BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

The Nucleotide Sequence, DNA Damage Location, and Protein Stoichiometry Influence the Base Excision Repair Outcome at CAG/CTG Repeats

Agathi-Vasiliki Goula,[†] Christopher E. Pearson,[‡] Julie Della Maria,[§] Yvon Trottier,[†] Alan E. Tomkinson,[§] David M. Wilson, III,[∥] and Karine Merienne*[,†]

[†]Department of Neurogenetics and Translational Medicine, Institute of Genetics and Molecular and Cellular Biology (IGBMC), UMR 7104-CNRS/INSERM/UdS, Illkirch, France

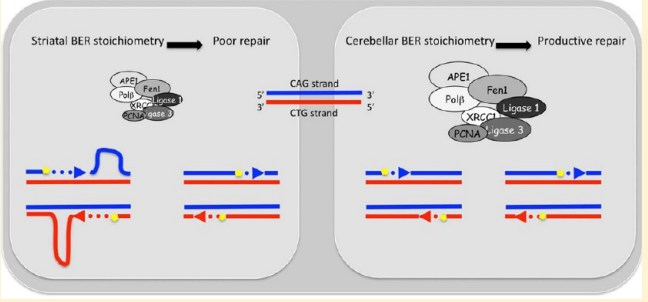
[‡]Genetics and Genome Biology, The Hospital for Sick Children, TMDT Building, 101 College Street, 15th Floor, Room 15-312, East Tower, Toronto, ON, M5G 1L7 Canada, and Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada

[§]Department of Radiation Oncology and Marlene and Stewart Greenebaum Cancer Center, University of Maryland School of Medicine, Baltimore, Maryland 21201, United States

[∥]Laboratory of Molecular Gerontology, National Institute on Aging, National Institutes of Health, Baltimore, Maryland 21224, United States

Supporting Information

ABSTRACT: Expansion of CAG/CTG repeats is the underlying cause of >14 genetic disorders, including Huntington's disease (HD) and myotonic dystrophy. The mutational process is ongoing, with increases in repeat size enhancing the toxicity of the expansion in specific tissues. In many repeat diseases, the repeats exhibit high instability in the striatum, whereas instability is minimal in the cerebellum. We provide molecular insights into how base excision repair (BER) protein stoichiometry may contribute to the tissue-selective instability of CAG/CTG repeats by using specific repair assays. Oligonucleotide substrates with an abasic site were mixed with either reconstituted BER protein stoichiometries mimicking the levels present in HD mouse striatum or cerebellum, or with protein extracts prepared from HD mouse striatum or cerebellum. In both cases, the repair efficiency at CAG/CTG repeats and at control DNA sequences was markedly reduced under the striatal conditions, likely because of the lower level of APE1, FEN1, and LIG1. Damage located toward the 5′ end of the repeat tract was poorly repaired, with the accumulation of incompletely processed intermediates as compared to an AP lesion in the center or at the 3′ end of the repeats or within control sequences. Moreover, repair of lesions at the 5′ end of CAG or CTG repeats involved multinucleotide synthesis, particularly at the cerebellar stoichiometry, suggesting that long-patch BER processes lesions at sequences susceptible to hairpin formation. Our results show that the BER stoichiometry, nucleotide sequence, and DNA damage position modulate repair outcome and suggest that a suboptimal long-patch BER activity promotes CAG/CTG repeat instability.



Trinucleotide repeat (TNR) expansions are responsible for >15 neurological, neuromuscular, and neurodegenerative genetic disorders.[1,2] This family of disorders includes CAG/CTG repeat-associated diseases, such as Huntington's disease (HD), and nine other polyglutamine (polyQ) disorders, as well as myotonic dystrophy type 1 (DM1). The TNR tracts need to reach a threshold length of 30−50 units to become genetically unstable and trigger pathogenesis. At this size, the length of the mutant allele continuously changes, in both germline and somatic tissues, most often incurring expansions. However, all tissues do not undergo repeat instability to an equal degree, and this tissue selectivity varies between diseases.[1,3] For instance, in HD and other polyQ disorders, both the striatum and the cortex exhibit increased CAG/CTG instability and the largest degree of expansion, while the same repeats remain stable and shorter in length in the cerebellum.[4−9] In DM1, continuous expansion of the mutated allele is observed in the heart, skeletal muscle, and cortex, whereas TNR expansion is limited in the cerebellum.[10−14] In DM1 and HD, instability is most prevalent in affected tissues, presumably accelerating the progression of the diseases.[5,15−17] It is therefore important to delineate the molecular mechanisms of tissue-specific TNR instability to improve our understanding of the pathogenesis of the different disorders.

The mechanisms underlying the tissue selectivity of TNR instability remain unclear. Gene-specific *cis* elements and tissue-

Received: March 31, 2012
Published: April 12, 2012

dx.doi.org/10.1021/bi300410d | *Biochemistry* 2012, 51, 3919−3932

specific epigenetic modifications have been implicated.[18] In addition, tissue-specific *trans* factors may contribute to the process. For instance, MSH2 and MSH3, which participate in the mismatch repair (MMR) pathway, have been implicated in CAG/CTG instability. Notably, mouse models for DM1 or HD that are deficient in *Msh2* or *Msh3* exhibit reduced CAG/CTG instability.[19−24] Because CAG/CTG repeats have a high propensity to form stable secondary DNA structures such as hairpins in vitro,[25,26] it has been hypothesized that the aberrant processing of these structures by MMR promotes instability. Interestingly, recent studies showed that downregulation of MMR genes upon differentiation of DM1-derived human embryonic stem cells correlates with decreased CAG/CTG instability,[27] suggesting that the contribution of MMR to CAG/CTG instability is regulated in a tissue-specific manner.

