cell lines were cultivated at 37 °C in a humidified atmosphere containing 5% CO₂.

Nuclear extract and protein preparation

HeLa nuclear extracts were prepared as previously described (23). DNA sequences coding for human Pol θ polymerase catalytic domain (Fig. 1A) and its insertion 2–deleted derivative Pol $\theta\Delta$ i2 were cloned into pLEXm and designated pLEXm-Pol θ and pLEXm-Pol $\theta\Delta$ i2, respectively. pLEXm-Pol $\theta\Delta$ i2 was derived from pSUMO3-Pol $\theta\Delta$ i2 (a generous gift from Dr Richard T. Pomerantz, Temple University) by inserting the Pol $\theta\Delta$ i2-coding sequence into pLEXm after double digestions with AgeI and XhoI. Pol θ and Pol $\theta\Delta$ i2 were expressed in HEK 293T cells and HEK 293GnTI⁻ cells, respectively, and purified as described previously (32). Pol δ and RFC were expressed in High Five insect cells, and Pol β , replication protein A, and PCNA were expressed in *Escherichia coli*; all proteins were purified as previously described (49, 50).

Hairpin primer extension assay

Unless otherwise mentioned, hairpin primer extension was assayed by Southern blot hybridization as described previously (23). Oligonucleotides that contained 15 CTG or 15 CAG repeats were annealed with ssDNA of M13mp18-(CAG)₁₀ or M13mp18-(CTG)₁₀ (51) to form CTG and a CAG hairpin substrate, respectively (Fig. 1B). Individual hairpin substrates were incubated with 30 μ g HeLa nuclear extracts in the presence or absence of a polymerase (δ , θ , or $\theta\Delta$ i2) for DNA

synthesis at 37 °C for 30 min in a 40- μ l reaction containing 110 mM KCl, 20 mM Tris-HCl, pH 7.6, 5 mM MgCl₂, 1.5 mM ATP, 0.1 mM of various dNTP, and 0.05 mg/ml BSA. In the purified system, each reaction contained RFC (110 fmol) and PCNA (2 pmol) in addition to the indicated polymerase (0.1 μ M). The resulting products were digested with BsrBI before electrophoresis through a 10% denaturing polyacrylamide gel, followed by Southern blotting using a ³²P-labeled probe as described (23). The products were detected by an Amersham Typhoon phosphor imager.

Preparation of WCL and chromatin-binding protein

Individual HD cells were harvested at 70% confluency and washed twice with ice-cold Dulbecco's PBS. Chromatin fractions were prepared as described (52) and incubated with 1:1 (v/v) ice-cold 0.2 M HCl for 10 min on ice to denature and disassociate proteins from chromatin. The acidic solution was neutralized with 1:1 (v/v) of 1 M Tris•HCl (pH 8.0):1 U/ μ l Benzonase Nuclease (Sigma) and incubated on ice for 1 h to allow DNA digestions. The digested genomic DNA and insoluble proteins were removed by centrifugation. WCLs were prepared by resuspending cells in 1:1 (v/v) ice-cold 0.2 M HCl treated as chromatin fractions. Lysates were clarified by high-speed centrifugation (21,000g) for 15 min at 4 °C. Protein concentrations were determined by using the Bio-Rad Bradford Protein Assay reagent. After electrophoresis, chromatin-binding proteins and WCLs were analyzed for Pol θ by Western blotting using a Pol θ -specific antibody (Novus). Pol θ bands were visualized and quantified by using a Bio-Rad ChemiDoc Imaging system.

