HHS Public Access

Author manuscript

Nat Struct Mol Biol. Author manuscript; available in PMC 2016 September 28.

Published in final edited form as:

Nat Struct Mol Biol. 2009 August ; 16(8): 869–875. doi:10.1038/nsmb.1638.

Incision-Dependent and Error-Free Repair of $(CAG)_n/(CTG)_n$ Hairpins in Human Cell Extracts

Caixia Hou[1], **Nelson L.S. Chan**[1], **Liya Gu**[1], and **Guo-Min Li**[1,2]

[1]Graduate Center for Toxicology and Markey Cancer Center, University of Kentucky College of Medicine, Lexington, Kentucky, USA

[2]Department of Molecular and Cellular Biochemistry, University of Kentucky College of Medicine, Lexington, Kentucky, USA

Abstract

Expansion of CAG/CTG trinucleotide repeats is associated with certain familial neurological disorders, including Huntington's disease. Increasing evidence suggests that formation of a stable DNA hairpin within CAG/CTG repeats during DNA metabolism contributes to their expansion. However, the molecular mechanism(s) by which cells remove CAG/CTG hairpins remain unknown. Here, we demonstrate that human cell extracts can catalyze error-free repair of CAG/CTG hairpins in a nick-directed manner. The repair system specifically targets CAG/CTG tracts for incisions in the nicked DNA strand, followed by DNA resynthesis using the continuous strand as a template, thereby ensuring CAG/CTG stability. PCNA is required for the incision step of the hairpin removal, which utilizes distinct endonuclease activities for individual CAG/CTG hairpins depending on their strand locations and/or secondary structures. The implication of these data for understanding the etiology of neurological diseases and trinucleotide repeat instability is discussed.

Keywords

Trinucleotide repeat; DNA hairpin repair; incision; PCNA; Huntington's disease

Introduction

Expansion of trinucleotide repeats (TNRs) is tightly associated with at least 15 human familial neurological, neurodegenerative and neuromuscular disorders, including CAG repeat expansion-caused Huntington's disease (HD) and CTG repeat expansion-caused myotonic dystrophy [1,2]. Each of these diseases is clinically distinct, and involves expansion

Correspondence should be addressed to Guo-Min Li at the Graduate Center for Toxicology, University of Kentucky College of Medicine, Lexington, KY 40536. gmli@uky.edu; phone: 859-257-7053; fax: 859-323-1059.

Author Contributions

C.H. performed all the experiments; N.L.C. helped to determine some incision intermediates; L.G. designed and constructed DNA hairpin substrates, and developed the in vitro HPR assay; G.-M.L. designed and analyzed all the experiments, and wrote the manuscript.

of a TNR at a unique site in the human genome. TNR expansion appears to share a similar biochemical/genetic mechanism, in which the expansion alters the function or expression of the gene in which it lies. For each disease, pathological symptoms are triggered and become progressively more severe when the number of repeats reaches and then exceeds a critical threshold. At the HD locus, CAG repeat lengths from 11 to 34 are not associated with disease pathology, while repeat lengths ≥35 result in clinical symptoms of HD. A particular mystery in the field is that TNR instability is associated with post-mitotic non-dividing cells (e.g., neurons) in an age-dependent and tract-length heterogeneous manner. At present, the mechanism by which TNR instability occurs has remained conjectural.

Since DNA expansions require DNA synthesis, TNR expansions must be associated with DNA replication and/or repair [1,2]. One model suggests that TNR sequences, e.g., CAG and CTG repeats, tend to form hairpins via strand slippage in the newly synthesized or nicked strand during DNA replication or repair, leading to TNR expansion [1-6]. This model is consistent with the observations that CAG and CTG repeats form very stable hairpins with a melting temperature higher than physiological temperature in mammalian cells [7-9]. Therefore, TNR hairpins are expected to persist *in vivo* once they form, and to require an active mechanism for dissolution or removal, if TNR expansion is to be prevented. Indeed, human cells possess a hairpin repair (HPR) system to remove CAG and CTG hairpins in a nick-dependent manner [10]. However, how the novel activity removes CAG/CTG hairpins is unknown.

To determine the molecular basis of this HPR system, we performed CAG/CTG HPR in human nuclear extracts using an *in vitro* assay that directly monitors repair intermediates and products. We show here that human cells conduct error-free repair of CAG/CTG hairpins in a strand-specific and PCNA-dependent manner. The repair is initiated by endonucleolytic incisions that specifically target the repeat tracts in the nicked strand, followed by DNA resynthesis using the continuous strand as a template. Our results support a notion that the HPR pathway is a primary system that ensures TNR stability in human cells and that defects in this system could lead to TNR instability and human diseases.

Results

CAG or CTG HPR in Human Cells is Error-Free

We developed an *in* vitro assay to study TNR HPR using nicked circular heteroduplex DNA substrates derived from the M13mp18 phage series (**Fig. 1a**). These heteroduplexes contain a $(CAG)_{25}$ or $(CTG)_{25}$ hairpin in the viral (V) or complementary (C) strand and a nick 5' to the hairpin in the C strand. Since repair of loop-containing heteroduplexes, including CAG or CTG hairpins, is targeted to the nicked DNA strand [10-12], $(CAG)_{25}$- or $(CTG)_{25}$-HPR in this study was scored by monitoring repeat length changes in the nicked strand using a strand-specific ^{32}P-labeled oligonucleotide as described [13,14].

HeLa nuclear extracts were incubated with the individual DNA substrates shown in **Fig. 1a**, and the repair results are shown in **Fig. 1b**. All four DNA substrates were efficiently repaired. For substrates with a hairpin on the V strand, i.e., V-$(CAG)_{25}$ and V-$(CTG)_{25}$, a slower migrating species was detected in reactions containing active HeLa extract (lanes 2

and 5, red squares), but not in reactions containing heat-inactivated (x) HeLa extract (lanes 1 and 4). This species is 75-nt longer than the original substrate, indicating that the continuous strand was used as a template for repair DNA synthesis. For substrates having an extruded CAG or CTG hairpin on the C strand, i.e., C-$(CAG)_{25}$ and C-$(CTG)_{25}$, the repair product is 75-nt shorter than the original substrate (lanes 8 and 11, red squares). However, a ^{32}P-labeled oligonucleotide specifically annealing to the V strand near the BglI site failed to detect any repeat number changes regardless of the presence of a hairpin in the non-nicked strand (data not shown); and repair assays performed with a covalently-closed circular substrate also showed little repair (data not shown). These results indicate that CAG/CTG HPR is indeed nick-directed [10]. No repair products were detected in the presence of aphidicolin, a potent inhibitor of DNA polymerases α, δ, and ε (**Fig. 1b**). These observations are consistent with a repair mechanism involving excision/incision in the nicked strand followed by DNA re-synthesis using the continuous strand as a template. Therefore, the human CAG/CTG HPR mechanism is an error-free repair system.

A previous study reported that human cell extracts carry out error-prone repair when a CAG or CTG hairpin is in the nicked strand [10]. The error-prone repair products, referred to as slipped intermediate heteroduplex DNAs (SI-DNAs), included CAG or CTG repeats of variable length. In contrast, the SI-DNAs were not detected in this study (**Fig. 1b**, lanes 7-12). A possible explanation for this discrepancy is given in Discussion.