Somatic CAG/CTG instability is reduced in HD mice deficient for the DNA glycosylase *Ogg1*, indicating that base excision repair (BER) also contributes to CAG/CTG instability.[28] In yeast, the flap endonuclease (FEN1) and DNA ligase I (LIG1), two proteins involved in long-patch BER (LP-BER), modulate the instability of CAG/CTG repeats.[29,30] Human cells and DM1 transgenic mice also revealed a role for LIG1 in CAG/CTG instability.[2,3] In classic BER, removal of an oxidative base lesion by a DNA glycosylase results in the formation of an abasic (AP) site, which is then cleaved by an AP endonuclease (APE1 in mammals).[31,32] This DNA strand break is subsequently processed by either single-nucleotide BER (SN-BER) or LP-BER. In SN-BER, POLβ incorporates a single nucleotide and excises the remaining 5′-abasic fragment, prior to ligation by DNA ligase III α (LIG3α) in complex with X-ray cross-complementing 1 (XRCC1) protein. In LP-BER, FEN1 removes the 5′-flap structure generated during the multinucleotide synthesis step mediated by POLβ or a replicative DNA polymerase prior to ligation by LIG1.[31,32]

We previously found that in wild-type and HD transgenic mice, BER proteins and associated enzymatic activities are reduced in the striatum in comparison to the cerebellum, though to different levels.[33] In particular, the level of FEN1 protein is greatly decreased in the striatum, leading to a lower FEN1:POLβ ratio compared to that in the cerebellum. We therefore proposed that a reduced level of FEN1 in the striatum would contribute to the high level of TNR instability seen in the striatum of HD animals, because of impaired coordination between DNA synthesis and 5′-flap removal in the striatum during LP-BER at CAG/CTG repeats. However, a previous study, which used in vitro repair assays and CAG oligonucleotide substrates with a tetrahydrofuran (THF) abasic site analogue that can be processed only by LP-BER, suggested that an increased level of FEN1 promotes CAG expansion by facilitating ligation of hairpins formed by strand slippage.[34] Thus, the roles of BER protein stoichiometry and LP-BER in tissue-selective TNR instability have remained unclear.

Repair outcomes at nicked CAG/CTG substrates with slipped-out repeats have been reported using mammalian cell extracts.[35−38] In these assays, repair outcome and efficiencies clearly depended upon nick location and slip-out sequence (CAG vs CTG). Interestingly, repair efficiency was significantly increased when the slip-out was located on the CAG strand in comparison to the CTG strand. Furthermore, the nick location (in the slipped vs continuous strand or 5′ vs 3′ of the slip-out) dramatically affected repair outcome. In all reports to date, in examining BER processing of trinucleotide repeats, the DNA lesion has been placed in the CAG strand at the 5′ end (the first repeat unit)[28,34] or within a CAG hairpin.[39,40] Thus, whether the position of an oxidative DNA lesion within a CAG/CTG repeat sequence influences repair is unknown.

To address the role of BER protein stoichiometry, nucleotide sequence context, and lesion position on damage-containing TNR processing, we use herein repair assays that employ oligonucleotide substrates harboring an abasic lesion located upstream (5′-oriented), downstream (3′-oriented), or centrally within the CAG or CTG strand. Moreover, repair assays were conducted either with purified BER protein mixtures that reflect the stoichiometry in the striatum or cerebellum of HD mice or with protein extracts prepared from the striatal or cerebellar tissue of HD mice. Our data demonstrate that lesions are less efficiently repaired at the striatal protein stoichiometry as compared to the cerebellar stoichiometry, the latter of which is characterized by high levels of the LP-BER enzymes FEN1 and LIG1. In addition, our studies indicate that lesions located at the 5′ end of the repeat lead to formation of increased amounts of intermediate repair products, indicative of incomplete repair, and the production of fewer full-length repair products, than when lesions are positioned at the 3′ end of the repeat. Furthermore, repair of lesions located toward the 5′ end in a CAG or CTG repeat involve multinucleotide synthesis, particularly at the cerebellar stoichiometry, suggesting that LP-BER conducts processing of substrates prone to formation of secondary structures. Our data provide evidence that the poor repair susceptibility of CAG/CTG sequences and the suboptimal striatal LP-BER activity both contribute to the inefficient repair at CAG/CTG repeats in the striatum. Our data support a model in which inefficient LP-BER of some postmitotic tissues increases the risk of somatic CAG/CTG repeat instability associated with specific diseases.