PCR amplification of CAG repeats and DNA sequencing analysis

PCR amplification of CAG repeats located in exon 1 of *HTT* in HD cell lines used forward (5'-ATGAAGGCCTTC-GAGTCCCTCAAGT CCTTC-3') and reverse (5'-CTGAGG-CAGCAGCGG CTGTGCTGCG-3') primers. PCRs were performed in a volume of 25 μ l containing 100 ng genomic DNA, 1.6 mM of each dNTP, 4 pmol of each primer, and 0.5 U Q5 DNA polymerase (New England Biolabs). After an initial denaturation of 4 min at 98 °C, 40 cycles of 45 s at 98 °C, 1 min at 68 °C, and 3 min at 72 °C were carried out, followed by a final extension of 10 min. PCR products were resolved in a 7.5% (wt/vol) denaturing polyacrylamide gel, followed by Southern blot analysis using a ³²P-labeled (CTG)₅ oligonucleotide probe. PCR products were then visualized by an Amersham Typhoon phosphor imager as described (49).

To analyze the expanded CTG repeats derived from Pol θ -catalyzed primer extension, DNA bands (³²P-labeled) were excised and eluted from gels, then PCR-reamplified for 35 cycles by REDTaq DNA polymerase (Sigma-Aldrich) using primers 5'-ACGTTGTAAAACGACGGCCA-3' (forward) and 5'-CATGATTACGAATTC-3' (reverse), essentially as described above. PCR products were cloned into a pGEM-T vector (Promega) and transfected into *E. coli* DH5-Alpha

(Thermo Fisher). Plasmid DNAs were isolated and subjected to Sanger DNA sequencing (Source BioScience).

Statistical analysis

All statistical assays and one-way ANOVA were performed by using GraphPad Prism 5.0 (GraphPad Software). Data were considered statistically significant if *p*-values were less than 0.05 or 0.001, as indicated.

Data availability

All data are contained within the article.

Supporting information—This article contains [supporting information](#) (53–57).

Acknowledgments—We thank Dr Wei Yang for conceptual discussion and providing the Pol θ expression system; Drs Isabel Mellon, Brett Spear, and Yvonne Fondufe-Mittendorf for helpful comments; Dr Fenghua Yuan for technical advice on purification of Pol δ and RFC; Dr Richard T. Pomerantz for providing pSUMO3-Pol $\theta\Delta$ 2; and Drs Damiana Chiavolini and Jonathan Feinberg for scientific editing. G.-M. L. is a CPRIT Scholar in Cancer Research and the holder of the Reece A. Overcash, Jr Distinguished Chair for Research on Colon Cancer.

Author contributions—K. Y. C., X. L., and G.-M. L. data curation; K. Y. C. software; K. Y. C. validation; K. Y. C., X. L., J. O., and L. G. investigation; K. Y. C., X. L., and J. O. methodology; K. Y. C. and G.-M. L. writing—original draft; K. Y. C., X. L., J. O., and G.-M. L. formal analysis; L. G. and G.-M. L. supervision; L. G. and G.-M. L. funding acquisition; L. G. and G.-M. L. project administration; G.-M. L. conceptualization; G.-M. L. resources; G.-M. L. writing—review and editing.

Funding and additional information—The work was supported in part by a grant from the Cancer Prevention and Research Institute of Texas (CPRIT) (RR160101, to G.-M. L.). K. Y. C. was supported by a training grant (ES007266) from the NIEHS, National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflict of interest—All other authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: 8-oxoG, 8-oxo-guanine; AP, abasic; BER, base excision repair; FBS, fetal bovine serum; HD, Huntington's disease; HEK, human embryonic kidney; OGG1, 8-oxo-guanine DNA glycosylase 1; PCNA, proliferating cellular nuclear antigen; Pol β , polymerase β ; Pol θ , DNA polymerase θ ; Pol $\theta\Delta$ 2, insertion 2–depleted Pol θ ; RFC, replication factor C; ROS, reactive oxygen species; TdT, terminal deoxynucleotidyl transferase; TNRs, trinucleotide repeats; WCLs, whole-cell lysates.