It is interesting to note that there is essentially no difference in total repair when a CTG- or a CAG-hairpin is present in the nicked strand (**Fig. 1b**, lanes 8 and 11); however, a CTG-hairpin is repaired much less efficiently than a CAG-hairpin when it is located in the V strand (compare lane 5 with lane 2). This may be because a CTG hairpin is more stable [15] and more resistant to unwinding than a CAG hairpin when acting as a template for repair. This result supports involvement of a DNA helicase in the processing a CTG/CAG hairpin in non-nicked strand [16].

Dual Incisions Remove CTG Hairpin in Nicked Strand

We examined the mechanism of TNR HPR in HeLa extracts by characterizing repair intermediates generated under reaction conditions that support DNA excision/incision, but not DNA synthesis, e.g., in the absence of dNTPs or in the presence of ddNTPs and aphidicolin. After incubation with HeLa nuclear extract, reaction products were cleaved with BsmBI and Bsu36I, and analyzed by Southern blot using a probe that anneals to the C strand near the BsmBI site (red bar, designated BsmBI probe hereafter) or near the BglI site (blue bar, designated BglI probe hereafter). Results of this analysis with C-$(CTG)_{25}$ are shown in **Fig. 2**. When the BsmBI probe was used for Southern analysis, unrepaired DNA molecules, i.e., unreacted and nick-ligated substrates, appear as a heavy doublet at the top of the blot (bracket, **Fig. 2a**, lanes 1-5). Two putative reaction intermediates, corresponding to bands I and II, were detected, and they are smaller than the BsmBI-HindIII fragment, but larger than the BsmBI-EcoRI fragment (**Fig. 2a**, lanes 3-5). Since $(CTG)_{35}$ repeats were cloned within HindIII and EcoRI sites (see **Fig. 1a**) and since bands I and II are discrete products, our results suggest that the intermediates are likely produced by incisions targeting the repeats (see **Fig. 2c**, diagrams 1 and 2), rather than by excision as previously proposed[10]. In fact, we

could not detect intermediates shorter than the intact substrate but longer than band I - which would correspond to products of excision starting at the nick - during the course of the reaction.

The same blot was also probed with the BglI probe (**Fig. 2b**, blue bar), which anneals to the C-strand near the BglI site, *i.e.*, the nick site. If the reaction proceeded via nick-directed excision, the target of this probe would have been degraded. Besides identifying unrepaired/unreacted DNA molecules (bracketed), the BglI probe also detected four putative reaction intermediates (**Fig. 2b**, bands a-d), indicating that incisions but not excisions must have occurred. Proposed structures for these four incision intermediates are shown in **Fig. 2c** (diagrams 3-6). We believe that intermediates c and d are derived from intermediates a and b, respectively, by ligation to the 78-nt Bsu36I-BglI fragment. Consistent with this prediction, intermediates a and b were the only bands detected when reaction products were digested with BsmBI and BglI (**Fig. 2d**). Furthermore, products c and d migrated between fragments Bsu36I-HindIII and Bsu36I-EcoRI (**Supplementary Fig. 1** on line), and were detected by a probe for the Bsu36I-BglI fragment (data not shown). Because products c and d are more abundant than products a and b (**Fig. 2b**), and because the nick is required for CAG/CTG HPR (data not shown, and [10]), it appears that the nick is rapidly ligated after incisions occur.

Based on the above results, we propose that HeLa cell extracts introduce dual incisions on either side of the CTG hairpin (**Fig. 2c**), resulting in release of the hairpin from the DNA substrate. We tested this possibility by performing C-$(CTG)_{25}$ hairpin repair in a time course, and the repair products/intermediates were analyzed with a ^{32}P-labeled $(CAG)_{10}$ probe (purple line in **Fig. 2e**). As expected, this probe identified CTG repeat-containing final repair product (red rectangle) and intermediate b, but not intermediate a, which contains little CTG repeats (**Fig. 2e**). In addition, the $(CAG)_{10}$ probe also detected a product (purple oval, **Fig. 2e**) that migrates faster than the $(CTG)_{35}$-containing BsrBI-HindIII fragment. The size and the CTG repeat-containing nature of the product suggest that it is likely the dual incision-released CTG hairpin. This notion is also supported by the fact that the product was undetectable by the BglI probe (**Fig. 2f**). Therefore, HeLa cell extracts remove CTG hairpins in the nicked strand by incisions on either side of the hairpin.

We estimated the lengths of the incision intermediates, using the migration distances of the molecular markers in **Fig. 2** (**e** and **f**) as standards, as described [17]. The estimated nucleotide lengths for products a and b are 168 nt and 230 nt, respectively (see **Supplementary Fig. 2**), suggesting that the 5' incision (generating products I and a) is ~ 20 nt 3' to the HindIII cleavage site (see black-boxed G in **Fig. 1a**), while the 3' incision (generating products II and b) is ~ 29 nt 5' to the EcoRI cleavage site (see black box near the 3' base of the CTG hairpin in **Fig. 1a**). These data provide further support for dual incisions of C-$(CTG)_{25}$ hairpin.

It is also noted that the incision to generate product a or I occurred initially, as early as 1 min, followed by the incision to generate product b or II (**Fig. 2a,b**). These incision intermediates were gradually converted to the final repair products (**Fig. 2e,f**), and eventually disappeared after 45-min incubation (data not shown).

Formation of a 5’ Flap During Nicked Strand CAG HPR

The repair intermediates of substrate C-$(CAG)_{25}$ were also examined using a similar approach. When the BsmBI probe was used in the Southern hybridization (**Fig. 3a**), three putative reaction intermediates (bands I, II, and III) were detected, as well as a smeared region between the unreacted DNA substrate (see small bracket) and band I. Band I is slightly smaller than the HindIII-BsmBI fragment; band II appears to be in the middle of the HindIII-EcoRI fragment, *i.e.*, the loop of the hairpin; band III has a similar size as the EcoRI-BsmBI fragment (**Fig. 3a**). Minor products smaller than band III were also seen (larger bracket). These DNA molecules could be produced by excision, multiple incisions, or both. When the same blot was hybridized with the BglI probe (**Fig. 3b**), two slightly smeared bands were observed, corresponding to bands a and b. The size of band a suggests incision between the HindIII and EcoRI sites (see markers in lane 1), and band b could be generated from band a by ligation to the 78 nt Bsu36I-BglI fragment (**Fig. 3c**, diagrams 4 and 5). Indeed, band b migrates between the HindIII-Bsu36I and EcoRI-Bsu36I fragments (**Fig. 3b**, lane 7). Because product II (**Fig. 3a**) and product a or b (**Fig. 3b**) correspond to cleavage sites between the HindIII and EcoRI sites, they are likely to result from incision at or near the loop of the CAG hairpin, with product a or b representing the 5’ portion and product II representing the 3’ portion of the same molecule (**Fig. 3c**, diagrams 2, 4, and 5).