MATERIALS AND METHODS

Materials. Human recombinant proteins LIG1, LIG3/XRCC1, PCNA, APE1, POLβ, and FEN1 were purified as previously described.[41−46] Uracil-DNA glycosylase (UNG) from *Escherichia coli* was a generous gift from the late D. Mosbaugh (Oregon State University, Corvallis, OR). Primary rabbit and mouse antibodies were purchased from MBL (mouse α-LIG1, K 0190-3), BD transduction Laboratories (mouse α-LIG3, 611876), Abcam (rabbit α-APE1, ab92744), Sigma (mouse α-PCNA, P 8825), and Chemicon (mouse $\alpha-\beta$-tubulin). Rabbit α-XRCC1 was a kind gift from P. J. McKinnon (St. Jude Children's Research Hospital, Memphis, TN). DNA oligonucleotides were purchased from Eurogentec. The radionucleotides, [γ-^{32}P]ATP (7000 μCi/mmol), [α-^{32}P]dCTP (3000 μCi/mmol), and [α-^{32}P]dGTP (3000 μCi/mmol), were from Perkin-Elmer.

Mice. Hemizygous R6/1 HD transgenic mice from the Jackson Laboratory were maintained on a mixed CBAxC57BL/6 genetic background and were genotyped as described previously.[47] The experiments were approved by the ethical committee CREMEAS (Comite Regional d'Ethique en Matiere d'Experimentation Animale de Strasbourg).

Western Blotting. Whole cell extracts from mouse cerebellum and striatum of R6/1 mice were prepared as previously described.[33] For Western blot analysis, rabbit and mouse antibodies against mouse endogenous and human recombinant BER proteins (see above) were used at 1:500 (α-LIG1), 1:1000 (α-LIG3, α-XRCC1, and α-PCNA), or 1:100000 (α-APE1) dilutions and were detected with appropriate α-rabbit or α-mouse peroxidase-conjugated secon-

dary antibodies (Jackson immunoResearch Laboratories) and the ECL chemiluminescence kit (Pierce or Millipore). Signals were imaged on radiographic films, captured with GeneSnap, and quantified with GeneTool on a Syngene Chemigenus XE machine.

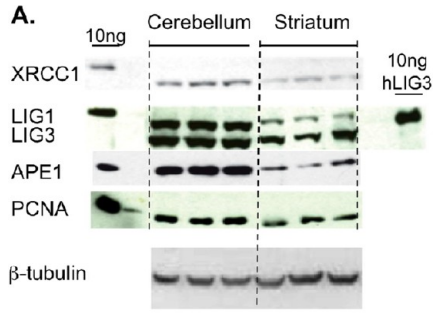
Preparation of AP Substrates. To create the AP–DNA duplex substrates, the target oligonucleotide strand harbored a uracil (U) modification that replaced a cytosine or a guanine residue, and this strand was hybridized with the complementary oligonucleotide at a molar ratio of 1:1 (see Figure 2). We note that the sequences of the CAG/CTG substrates and the control substrates were based upon the sequences described by Liu et al.[34] and Petermann et al.,[48] respectively. Subsequently, the optimal conditions for the complete removal of uracil from the various DNA substrates by UNG were determined (Figure S1 of the Supporting Information). In brief, 0.5 pmol of double-strand substrate was incubated with 2.9 units of UNG in 81 mM KCl, 2.5 mM $MgCl_2$, 3.125 mM HEPES-KOH (pH 7.7), 1% glycerol, and 0.25 mM EDTA for 1 h at 37 °C, and the reaction mixture was stored on ice for 10 min until it was further needed. For 5′-γ-^{32}P labeling experiments, U-containing substrates were radiolabeled at the 5′ end using [γ-^{32}P]ATP and T4 polynucleotide kinase (New England Biolabs) for 40 min at 37 °C, prior to annealing with the template strand.

Reconstitution Repair Assays. The AP-DNA substrates were incubated with a mixture of BER proteins as described throughout. The repair reactions were performed by incubating 0.5 pmol of AP-DNA substrate with 0.1 nmol of POLβ, 0.5 nmol of APE1, 0.75 nmol of FEN1, 0.1 nmol of PCNA, 0.05 nmol of LIG1, and 0.75 nmol each of LIG3 and XRCC1 (striatal stoichiometry) or 0.1 nmol of POLβ, 1.5 nmol of APE1, 0.6 nmol of FEN1, 0.2 nmol of PCNA, 0.2 nmol of LIG1, and 0.2 nmol each of LIG3 and XRCC1 (cerebellar stoichiometry). The reactions were conducted at 37 °C for the indicated period of time in 75 mM KCl, 25 mM $MgCl_2$, 3.125 mM HEPES-KOH (pH 7.7), 1% glycerol, and 0.25 mM EDTA, as previously described.[33] For radioincorporation repair assays, the reaction buffer was supplemented with 20 μM dNTPs and 125 nM [α-^{32}P]dCTP or [α-^{32}P]dGTP, depending on the substituted base (a cytosine or guanine). For 5′-γ-^{32}P labeling experiments, the reaction buffer was supplemented with 20 μM dNTPs only. The reactions were stopped by addition of stop buffer (98% formamide, 20 mM EDTA, bromophenol blue, and xylene cyanol) and heating at 95 °C for 5 min. The reaction products were resolved by 15% polyacrylamide urea denaturing gel electrophoresis, imaged on a Typhoon phosphoimager, and quantified with ImageQuant TL.