References

- 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
- Mirkin, S. M. (2007) Expandable DNA repeats and human disease. *Nature* **447**, 932–940
- Khrstich, A. N., and Mirkin, S. M. (2020) On the wrong DNA track: Molecular mechanisms of repeat-mediated genome instability. *J. Biol. Chem.* **295**, 4134–4170
- Schmidt, M. H., and Pearson, C. E. (2016) Disease-associated repeat instability and mismatch repair. *DNA Repair* **38**, 117–126
- Liu, G., Chen, X., Bissler, J. J., Sinden, R. R., and Leffak, M. (2010) Replication-dependent instability at (CTG) \times (CAG) repeat hairpins in human cells. *Nat. Chem. Biol.* **6**, 652–659
- 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
- McMurray, C. T. (2010) Mechanisms of trinucleotide repeat instability during human development. *Nat. Rev. Genet.* **11**, 786–799
- Usdin, K., and Grabczyk, E. (2000) DNA repeat expansions and human disease. *Cell Mol. Life Sci.* **57**, 914–931
- 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
- Guo, J., Chen, L., and Li, G. M. (2017) DNA mismatch repair in trinucleotide repeat instability. *Sci. China Life Sci.* **60**, 1087–1092
- 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
- Gacy, A. M., and McMurray, C. T. (1998) Influence of hairpins on template reannealing at trinucleotide repeat duplexes: A model for slipped DNA. *Biochemistry* **37**, 9426–9434
- 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
- Panigrahi, G. B., Lau, R., Montgomery, S. E., Leonard, M. R., and Pearson, C. E. (2005) Slipped (CTG) \times (CAG) repeats can be correctly repaired, escape repair or undergo error-prone repair. *Nat. Struct. Mol. Biol.* **12**, 654–662
- Madabhushi, R., Pan, L., and Tsai, L. H. (2014) DNA damage and its links to neurodegeneration. *Neuron* **83**, 266–282
- 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
- Bruner, S. D., Norman, D. P., and Verdine, G. L. (2000) Structural basis for recognition and repair of the endogenous mutagen 8-oxoguanine in DNA. *Nature* **403**, 859–866
- Li, G. M. (2010) Novel molecular insights into the mechanism of GO removal by MutM. *Cell Res.* **20**, 116–118
- Friedberg, E. C., Walker, G. C., Siede, W., Wood, R. D., Schultz, R. A., and Ellenberger, T. (2006) *DNA Repair and Mutagenesis*, ASM Press, Washington, DC
- Krishnamurthy, N., Haraguchi, K., Greenberg, M. M., and David, S. S. (2008) Efficient removal of formamidopyrimidines by 8-oxoguanine glycosylases. *Biochemistry* **47**, 1043–1050
- Prasad, R., Shock, D. D., Beard, W. A., and Wilson, S. H. (2010) Substrate channeling in mammalian base excision repair pathways: Passing the baton. *J. Biol. Chem.* **285**, 40479–40488
- 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
- Chan, N. L., Guo, J., Zhang, T., Mao, G., Hou, C., Yuan, F., Huang, J., Zhang, Y., Wu, J., Gu, L., and Li, G. M. (2013) Coordinated processing of 3' slipped (CAG) n /(CTG) n hairpins by DNA polymerases beta and delta preferentially induces repeat expansions. *J. Biol. Chem.* **288**, 15015–15022
- Bogdanov, M. B., Andreassen, O. A., Dedeoglu, A., Ferrante, R. J., and Beal, M. F. (2001) Increased oxidative damage to DNA in a transgenic mouse model of Huntington's disease. *J. Neurochem.* **79**, 1246–1249
- Kennedy, L., and Shelbourne, P. F. (2000) Dramatic mutation instability in HD mouse striatum: Does polyglutamine load contribute to cell-specific vulnerability in Huntington's disease? *Hum. Mol. Genet.* **9**, 2539–2544