In **Fig. 3a**, band I is abundant and is likely generated by cleavage at a site closer to the nick (the BglI site) than bands II and III. However, a product corresponding to band I was not detected when the same blot was probed with the BglI probe (**Fig. 3b**). Furthermore, there is a smear on the blot between the unrepaired DNA substrate and band I (**Fig. 3a**), corresponding to multiple incision or excision events between the nick (BglI) site and the 5'-end of the CAG hairpin (**Fig. 3a**, also dashed line in **c**). These observations suggest that band I may be generated by excision that is terminated at or near the 5' base of the CAG hairpin, leaving the hairpin unremoved. However, whether the excision is terminated by the physical structure of the CAG hairpin or a protein-DNA complex at the 5’ base of the hairpin is unknown. Band II may be produced by cleavage at the loop of the CAG hairpin, such that band II would include approximately half of the CAG repeat units. We propose that bands I and II would require further processing to ensure error-free HPR, and that the remaining CAG repeats in these molecules might form a 5’ flap.

Band III in **Fig. 3a** appeared after incubation for 15 to 30 min, and its size corresponds to incision near the EcoRI site. We therefore hypothesize that band III could be derived from bands I and II via removing the CAG repeat-containing flap structure by a 5’ flap endonuclease. We tested this idea by incubating repair intermediates generated in HeLa extracts with purified recombinant FEN-1. As shown in **Fig. 3d**, incubation of C-$(CAG)_{25}$ intermediates with FEN-1 produces a DNA molecule similar in size to the EcoR1-BsmBI fragment (left panel, lane 5). Interestingly, when the same blot was probed with a ^{32}P-labeled $(CTG)_{10}$ probe, a series of small fragments were identified in the reaction containing FEN-1 (right panel, lane 5) but not in reactions lacking FEN-1 (right panel, lane 4). These observations strongly suggest that a CAG flap forms during repair of the C-$(CAG)_{25}$ hairpin in HeLa nuclear extracts, which can be further processed by a flap endonuclease.

It is worth mentioning that although product II and band a/b were derived from the same incision (**Fig. 3c**, diagrams 2, 4, and 5), the signals for band a/b (**Fig. 3b**) are much weaker than for band II (**Fig. 3a**). We believe that the major reason is that some of the CAG substrate undergo 5'→3' excision from the strand break toward to the HindIII site (see dashed lines in **Fig. 3c**), which generates band I (**Fig. 3a,c**). As described above, the incision product for band I was not identifiable by the BglI probe (**Fig. 3b**), as the probe target was destroyed by the excision activity. Likewise, the excision activity may have removed the BglI probe target sequence of band a before and/or after the incision to generate band a is made (**Fig. 3c**, diagram 4), leading to its reduced detection. This is supported by the fact that band a consists of a series of smeared species (**Fig. 3b**). Because of the excision, less band a can be converted to band b by ligations, resulting in reduced formation of band b.

Repair of Continuous Strand Hairpins via Incisions

The intermediates that form during repair of $(CAG)_{25}$ and $(CTG)_{25}$ hairpins in the continuous V strand were also examined (**Fig. 4**). When the Southern blots were analyzed with the BsmBI probe (**Fig. 4a**), two major putative reaction intermediates was detected within the HindIII-EcoRI region for both substrates, indicative of incision or incision/excision opposite the hairpin. When the same membrane was probed for incisions with the BglI probe, both substrates showed putative intermediates between the HindIII and EcoRI markers (**Fig. 4b**), suggesting targeted incision(s) within repeat sequences opposite the hairpin, with or without ligation of the nick at the BglI site. However, the abundance of reaction intermediates appeared to be higher for the $(CAG)_{25}$ DNA substrate, consistent with the fact that V-$(CAG)_{25}$ is repaired more efficiently than V-$(CTG)_{25}$ (see **Fig. 1b**).

We noted some subtle differences in reaction intermediates detected by the BsmBI probe (**Fig. 4a**) and the BglI probe (**Fig. 4b**). For substrate V-$(CAG)_{25}$, both probes identified two major repair intermediates, indicating that two incisions occur during repair of V-$(CAG)_{25}$ heteroduplex. However, for substrate V-$(CTG)_{25}$, a putative intermediate near the HindIII site (asterisk, lanes 8-10 in **Fig. 4a**), which represents ~ 15% of the partially repaired DNA molecules, was identified by the BsmBI probe, but the corresponding intermediate was not detected in **Fig. 4b** (lanes 8-10). This discrepancy suggests that this intermediate is generated by excision, because excision removes the target sequence for the BglI probe used in **Fig. 4b**. Furthermore, minor faster-migrating bands (see brackets) detected in **Fig. 4a** may be produced by non-specific DNA degradation, which occurs when reactions do not support DNA repair synthesis. Taken together, the results shown here suggest that repair of V-$(CAG)_{25}$ and V-$(CTG)_{25}$ occurs predominantly via incisions in the nicked strand opposite the hairpin.

Requirement for PCNA in CAG/CTG HPR

Proliferating cellular nuclear antigen (PCNA) is required at the initiation stage of nick-directed mismatch repair [18]. Therefore, we tested whether PCNA might play a role in CAG/CTG HPR *in vitro.* Repair assays were performed in the presence of a purified p21 C-terminal peptide (p21C), which strongly inhibits PCNA functions in DNA replication and repair [18-20]. As shown in **Fig. 5a**, p21C almost completely inhibited repair of $(CTG)_{25}$ and $(CAG)_{25}$ hairpins by HeLa nuclear extracts (lanes 3, 7, 11, and 15); and the inhibition was

reversed by adding purified PCNA to the *in vitro* repair assay (lanes 4, 8, 12, and 16). Furthermore, in the presence of aphidicolin, p21C strongly inhibited incisions between the HindIII and EcoRI sites (**Fig. 5b**, lanes 3 and 7 and **Fig. 5c**, lanes 3 and 7), and this inhibition was reversed by addition of purified PCNA to the reaction (**Fig. 5b**, lanes 4 and 8 and **Fig. 5c**, lanes 4 and 8). These observations strongly suggest that PCNA is required at or prior to the incision step for repair of all CAG/CTG hairpins in human cell extracts.

Discussion

We demonstrate here that human cell extracts catalyze error-free nick-directed removal of CAG/CTG hairpins in a manner dependent on PCNA and endonucleolytic incisions. The repair mechanism by which the CAG/CTG hairpin is incised is greatly influenced by the strand location, sequence specificity and secondary structure of the hairpin. These observations suggest that TNR HPR is a complex process, which may involve subtly different enzymes in different biological contexts. Thus, it may be inappropriate to propose a single general model to explain expansion of different TNR sequences associated with human neurological diseases.

Mechanisms of CAG/CTG Hairpin Removal

A previous study proposed that CAG/CTG hairpin is removed by exonuclease activities [10]. However, evidence presented in our study strongly suggests that incisions, rather than excisions, are primarily responsible for CAG/CTG hairpin removal. A model for 5’ nick-directed CAG/CTG HPR via incisions in human cells is presented in **Fig. 6**.