Tissue-Based Repair Assays. Tissue extracts were prepared as previously described.[33] Briefly, striata and cerebella were dissected from HD transgenic mice and control wild-type littermates, homogenized in 10 mM HEPES-KOH (pH 7.7), 0.5 mM $MgCl_2$, 10 mM KCl, and 1 mM DTT, and centrifuged at 2000*g* and 4 °C for 10 min. The pellet was resuspended in 20 mM HEPES-KOH (pH 7.7), 0.5 mM $MgCl_2$, 420 mM NaCl, 0.2 mM EDTA, 25% glycerol with protease inhibitors (Complete Protease Inhibitor Cocktail EDTA-free, Roche), and 1 mM DTT and gently stirred at 4 °C for 20–30 min. The suspension was centrifuged at 14000*g* and 4 °C for 15 min. The supernatant was dialyzed against 40 mM HEPES-KOH (pH 7.7), 50 mM KCl, and 2 mM DTT buffer overnight at 4 °C. Cell extract concentrations were determined using a Bradford assay, 0.5 pmol of [γ-^{32}P]ATP 5′-radiolabeled UNG-pretreated DNA substrate was incubated with 2 μg of cell extract and assay buffer [81 mM KCl, 2.5 mM $MgCl_2$, 750 μM dNTPs, 3.125 mM HEPES-KOH (pH 7.7), 1% glycerol, and 0.25 mM DTT] at 37 °C for 90 min. The samples were then treated with 0.5% SDS and 0.8 mg/mL proteinase K, heated at 55 °C for 15 min, and purified by phenol/chloroform extraction. The reactions were further prepared via addition of an equal volume of stop buffer (98% formamide, 10 mM EDTA, bromophenol blue, and xylene cyanol). The reaction products were resolved and quantified as described for the reconstitution repair assays.

RESULTS

LIG1 and FEN1 Protein Levels Are Higher in the Cerebellum Than in the Striatum. To determine the potential role of BER protein stoichiometry in the tissue-selective instability of CAG/CTG repeats in HD, we determined the relative molar levels of key BER proteins, APE1, LIG1, LIG3, XRCC1, and PCNA, in the cerebellum and striatum of HD transgenic mice by Western blot analysis (Figure 1A). β-Tubulin was used as a loading control (Figure 1A). The molar levels of POLβ, the main BER polymerase in the brain,[49,50] and FEN1 were determined previously (ref 33



B.

Protein	Cb	St
POLβ	1	1
FEN1	6	1.5
XRCC1	3.5	1.5
LIG1	2	0.4
LIG3	2	1.5
APE1	15	5
PCNA	2	1

Numbers correspond to protein molar ratios relative to POLβ

Figure 1. BER protein stoichiometry in the cerebellum and striatum of HD mice. (A) Steady state levels of XRCC1, LIG1, LIG3, APE1, and PCNA proteins in the striatum and cerebellum of HD mice were determined by Western blot analysis, using purified human recombinant proteins as the reference; 100 μg of whole cell extracts prepared from the striatum or cerebellum of the same HD mice was run on a sodium dodecyl sulfate–polyacrylamide gel (three mice were used per gel) with 10 ng of recombinant XRCC1, LIG1, LIG3, APE1, or PCNA and probed with α-XRCC1, α-LIG1, α-LIG3, α-APE1, and α-PCNA antibodies. The α–β-tubulin antibody was used as a loading control. Representative images are shown. Band intensities were quantified relative to corresponding recombinant protein. The levels of FEN1 and POLβ in HD mouse striatum (St) and cerebellum (Cb) were previously determined,[33] thereby allowing determination of the XRCC1:LIG1:LIG3:APE1:PCNA:FEN1:POLβ molar ratios in the two tissues. (B) XRCC1:LIG1:LIG3:APE1:PCNA:FEN1:POLβ stoichiometry in the striatum and cerebellum of HD mice. The level of POLβ, which is similar in the two tissues of HD mice,[33] was set as 1, and the molar ratio of the other BER proteins was calculated relative to POLβ.

dx.doi.org/10.1021/bi300410d | *Biochemistry* 2012, 51, 3919–3932

	nt number	Sequence 5′-3′*
Target		
D1	100	5′CGAGTCATCTAGCATCCGTAUAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG CAGCAGCAGCAGCAGCAGTACGTAGACTTACTCATTGC 3′
M1	100	5′CGAGTCATCTAGCATCCGTAUTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG CTGCTGCTGCTGCTGCTGTACGTAGACTTACTCATTGC 3′
M2	100	5′CGAGTCATCTAGCATCCGTACAUCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG CAGCAGCAGCAGCAGCAGTACGTAGACTTACTCATTGC 3′
M3	100	5′CGAGTCATCTAGCATCCGTACAGCAGCAGCAGCAGCAGCAGCAGCAUCAGCAGCAG CAGCAGCAGCAGCAGCAGTACGTAGACTTACTCATTGC 3′
M4	100	5′CGAGTCATCTAGCATCCGTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG CAGCAGCAGCAGCAGCAUTACGTAGACTTACTCATTGC 3′
M5	100	5′CGAGTCATCTAGCATCCGTACTUCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG CTGCTGCTGCTGCTGCTGTACGTAGACTTACTCATTGC 3′
M6	100	5′CGAGTCATCTAGCATCCGTACTGCTGCTGCTGCTGCTGCTGCTUCTGCTGCTGCTG CTGCTGCTGCTGCTGCTGTACGTAGACTTACTCATTGC 3′
M7	100	5′CGAGTCATCTAGCATCCGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG CTGCTGCTGCTGCTGCTUTACGTAGACTTACTCATTGC 3′
U14	33	5′CGTGACCTAATGCUCCTTTAAGTCTGAAGGCTG 3′
U16	33	5′CGTGACCTAATGCGCUTTTAAGTCTGAAGGCTG 3′
Template		
D1c/M2c	100	5′GCAATGAGTAAGTCTACGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT GCTGCTGCTGCTGCTGCTGTACGGATGCTAGATGACTCG 3′
M1c/M5c	100	5′GCAATGAGTAAGTCTACGTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA GCAGCAGCAGCAGCAGCAGTACGGATGCTAGATGACTCG 3′
G16	33	5′CAGCCTTCAGACTTAAAGGCGCATTAGGTCACG 3′