Polymerase θ promotes CAG•CTG repeat expansion

26. Crespan, E., Hubscher, U., and Maga, G. (2015) Expansion of CAG triplet repeats by human DNA polymerases lambda and beta *in vitro*, is regulated by flap endonuclease 1 and DNA ligase 1. *DNA Repair* **29**, 101–111
27. Liu, Y., and Wilson, S. H. (2012) DNA base excision repair: A mechanism of trinucleotide repeat expansion. *Trends Biochem. Sci.* **37**, 162–172
28. Wood, R. D., and Doublet, S. (2016) DNA polymerase theta (POLQ), double-strand break repair, and cancer. *DNA Repair* **44**, 22–32
29. Yoshimura, M., Kohzaki, M., Nakamura, J., Asagoshi, K., Sonoda, E., Hou, E., Prasad, R., Wilson, S. H., Tano, K., Yasui, A., Lan, L., Seki, M., Wood, R. D., Arakawa, H., Buerstedde, J. M., *et al.* (2006) Vertebrate POLQ and POLbeta cooperate in base excision repair of oxidative DNA damage. *Mol. Cell* **24**, 115–125
30. Arana, M. E., Seki, M., Wood, R. D., Rogozin, I. B., and Kunkel, T. A. (2008) Low-fidelity DNA synthesis by human DNA polymerase theta. *Nucleic Acids Res.* **36**, 3847–3856
31. Hogg, M., Seki, M., Wood, R. D., Doublet, S., and Wallace, S. S. (2011) Lesion bypass activity of DNA polymerase theta (POLQ) is an intrinsic property of the pol domain and depends on unique sequence inserts. *J. Mol. Biol.* **405**, 642–652
32. Lee, Y. S., Gao, Y., and Yang, W. (2015) How a homolog of high-fidelity replicases conducts mutagenic DNA synthesis. *Nat. Struct. Mol. Biol.* **22**, 298–303
33. Zahn, K. E., Averill, A. M., Aller, P., Wood, R. D., and Doublet, S. (2015) Human DNA polymerase theta grasps the primer terminus to mediate DNA repair. *Nat. Struct. Mol. Biol.* **22**, 304–311
34. Guo, J., Gu, L., Leffak, M., and Li, G. M. (2016) MutSbeta promotes trinucleotide repeat expansion by recruiting DNA polymerase beta to nascent (CAG)_n or (CTG)_n hairpins for error-prone DNA synthesis. *Cell Res.* **26**, 775–786
35. Kent, T., Chandramouly, G., McDevitt, S. M., Ozdemir, A. Y., and Pomerantz, R. T. (2015) Mechanism of microhomology-mediated end-joining promoted by human DNA polymerase theta. *Nat. Struct. Mol. Biol.* **22**, 230–237
36. 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
37. Kent, T., Mateos-Gomez, P. A., Sfeir, A., and Pomerantz, R. T. (2016) Polymerase theta is a robust terminal transferase that oscillates between three different mechanisms during end-joining. *Elife* **5**, e13740
38. 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)_n(CTG)_n repeats: Slipped-out repeats and slip-out junctions. *Nucleic Acids Res.* **30**, 4534–4547
39. 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
40. Poranen, M. M., Salgado, P. S., Koivunen, M. R., Wright, S., Bamford, D. H., Stuart, D. I., and Grimes, J. M. (2008) Structural explanation for the role of Mn²⁺ in the activity of phi6 RNA-dependent RNA polymerase. *Nucleic Acids Res.* **36**, 6633–6644
41. Vashishtha, A. K., Wang, J., and Konigsberg, W. H. (2016) Different divalent cations alter the kinetics and fidelity of DNA polymerases. *J. Biol. Chem.* **291**, 20869–20875
42. Frank, E. G., and Woodgate, R. (2007) Increased catalytic activity and altered fidelity of human DNA polymerase iota in the presence of manganese. *J. Biol. Chem.* **282**, 24689–24696
43. Wu, W.-J., Yang, W., and Tsai, M.-D. (2017) How DNA polymerases catalyze replication and repair with contrasting fidelity. *Nat. Rev. Chem.* **1**, 0068