For heteroduplexes with a CTG hairpin in the nicked strand (**Fig. 6a**), an incision at each side of the hairpin releases the heterology (**Fig. 2**). The resulting small gap is filled by an aphidicolin-sensitive polymerase, followed by strand ligation. For heteroduplexes containing a CTG hairpin in the continuous strand (**Fig. 6b**), a single incision occurs opposite the hairpin in the nicked strand (**Fig. 4a,b**, right panels). The incision-generated strand break may facilitate helicase-mediated unwinding of the CTG hairpin. A minor product (~15%), likely derived from excision (see asterisk in **Fig. 4a**), could also undergo unwinding to removal the CTG hairpin. Previous studies in yeast suggest that Srs2 helicase may promote CAG/CTG stability, presumably by unwinding CAG/CTG hairpins [16]. A similar mechanism is also used for repair of CAG hairpins in the continuous strand (**Fig. 6c**), although two incisions appear to occur opposite the hairpin (**Fig. 4a,b**, left panels). For a CAG hairpin in the nicked strand (**Fig. 6d**), an incision in the hairpin loop and an excision from the nick to the 5’ base of the hairpin appear to occur simultaneously (**Fig. 3**). The coordination of these activities is essential for error-free repair of the CAG hairpin. In this case, a 5’ flap containing CAG repeats may form, and can be removed by a flap endonuclease such as FEN-1 (**Fig. 3c-e**) before gap-filling and ligation. PCNA is required at or prior to incisions for all CAG/CTG hairpin substrates tested in this study (**Fig. 5**).

Consistent with our finding here that TNR HPR involves incisions, FEN-1 and the nuclear excision repair (NER) pathway, where DNA lesions are removed via endonucleases [21-23], have been implicated in TNR stability [1,2]. However, the published results are quite controversial. Whereas a lack of the NER function dramatically increases the instability of

TNR sequences in *E. coli* [24,25], a depletion of NER genes (including ERCC1 and XPG) in human cells stabilizes TNR sequences [26]. A recent study suggests that NER proteins play no roles in CAG/CTG HPR in human cell extracts [10]. Whether or not these discrepancies are related to organism- or species-specificity and/or TNR sequence-specificity remains to be determined. Our results in this study suggest that if NER is involved in TNR HPR, it may participate in repair of TNR hairpins located in the nicked DNA strand. This is because the dual incisions observed in repair of the CTG hairpin in the nicked strand (**Fig. 2**) resemble endonuclease cleavages of DNA lesions by NER [21-23]. Further investigations are needed to clarify this issue. Controversies also exist concerning the involvement of FEN-1 in TNR stability. Genetic studies in yeast indicate that deletion of Rad27 (yeast FEN-1 homolog) destabilizes CAG/CTG tracts and other simple repeats [27-29], but cells from *FEN-1* knockout mice do not display TNR instability [30]. Studies using purified FEN-1 protein to process TNR hairpins also gave distinct results [31-33]. Our data suggests that FEN-1, if involved, may only play a role in processing CAG hairpins formed in the nicked strand (**Fig. 3**), and this substrate specificity may explain discrepancies observed in previous studies.

HPR ensures replication fidelity of TNRs

A recent study reported that CAG/CTG HPR occurs in either an error-free and or error-prone manner in human cells [10]. Interestingly, the error-free repair is always observed in DNA substrates with a CAG/CTG hairpin in the continuous strand, and the error-prone repair is always associated with substrates with a hairpin in the nicked strand. The error-prone repair is believed to result from incomplete excision of CAG/CTG hairpins, followed by gap-filling and ligation [10]. These products were referred to as SI-DNAs. However, the SI-DNAs were not observed in our study (**Fig. 1b**, lanes 8 and 11). This discrepancy is likely due to differences in the assay systems used in these two studies. In our study, repair products and intermediates were analyzed by Southern blot and detected using highly specific ^{32}P-labeled DNA probes that recognize repair products and intermediates. In contrast, HPR in the previous study was scored by incorporations of [α-^{32}P]-dNTPs [10]. Although we cannot rule out the possibility that the [α-^{32}P]-dNTP incorporation assay is more sensitive than Southern blot analysis to pick up minor intermediates/products, previous cell-free studies have documented that the ^{32}P-incorporation approach results in substantial products unrelated to heteroduplex repair [34,35]. Therefore, the SI-DNAs could be non-specific minor products and/or intermediates of TNR repair. In particular, they may be derived from primer extensions using incision-generated $(CTG)_n$ fragments (**Fig. 2f**) or $(CAG)_n$ fragments (**Fig. 3d**) as primers in the presence of [α-^{32}P]-dNTPs, side products that would not be detected in our assay.

The nick-directed strand-specific HPR resembles the nature of DNA mismatch repair (MMR), which ensures replication fidelity by targeting repair in the newly synthesized (nicked) strand in a nick-directed manner [36]. Coincidently, PCNA is required for the initiation step of both MMR [18,19] and TNR HPR (**Fig. 5**). It is likely that PCNA, an important factor in DNA replication, may act to direct the strand-specificity for both repair reactions. We therefore conclude that like MMR, the HPR activity described here promotes replication fidelity of TNR tracts in human cells. Therefore, it is expected that defects in this repair pathway will cause TNR instability and human diseases.

It is worth mentioning that despite its role in removing replication-associated insertion/deletion mispairs, mismatch recognition protein MutSβ (MSH2-MSH3) has been shown to promote CAG repeat instability in transgenic animals [37,38]. It is proposed that binding to $(CAG)_n$-hairpins alters MutSβ biochemical and biophysical activities required for MMR, leading to hijacking the MMR process and inhibiting CAG hairpin removal [39]. However, our recent studies reveal that MutSβ exhibits identical properties during its interaction with a mismatch and a CAG hairpin and that MutSβ binding does not interfere with CAG/CTG HPR in human cells extracts [40]. Therefore, if MutSβ is indeed involved in promoting CAG repeat expansions in human cells, it may not be through inhibition of HPR. A recently study suggests that MutSβ may influence CAG/CTG repeat instability by a mechanism involving transcription [41]. Further investigations are required to evaluate the role of MutSβ in TNR stability.

Methods

Preparation of CAG/CTG Hairpin Substrates

We cloned oligonucleotide 111-mer duplexes containing $(5'\text{-CAG-}3')_{35}/(3'\text{-GTC-}5')_{35}$ or $(5'\text{-CTG-}3')_{35}/(3'\text{-GAC-}5')_{35}$ and oligonucleotide 36-mer duplexes containing $(5'\text{-CAG-}3')_{10}/(3'\text{-GTC-}5')_{10}$, or $(5'\text{-CTG-}3')_{10}/(3'\text{-GAC-}5')_{10}$ into HindIII and EcoRI sites of bacterial phage M13mp18 replication form (RF) DNA (**Fig. 1a**) to create M13mp18 derivatives M13mp18-$(CAG)_{35}$, M13mp18-$(CTG)_{35}$, M13mp18-$(CAG)_{10}$, or M13mp18-$(CTG)_{10}$, respectively. We confirmed the individual derivatives by DNA sequencing. Nicked-heteroduplex containing a $(CAG)_{25}$ or $(CTG)_{25}$ hairpin either in the nicked or continuous strand was constructed essentially as described for mismatched heteroduplexes [19]. For example, to generate a heteroduplex containing a $(CAG)_{25}$ hairpin in the complementary (C) strand, M13mp18-$(CTG)_{35}$ RF DNA (containing 35 repeats of CAG in C strand) was first linearized with BglI and then hybridized with M13mp18-$(CTG)_{10}$ single-stranded phage DNA. This hybridization forms a heteroduplex containing $(CAG)_{35}$ in the C strand and $(CTG)_{10}$ in the viral (V) strand, resulting in a $(CAG)_{25}$ hairpin in the C strand. This substrate was designated 5' C-$(CAG)_{25}$ (see **Fig. 1a**), meaning that it has a $(CAG)_{25}$ hairpin in the C strand and a nick 5' to the hairpin. Conversely, substrate 5' C-$(CTG)_{25}$ stands for having a $(CTG)_{25}$ hairpin in C strand and a strand break 5' to the heterology. Two additional substrates, 5' V-$(CAG)_{25}$ and 5' V-$(CTG)_{25}$, were also similarly prepared (see **Fig. 1a**). The nick for all substrates is at the BglI site in the C strand, and is 164 bp or 149 bp away from the hairpin, depending if the hairpin is in the C or V strand, respectively.