Figure 2. Sequences of oligonucleotide substrates. The names, sizes, and sequences of the target oligonucleotides are shown, as well as the location of the AP site within the sequences (designated as U, which is the uracil residue excised by UNG treatment). The complementary (template) oligonucleotides used for double-stranded DNA preparation are also listed.

and data not shown). The amount of each BER protein (see above) was measured relative to a known quantity of the corresponding purified recombinant protein, thereby permitting determination of the molar ratio of BER proteins in mouse striatum and cerebellum (Figure 1B). BER protein levels were globally higher in the cerebellum than in the striatum of HD mice (Figure 1), with wild-type and HD mice exhibiting similar tissue-specific protein ratios (data not shown). Levels of LIG1, FEN1, and APE1, which showed the greatest difference in expression between tissues, were 5-, 4-, and 3-fold more elevated in the cerebellum than in the striatum, respectively (Figure 1B and ref 33). Levels of XRCC1 and PCNA were ∼2-fold higher in the cerebellum, while levels of POLβ and LIG3 were similar in the two tissues (Figure 1 and ref 33). Thus, in agreement with our previous study showing that the activities related to several BER proteins are different in the striatum and cerebellum,[33] we show herein that the relative molar concentrations of key BER proteins vary between the striatum and cerebellum.

The Tissue-Specific BER Stoichiometry Affects the Repair of AP Site-Containing CAG Repeat Substrates. Tissue-specific differences in BER protein stoichiometries could lead to different DNA repair outcomes on CAG repeat substrates. To test this hypothesis, we used a CAG-containing oligonucleotide duplex (D1D1c) and, as a control, a random sequence oligonucleotide duplex (U16G16), which were both treated with the monofunctional DNA glycosylase UNG to excise a strategically positioned uracil residue and generate an AP site (Figure 2; see Materials and Methods). We introduced a natural AP site, as opposed to the synthetic abasic site analogue THF, to allow for processing by either SN-BER or LP-BER. These AP-DNAs were then incubated with a mixture of recombinant BER proteins, which included APE1, POLβ, FEN1, PCNA, the XRCC1−LIG3α complex, and LIG1, at either the cerebellum- or striatum-specific stoichiometry. Repair was monitored by incorporation of the [α-^{32}P]dCTP radionucleotide over time (see Materials and Methods), unless otherwise indicated.

Repair of the CAG (D1D1c) substrate yielded full-length repair products, as well as BER intermediate products that involved incorporation of one or several (up to nine) nucleotides (Figure 3A). The production of full-length repair product, considered herein a measurement of complete repair efficiency, was quantified over time in the presence or absence of ATP (Figure 3B, top panel). Because a significant fraction of the protein molecules in the purified preparations of LIG1 and LIG3 were preadenylated, full-length repair products were detected in a somewhat time-dependent manner without supplemented ATP. Interestingly, in the absence of added ATP, the level of production of full-length repair product was lowest at the striatal BER protein ratios and was inversely correlated with the level of accumulated intermediate products (Figure 3A). A similar effect was observed for both D1D1c and the control DNA U16G16 (Figure S2 of the Supporting Information), indicating that the striatal BER protein stoichiometry results in less efficient repair, as shown by the