44. Gusella, J. F., Wexler, N. S., Conneally, P. M., Naylor, S. L., Anderson, M. A., Tanzi, R. E., Watkins, P. C., Ottina, K., Wallace, M. R., Sakaguchi, A. Y., Young, A. B., Shoulson, I., Bonilla, E., and Martin, J. B. (1983) A polymorphic DNA marker genetically linked to Huntington's disease. *Nature* **306**, 234–238
45. Ha, A. D., and Jankovic, J. (2011) Exploring the correlates of intermediate CAG repeats in Huntington disease. *Postgrad. Med.* **123**, 116–121
46. Semaka, A., Collins, J. A., and Hayden, M. R. (2010) Unstable familial transmissions of Huntington disease alleles with 27–35 CAG repeats (intermediate alleles). *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **153B**, 314–320
47. Kumar, A., and Ratan, R. R. (2016) Oxidative stress and Huntington's disease: The good, the bad, and the ugly. *J. Huntingtons Dis.* **5**, 217–237
48. Brocklebank, D., Gayan, J., Andresen, J. M., Roberts, S. A., Young, A. B., Snodgrass, S. R., Penney, J. B., Ramos-Arroyo, M. A., Cha, J. J., Rosas, H. D., Hersch, S. M., Feigin, A., Cherny, S. S., Wexler, N. S., Housman, D. E., *et al.* (2009) Repeat instability in the 27–39 CAG range of the HD gene in the Venezuelan kindreds: Counseling implications. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **150B**, 425–429
49. Ortega, J., Lee, G. S., Gu, L., Yang, W., and Li, G. M. (2021) Mismatch-bound human MutS-MutL complex triggers DNA incisions and activates mismatch repair. *Cell Res.* **31**, 542–553
50. Zhang, Y., Yuan, F., Presnell, S. R., Tian, K., Gao, Y., Tomkinson, A. E., Gu, L., and Li, G. M. (2005) Reconstitution of 5'-directed human mismatch repair in a purified system. *Cell* **122**, 693–705
51. Zhang, T., Huang, J., Gu, L., and Li, G. M. (2012) *In vitro* repair of DNA hairpins containing various numbers of CAG/CTG trinucleotide repeats. *DNA Repair* **11**, 201–209
52. Fang, J., Huang, Y., Mao, G., Yang, S., Rennert, G., Gu, L., Li, H., and Li, G. M. (2018) Cancer-driving H3G34V/R/D mutations block H3K36 methylation and H3K36me3-MutSalpha interaction. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 9598–9603
53. Cohen-Carmon, D., Sorek, M., Lerner, V., Divya, M. S., Nissim-Rafinia, M., Yarom, Y., and Meshorer, E. (2020) Progerin-induced transcriptional changes in Huntington's disease human pluripotent stem cell-derived neurons. *Mol. Neurobiol.* **57**, 1768–1777
54. de Mezer, M., Wojciechowska, M., Napierala, M., Sobczak, K., and Krzyzosiak, W. J. (2011) Mutant CAG repeats of Huntington transcript fold into hairpins, form nuclear foci and are targets for RNA interference. *Nucleic Acids Res.* **39**, 3852–3863
55. Victor, M. B., Richner, M., Olsen, H. E., Lee, S. W., Monteys, A. M., Ma, C., Huh, C. J., Zhang, B., Davidson, B. L., Yang, X. W., and Yoo, A. S. (2018) Striatal neurons directly converted from Huntington's disease patient fibroblasts recapitulate age-associated disease phenotypes. *Nat. Neurosci.* **21**, 341–352
56. Ferlazzo, M. L., Sonzogni, L., Granzotto, A., Bodgi, L., Lartin, O., Devic, C., Vogin, G., Pereira, S., and Foray, N. (2014) Mutations of the Huntington's disease protein impact on the ATM-dependent signaling and repair pathways of the radiation-induced DNA double-strand breaks: corrective effect of statins and bisphosphonates. *Mol. Neurobiol.* **49**, 1200–1211
57. Mollica, P. A. (2015) *DNA Repair Deficiency in Huntington's Disease Fibroblasts and Induced Pluripotent Stem Cells*. Doctor of Philosophy (PhD), Dissertation, Biological Sciences, Old Dominion University