Cell Culture and Nuclear Extract Preparation

We cultured HeLa S_3 cells to a density of 5×10^5 cells ml^{-1} in RPMI 1640 with 5% (v/v) FBS (Hyclone) and 4 mM glutamine at 37 °C in a 5% CO_2 atmosphere, and prepared nuclear extracts according to Holmes et al. [34].

CAG/CTG hairpin repair and Analysis of Repair Intermediates

Unless mentioned otherwise, we performed CAG/CTG hairpin repair (HPR) by Southern blot analysis as described [13,14]. Briefly, 42 fmol of DNA heteroduplex were incubated with

100 μg of HeLa nuclear extracts in a 40-μl reaction containing 20 mM Tris-HCl (pH 7.6), 110 mM KCl, 5 mM $MgCl_2$, 1.5 mM ATP, 1 mM glutathione, and 0.1 mM each of the four dNTPs at 37°C for 30 min. Reactions were terminated by incubating with protease K (30 μg ml^{-1}) at 37°C for 20 min. DNA was recovered by sequential phenol extraction and ethanol precipitation and digested with BsrBI and BglI. The resulting DNA products were separated on a 6% denaturing polyacrylamide gel, followed by electrotransferring to nylon membrane. The membrane was probed with a ^{32}P-end labeled oligonucleotide specifically annealing to the BsrBI-BglI fragment in the nicked strand (see **Fig. 1**) to score for conversion of 35 CAG/CTG repeats to 10 CAG/CTG repeats or vice versa. Repair products, as well as unrepaired molecules, were visualized by exposing to X-Ray film. Repair efficiency was quantified by KODAK Image Station 2000 (Kodak, Rochester, NY).

To analyze the repair intermediates, we conducted *in vitro* assay in the absence of exogenous dNTPs and the presence of 0.15 mM of aphidicolin and 0.5 mM ddNTPs to block repair DNA synthesis. Unless specified, DNA samples were digested with BsmBI and Bau36I. After electrophoresis, reaction intermediates were subjected to Southern blot analysis using ^{32}P-labeled oligonucleotide probes specifically annealing to the BsmBI-Bau36I fragment in the nicked strand.

Purification of PCNA, p21C and FEN-1 Proteins

We expressed human recombinant PCNA, C-terminal domain of $p21^{Cip1/WAF}$ (p21C) in *E. coli* BL21 (DE3) cells and purified them to homogeneity as previously described [19]. The human FEN-1 expression vector was a gift from Binghui Shen (City of Hope) and the protein was overexpressed in *E. coli* BL21 (DE3) cells and purified to homogeneity as described [43].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgement

We thank Binghui Shen (City of Hope National Medical Center), Jerard Hurwitz (Memorial Sloan Kettering Cancer Center), and Anindya Dutta (University of Virginia) for FEN-1, PCNA, and p21C expression vectors, respectively, and Dangsheng Li (Shanghai Institutes for Biological Sciences, China) for helpful discussions. The work was supported in part by NIH grants GM072756 (to G.-M.L.) and CA104333 (to L.G.). G.-M.L. is the James-Gardner Endowed Chair in Cancer Research.

References

1. Mirkin SM. Expandable DNA repeats and human disease. Nature. 2007; 447:932–40. [PubMed: 17581576]
2. Pearson CE, Nichol Edamura K, Cleary JD. Repeat instability: mechanisms of dynamic mutations. Nat Rev Genet. 2005; 6:729–42. [PubMed: 16205713]
3. Gordenin DA, Kunkel TA, Resnick MA. Repeat expansion--all in a flap? Nat Genet. 1997; 16:116–8. [PubMed: 9171819]
4. Kang S, Jaworski A, Ohshima K, Wells RD. Expansion and deletion of CTG repeats from human disease genes are determined by the direction of replication in E. coli. Nat Genet. 1995; 10:213–8. [PubMed: 7663518]