dx.doi.org/10.1021/bi300410d | Biochemistry 2012, 51, 3919−3932

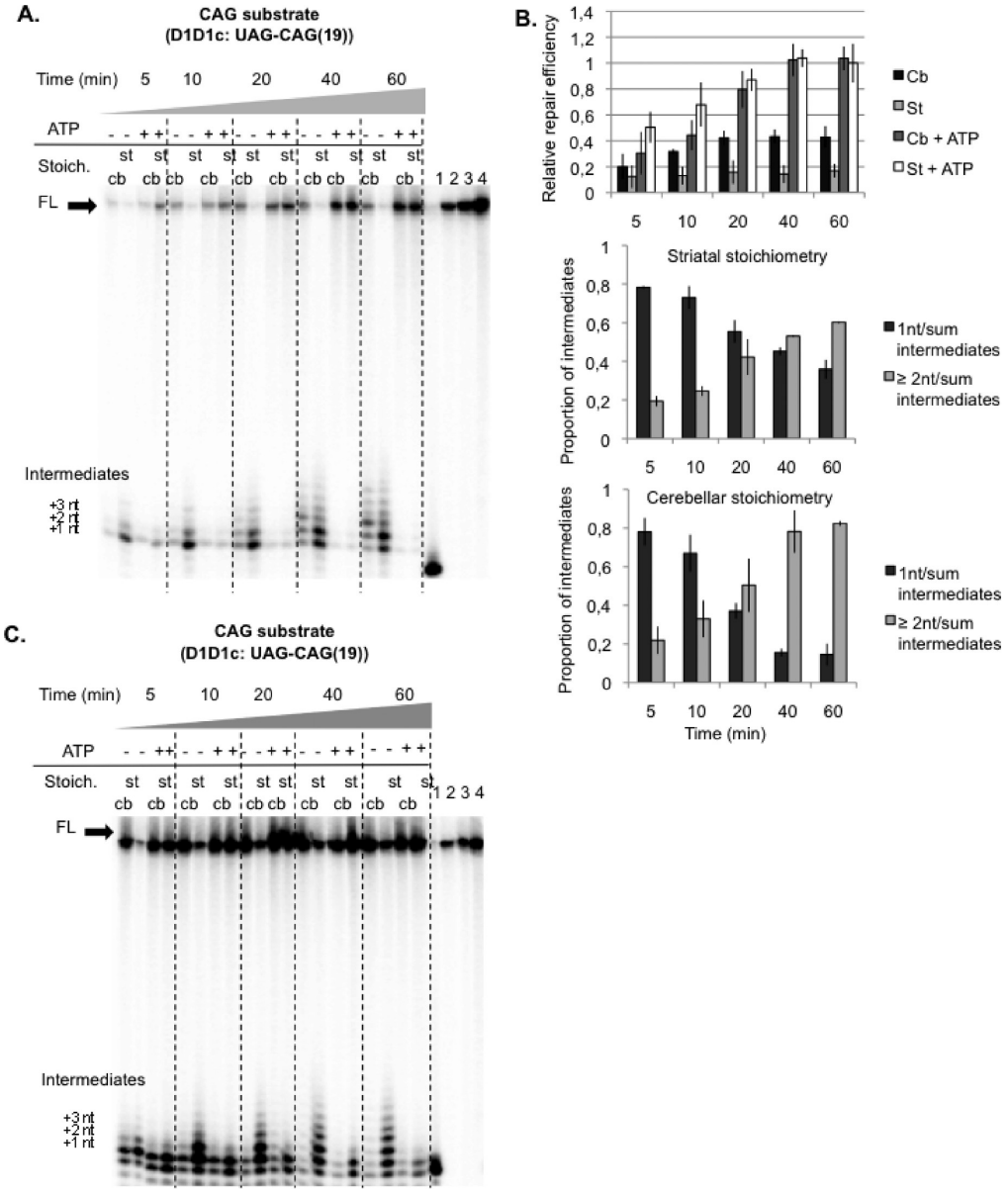


Figure 3. Repair of the CAG substrate is modulated by BER stoichiometry. (A) Radioincorporation experiment showing the time course of repair of the CAG substrate D1D1c under striatal (st) and cerebellar (cb) repair conditions, with or without 1 mM ATP supplementation. A representative polyacrylamide denaturing gel of the repair reaction is shown. The full-length repaired products (FL) are indicated with bold arrows (→), and +1-, +2-, and +3-nucleotide intermediate products are shown. Control reactions included the 5′-labeled U-containing oligonucleotide substrate alone (lane 4) or the substrate treated with the cerebellar BER protein ratio in the absence of UNG (lane 3), with the striatal BER protein ratio in the absence of UNG (lane 2), or with UNG and APE1 only (lane 1). (B) Graph representing repair efficiencies of the CAG D1D1c susbstrate at the striatal or cerebellar BER protein stoichiometries (top). Relative repair efficiencies correspond to the ratios (full-length repair product at a specific time point and BER stoichiometry)/(full-length repair product at 60 min at the striatal stoichiometry + ATP). Two independent experiments were quantified. Error bars are standard deviations. Graphs representing the relative proportion of 1- and ≥2-nucleotide intermediate products at the striatal (middle) or cerebellar (bottom) stoichiometries. (C) Reconstitution repair assay using the 5′-labeled CAG substrate. Time course of repair of the CAG substrate D1D1c at cerebellar or striatal stoichiometries using the 5′-radiolabeled oligonucleotide substrate. The full-length substrates and repair products are indicated with bold arrows (→). For control reactions, the substrate was incubated alone (lane 4), with BER proteins but without UNG (lanes 2 and 3), or with UNG and APE1 only (lane 1).

large number of intermediates and low level of full-length repair product, irrespective of the DNA substrate. Supplementation with ATP, which allows the DNA ligases to catalyze more than one ligation event, substantially increased the repair efficiency over time at both tissue-specific BER stoichiometries and for both substrates, consistent with ligation being a major determinant of repair progression (Figure 3A and Figure S2 of the Supporting Information).