5. Miret JJ, Pessoa-Brandao L, Lahue RS. Orientation-dependent and sequence-specific expansions of CTG/CAG trinucleotide repeats in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A. 1998; 95:12438–43. [PubMed: 9770504]
6. Richards RI, Sutherland GR. Simple repeat DNA is not replicated simply. Nat Genet. 1994; 6:114–6. [PubMed: 8162063]
7. Gacy AM, Goellner G, Juranic N, Macura S, McMurray CT. Trinucleotide repeats that expand in human disease form hairpin structures in vitro. Cell. 1995; 81:533–40. [PubMed: 7758107]
8. Moore H, Greenwell PW, Liu CP, Arnheim N, Petes TD. Triplet repeats form secondary structures that escape DNA repair in yeast. Proc Natl Acad Sci U S A. 1999; 96:1504–9. [PubMed: 9990053]
9. Petruska J, Arnheim N, Goodman MF. Stability of intrastrand hairpin structures formed by the CAG/CTG class of DNA triplet repeats associated with neurological diseases. Nucleic Acids Res. 1996; 24:1992–8. [PubMed: 8668527]
10. Panigrahi GB, Lau R, Montgomery SE, Leonard MR, Pearson CE. Slipped (CTG)*(CAG) repeats can be correctly repaired, escape repair or undergo error-prone repair. Nat Struct Mol Biol. 2005; 12:654–62. [PubMed: 16025129]
11. Corrette-Bennett SE, Parker BO, Mohlman NL, Lahue RS. Correction of large mispaired DNA loops by extracts of Saccharomyces cerevisiae. J Biol Chem. 1999; 274:17605–11. [PubMed: 10364197]
12. Littman SJ, Fang WH, Modrich P. Repair of large insertion/deletion heterologies in human nuclear extracts is directed by a 5′ single-strand break and is independent of the mismatch repair system. J Biol Chem. 1999; 274:7474–81. [PubMed: 10066813]
13. McCulloch SD, Gu L, Li GM. Nick-dependent and -independent processing of large DNA loops in human cells. J Biol Chem. 2003; 278:50803–9. [PubMed: 14522965]
14. Zhang Y, et al. Reconstitution of 5′-directed human mismatch repair in a purified system. Cell. 2005; 122:693–705. [PubMed: 16143102]
15. Hartenstine MJ, Goodman MF, Petruska J. Base stacking and even/odd behavior of hairpin loops in DNA triplet repeat slippage and expansion with DNA polymerase. J Biol Chem. 2000; 275:18382–90. [PubMed: 10849445]
16. Bhattacharyya S, Lahue RS. Srs2 helicase of Saccharomyces cerevisiae selectively unwinds triplet repeat DNA. J Biol Chem. 2005; 280:33311–7. [PubMed: 16085654]
17. Easteal, S.; McLeod, N.; Reed, K. DNA profiling: principles, pitfalls and potential. Harwood Academic Publishers; 1993.
18. Umar A, et al. Requirement for PCNA in DNA mismatch repair at a step preceding DNA resynthesis. Cell. 1996; 87:65–73. [PubMed: 8858149]
19. Guo S, et al. Differential requirement for proliferating cell nuclear antigen in 5′ and 3′ nick-directed excision in human mismatch repair. J Biol Chem. 2004; 279:16912–7. [PubMed: 14871894]
20. Chen J, Jackson PK, Kirschner MW, Dutta A. Separate domains of p21 involved in the inhibition of Cdk kinase and PCNA. Nature. 1995; 374:386–8. [PubMed: 7885482]
21. Friedberg, EC., et al. DNA Repair and Mutagenesis. ASM Press; Washington, D. C.: 2006.
22. Sancar A, Lindsey-Boltz LA, Unsal-Kacmaz K, Linn S. Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. Annu Rev Biochem. 2004; 73:39–85. [PubMed: 15189136]
23. Wood RD, Mitchell M, Sgouros J, Lindahl T. Human DNA repair genes. Science. 2001; 291:1284–9. [PubMed: 11181991]
24. Oussatcheva EA, Hashem VI, Zou Y, Sinden RR, Potaman VN. Involvement of the nucleotide excision repair protein UvrA in instability of CAG*CTG repeat sequences in Escherichia coli. J Biol Chem. 2001; 276:30878–84. [PubMed: 11413147]
25. Parniewski P, Bacolla A, Jaworski A, Wells RD. Nucleotide excision repair affects the stability of long transcribed (CTG*CAG) tracts in an orientation-dependent manner in Escherichia coli. Nucleic Acids Res. 1999; 27:616–23. [PubMed: 9862988]
26. Lin Y, Wilson JH. Transcription-induced CAG repeat contraction in human cells is mediated in part by transcription-coupled nucleotide excision repair. Mol Cell Biol. 2007; 27:6209–17. [PubMed: 17591697]

27. Freudenreich CH, Kantrow SM, Zakian VA. Expansion and length- dependent fragility of CTG repeats in yeast. Science. 1998; 279:853–6. [PubMed: 9452383]
28. Liu Y, Zhang H, Veeraraghavan J, Bambara RA, Freudenreich CH. Saccharomyces cerevisiae flap endonuclease 1 uses flap equilibration to maintain triplet repeat stability. Mol Cell Biol. 2004; 24:4049–64. [PubMed: 15082797]
29. Tishkoff DX, Filosi N, Gaida GM, Kolodner RD. A novel mutation avoidance mechanism dependent on S. cerevisiae RAD27 is distinct from DNA mismatch repair. Cell. 1997; 88:253–63. [PubMed: 9008166]
30. van den Broek WJ, Nelen MR, van der Heijden GW, Wansink DG, Wieringa B. Fen1 does not control somatic hypermutability of the (CTG)(n)*(CAG)(n) repeat in a knock-in mouse model for DM1. FEBS Lett. 2006; 580:5208–14. [PubMed: 16978612]
31. Henricksen LA, Tom S, Liu Y, Bambara RA. Inhibition of flap endonuclease 1 by flap secondary structure and relevance to repeat sequence expansion. J Biol Chem. 2000; 275:16420–7. [PubMed: 10748145]
32. Singh P, Zheng L, Chavez V, Qiu J, Shen B. Concerted action of exonuclease and Gap-dependent endonuclease activities of FEN-1 contributes to the resolution of triplet repeat sequences (CTG)n- and (GAA)n-derived secondary structures formed during maturation of Okazaki fragments. J Biol Chem. 2007; 282:3465–77. [PubMed: 17138563]
33. Spiro C, et al. Inhibition of FEN-1 processing by DNA secondary structure at trinucleotide repeats. Mol Cell. 1999; 4:1079–85. [PubMed: 10635332]
34. Holmes J Jr. Clark S, Modrich P. Strand-specific mismatch correction in nuclear extracts of human and Drosophila melanogaster cell lines. Proc Natl Acad Sci U S A. 1990; 87:5837–41. [PubMed: 2116007]
35. Thomas DC, Roberts JD, Kunkel TA. Heteroduplex repair in extracts of human HeLa cells. J Biol Chem. 1991; 266:3744–51. [PubMed: 1995629]
36. Li GM. Mechanisms and functions of DNA mismatch repair. Cell Res. 2008; 18:85–98. [PubMed: 18157157]
37. Manley K, Shirley TL, Flaherty L, Messer A. Msh2 deficiency prevents in vivo somatic instability of the CAG repeat in Huntington disease transgenic mice. Nat Genet. 1999; 23:471–3. [PubMed: 10581038]
38. Owen BA, et al. (CAG)(n)-hairpin DNA binds to Msh2-Msh3 and changes properties of mismatch recognition. Nat Struct Mol Biol. 2005; 12:663–70. [PubMed: 16025128]
39. McMurray CT. Hijacking of the mismatch repair system to cause CAG expansion and cell death in neurodegenerative disease. DNA Repair (Amst). 2008; 7:1121–34. [PubMed: 18472310]
40. Tian L, et al. Mismatch recognition protein MutSβ does not hijack (CAG)n hairpin repair in vitro. J Biol Chem In press. 2009
41. Lin Y, Dion V, Wilson JH. Transcription promotes contraction of CAG repeat tracts in human cells. Nat Struct Mol Biol. 2006; 13:179–80. [PubMed: 16388310]
42. Pearson CE, et al. Slipped-strand DNAs formed by long (CAG)*(CTG) repeats: slipped-out repeats and slip-out junctions. Nucleic Acids Res. 2002; 30:4534–47. [PubMed: 12384601]
43. Frank G, Qiu J, Zheng L, Shen B. Stimulation of eukaryotic flap endonuclease-1 activities by proliferating cell nuclear antigen (PCNA) is independent of its in vitro interaction via a consensus PCNA binding region. J Biol Chem. 2001; 276:36295–302. [PubMed: 11477073]

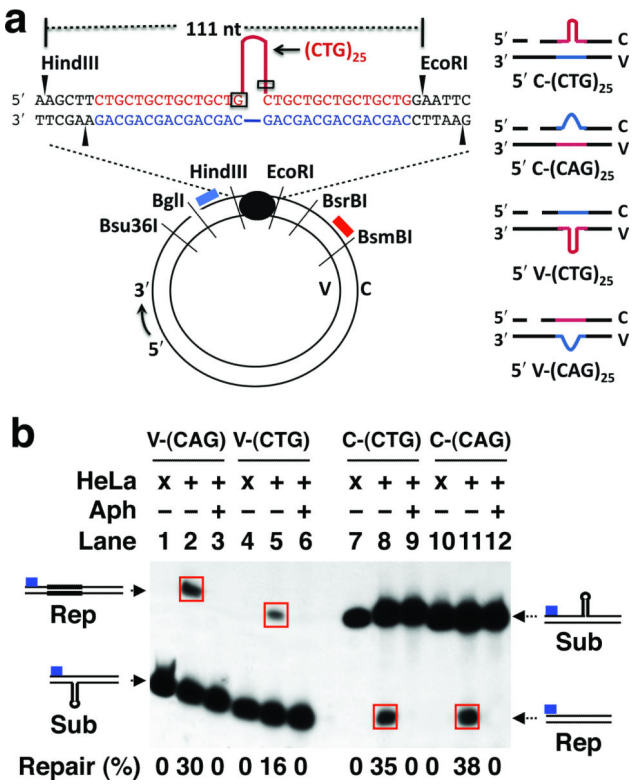

Figure 1. Error-Free CAG/CTG HPR in Human Extracts

(**a**) DNA hairpin substrates. Circular DNA substrates contain a $(CAG)_{25}$ or $(CTG)_{25}$ hairpin either in the complementary (C) or viral (V) strand and a strand break 5' to the hairpin (see Methods). The CAG (blue type/line) and CTG (red type/line) repeats are located between HindIII and EcoRI restriction enzyme sites. Sequence compositions of the $(CTG)_{25}$ hairpin in C strand and the predicted secondary structure of the hairpin are given, while other substrates are depicted by colored lines at right. Different shapes for CAG and CTG hairpin heteroduplexes reflect their secondary structure as a random-coil and hairpin, respectively [42]. Blue and red bars represent oligonucleotide probes complementary to the indicated locations in the nicked strand near the BglI and BsmBI sites, respectively. (**b**) CAG/CTG HPR assay. Individual heteroduplexes were incubated with active (+) or heat-inactivated (x) HeLa nuclear extracts in the presence or absence of aphidicolin (Aph), as indicated. Bands with red squares indicate HPR products. Diagrams on the left or right side of the gel show substrates (Sub) and repair products (Rep) for hairpins in the V or C strand, respectively. Blue bar shows ^{32}P-labed oligonucleotide probe annealing near to the BglI site (see **a**) in the nicked strand.

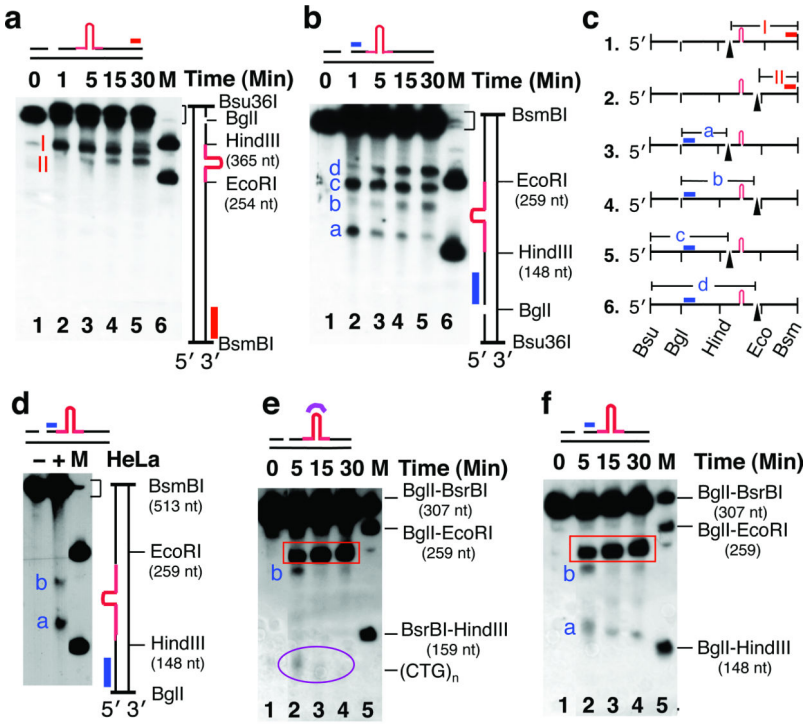

Figure 2. Removal of Nicked Strand CTG Hairpin by Dual Incisions

(**a**) and (**b**) Analyses of repair intermediates. Repair intermediates were obtained by incubating C-$(CTG)_{25}$ substrate with HeLa extracts under conditions of no DNA synthesis, and DNA samples were digested with Bsu36I and BsmBI and analyzed by Southern hybridization using an oligonucleotide probe (the BsmBI probe, see **Fig. 1**) complementary to the nicked strand near the BsmBI site (**a**) or a probe (the BglI probe, see **Fig. 1a**) complementary to nicked strand near the BglI site (**b**). (**c**) Schematic diagrams of individual repair intermediates identified in **a** (red type) and **b** (blue type). (**d**) Detecting repair intermediates by the BglI probe after digesting DNA samples with BglI and BsmBI. (**e**) and (**f**) Repair assays in time course. Reactions were performed under normal repair conditions as described in **Fig. 1b**, except for the incubation times as indicated. Repair products were detected by a ^{32}P-labeled $(CAG)_{10}$ oligonucleotide probe (**e**) or the BglI probe (**f**).

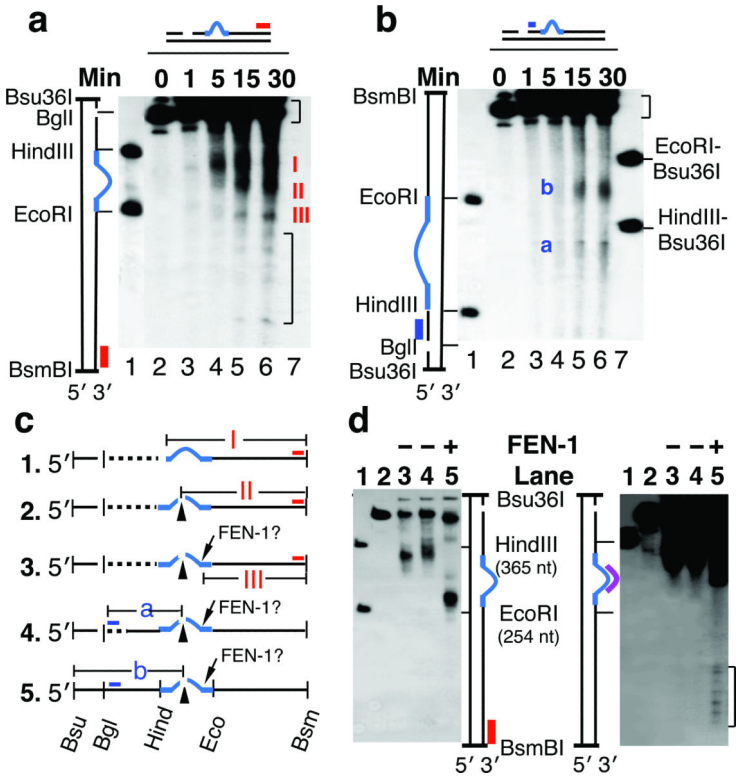

Figure 3. Analysis of Repair Intermediates of Substrate C-$(CAG)_{25}$

Repair intermediates were obtained and analyzed as in **Fig. 2**. (**a**) Intermediates detected by an oligonucleotide probe complementary to the nicked strand near the BsmBI site (the BsmBI probe, see **Fig. 1a**). (**b**) Intermediates detected by a probe complementary to nicked strand near the BglI site (the BglI probe, see **Fig. 1a**). (**c**) Schematic diagrams of individual intermediates obtained in **a** (diagrams 1-3) and **b** (diagrams 4 and 5). Dashed fragments indicate regions where DNA excision may occur, which produces product I. (**d**) Cleavage of C-$(CAG)_{25}$ repair intermediates by FEN-1. Repair intermediates (lane 3) were obtained by incubating C-$(CAG)_{25}$ with HeLa extracts for 5 min and treated with or without purified FEN-1 as indicated, followed by Southern analysis using the BsmBI probe (left panel) or a $(CTG)_{10}$ oligonucleotide probe (right panel, purple line).