Repair of the CAG substrate (D1D1c) at the striatal or cerebellar stoichiometry mainly involved multinucleotide synthesis, particularly in the absence of ATP supplementation, as revealed by the presence of ≥2-nucleotide intermediate products, indicating that the AP lesion in this DNA context was preferentially processed by LP-BER (Figure 3A). While multinucleotide synthesis occurred at both stoichiometries, the proportion of ≥2-nucleotide intermediate products relative to

dx.doi.org/10.1021/bi300410d | *Biochemistry* 2012, 51, 3919−3932

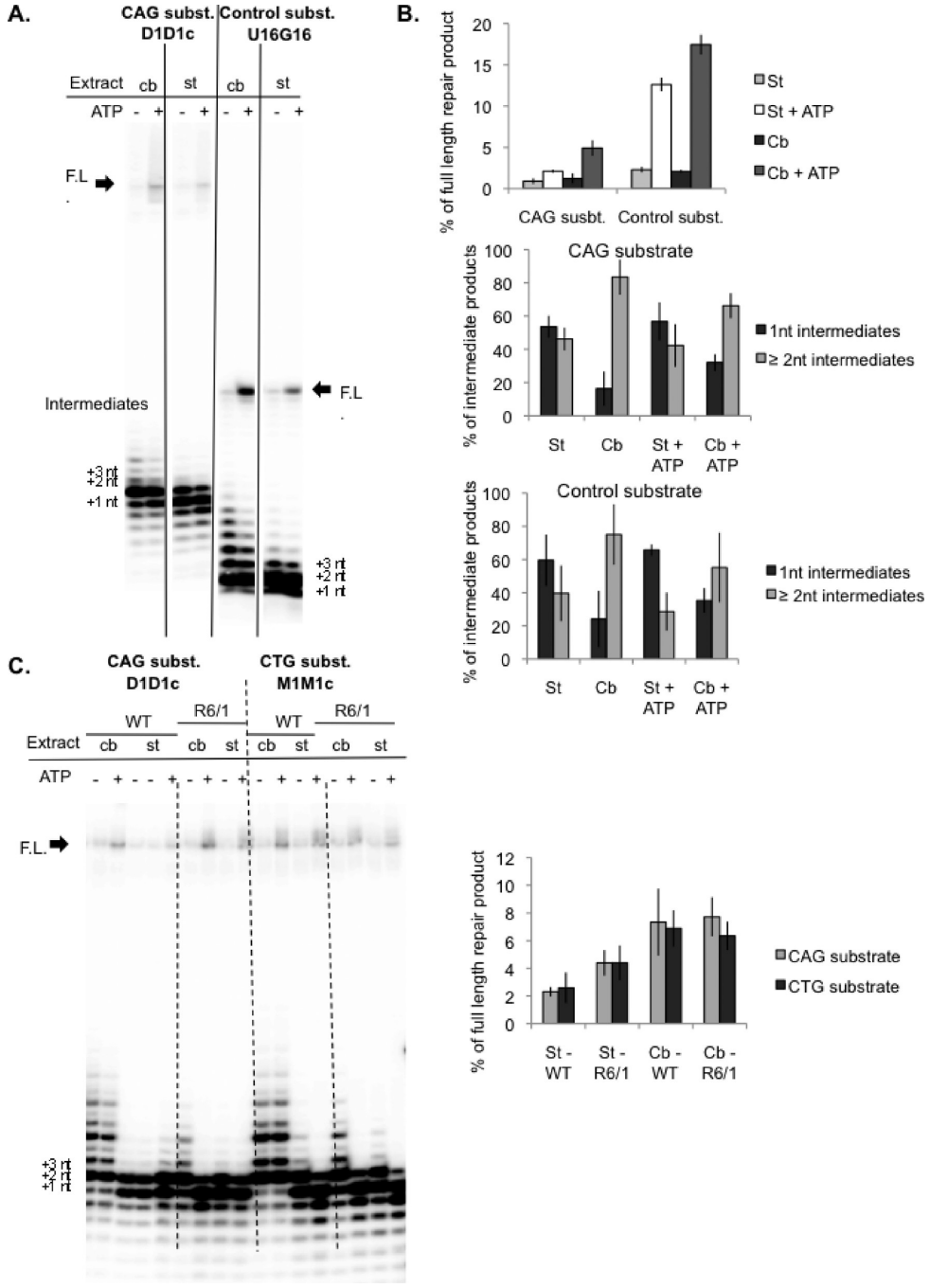


Figure 4. Repair efficiency of DNA substrates is tissue- and substrate-dependent. (A) Tissue-based repair assay using the 5′-labeled CAG D1D1c substrate or the control U16G16 substrate, with or without 1 mM ATP. Cerebellar (cb) or striatal (st) protein extracts prepared from wild-type mice were used. The full-length repaired products (FL) are indicated with bold arrows (→); +1-, +2-, and +3-nucleotide intermediate products are shown. (B) Graph representing repair efficiencies of the CAG or control susbstrates using the striatal or cerebellar protein extracts (top). Repair efficiencies correspond to the percentage of full-length repair products. Error bars are standard deviations. Graphs representing the percentage of +1 and ≥2-nucleotide intermediate products that are generated using striatal (middle) or cerebellar (bottom) protein extracts. (C) Tissue-based repair assay using the 5′-labeled CAG D1D1c or CTG M1M1c substrates with or without 1 mM ATP, using protein extracts prepared from the cerebellum (cb) or striatum (st) of wild-type (WT) or HD (R6/1) mice (left). Graph representing repair efficiencies of the CAG or CTG substrates using the striatal or cerebellar protein extracts (right). Repair efficiencies correspond to the percentage of full-length repair products. Error bars are standard deviations.