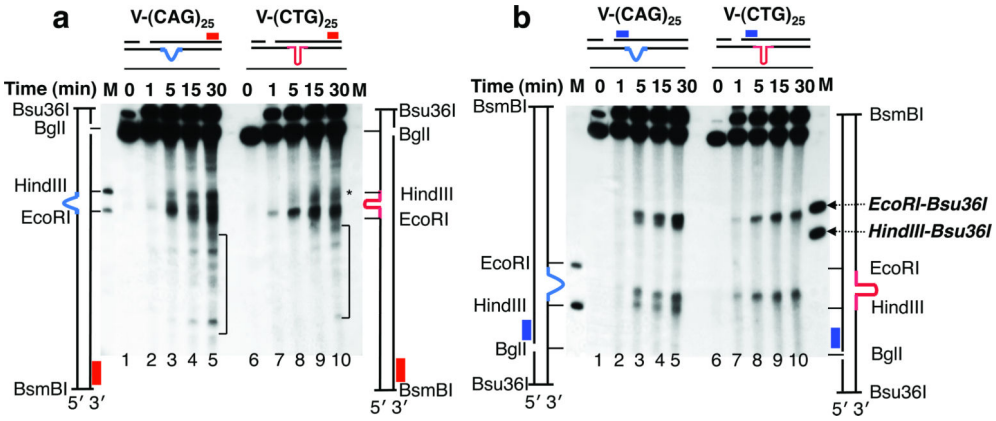

Figure 4. Analysis of Repair Intermediates of Substrates V-(CAG)$_{25}$ and V-(CTG)$_{25}$

Repair intermediates were obtained as described in **Fig. 2**, and analyzed by Southern hybridization using the BsmBI probe (**a**) or the BglI probe (**b**). Reactions 1-5 and 6-10 in each panel show repair intermediates for substrates V-(CAG)$_{25}$ and V-(CTG)$_{25}$, respectively.

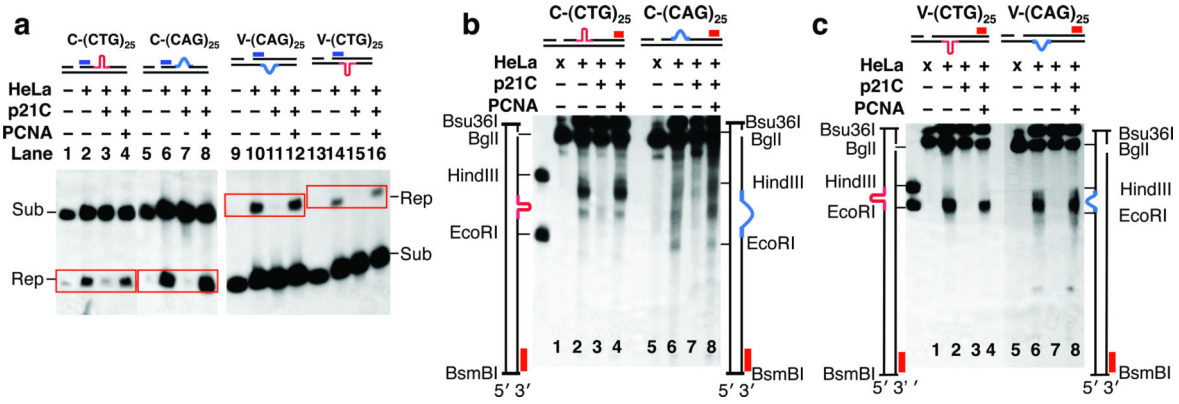

Figure 5. PCNA Is Required for CAG/CTG HPR

(**a**) Inhibition of CAG/CTG HPR by p21C. Repair assays were performed as described in **Fig. 1b**, but in the presence or absence of p21C and/or PCNA. Repair products are highlighted by red rectangles. Sub and Rep stand for substrate and repair product, respectively. (**b**) and (**c**) PCNA is required at or prior to the incision reaction. Repair reactions were carried out in p21C-inactivated HeLa extracts in the presence or absence of exogenous PCNA, as indicated, and under the conditions of no DNA synthesis. Repair intermediates of individual substrates were analyzed by Southern hybridization using the BsmBI probe. Letter “x” indicates heat-inactivated HeLa extracts. Schematic diagram of DNA fragments on the left or right side of an autograph is specific for the left or right part of the gel, respectively.

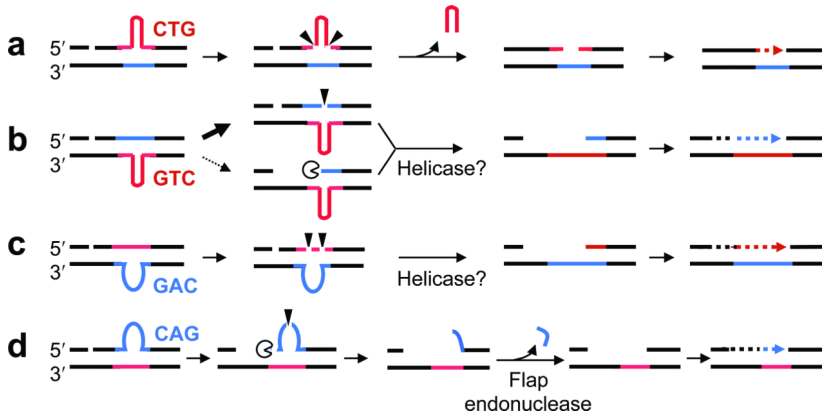

Figure 6. Model of CAG/CTG HPR in Human Cells

CAG/CTG HPR specifically targets repeat tracts for incisions in the nicked DNA strand, and the resulting gap is filled by an aphidicolin-sensitive DNA polymerase using the continuous strand as a template. However, distinct endonuclease activities are required for processing structure- and sequence-specific CAG/CTG hairpins. While repair of a CTG hairpin located in the nicked strand involves an incision on either side of the hairpin (**a**), removal of a CTG hairpin located in the continuous strand is via a single incision in the nicked strand opposite the hairpin (**b**). Conversely, removing a CAG hairpin in the continuous strand seems to require only one type of endonuclease activity (**c**), but removing the same hairpin in the nicked strand involves at least three different nuclease activities (**d**). For more details, see Discussion.