+1-nucleotide intermediate products increased more noticeably with time at the cerebellar BER stoichiometry than at the striatal stoichiometry (Figure 3B, middle and bottom panels). The fact that there was also an increased level of production of the full-length repair product at the cerebellar stoichiometry implies that TNR-containing DNA substrates can be efficiently processed by effective LP-BER, which likely exists in the cerebellum because of the higher levels of the LP-BER enzymes FEN1 and LIG1 (Figure 1). Notably, repair of the AP lesion in the control random sequence (U16G16) led to the synthesis of predominantly +1-nucleotide intermediate products at both BER protein stoichiometries, with or without ATP, suggesting

dx.doi.org/10.1021/bi300410d | *Biochemistry* 2012, 51, 3919−3932

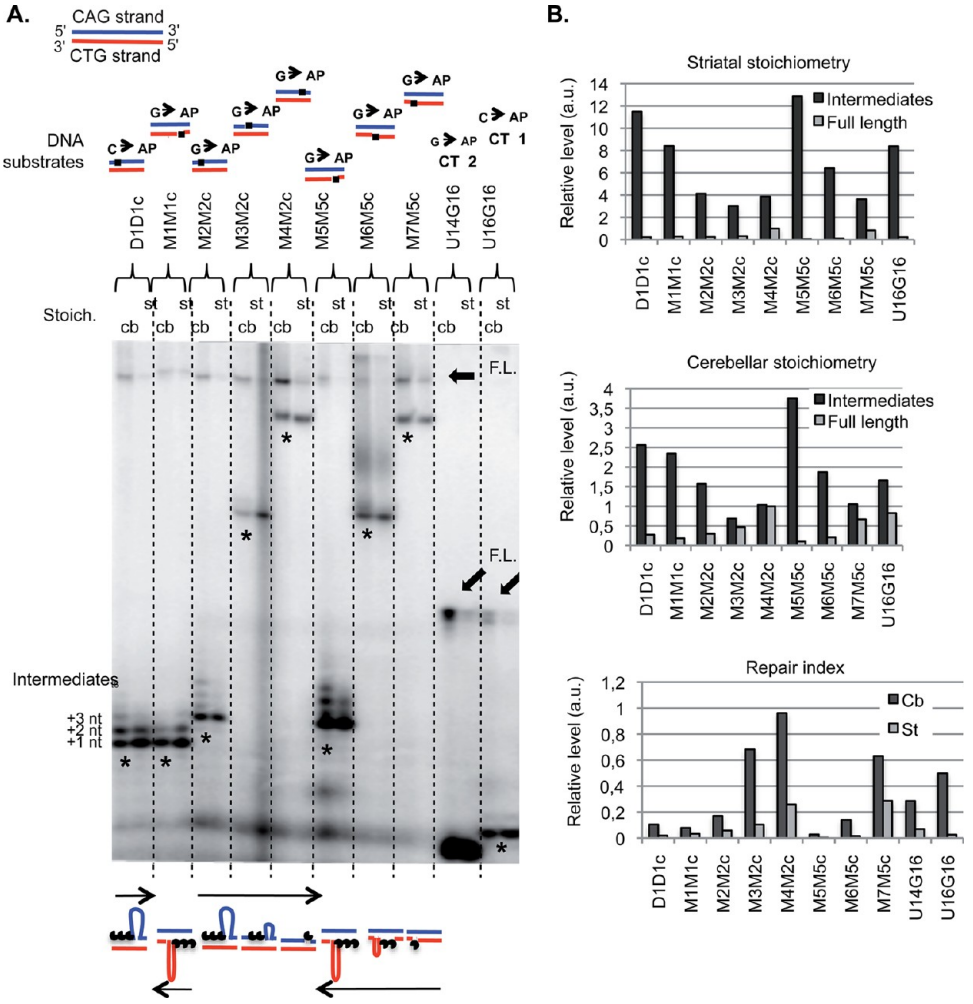


Figure 5. Position of the lesion within the CAG/CTG repeat stretch modulates repair using the reconstitution repair assay. (A) Radioincorporation experiment showing repair of CAG or CTG substrates with an AP site located at various positions within the repeat sequence (top). The scheme represents the substrates before BER. The CAG and CTG strands are colored blue and red, respectively. The dark square represents the AP lesion, and the C or G indicates that the AP lesion replaces a cytosine or guanine, respectively. Control substrates are denoted CT1 and CT2. Reactions were performed at striatal (st) or cerebellar (cb) BER protein stoichiometries for 40 min without ATP supplementation. Bold arrows (→) indicate the full-length repaired product; asterisks indicate the intermediate products. Schematic representation of the different substrates during BER (bottom). The CAG and CTG strands are colored blue and red, respectively. The dark circles represent DNA synthesis by POLβ, and the arrows show the directionality of synthesis. Long CAG or CTG tracts form more stable hairpins than shorter tracts; CTG tracts form more stable hairpins than CAG tracts. (B) Graph representing the relative levels of intermediate products (black bars) and full-length repaired (gray bars) products at the striatal BER stoichiometry (top). For a given substrate, the level of intermediate products corresponds to the sum of all the intermediate products (i.e., +1-, +2-, and +3-nucleotide products, when detected). The level of the full-length repaired products corresponding to the M4M2c substrate at the striatal BER stoichiometry was arbitrarily set to 1. The levels of the full-length and intermediate products for all substrates were calculated relative to that of the M4M2c full-length product. Graph representing the relative levels of intermediate [+1, +2, and +3 nucleotides (black bars)] and full-length repaired (gray bars) products under the cerebellar BER stoichiometry (middle). The level of the full-length repaired products measured for the M4M2c substrate at the cerebellar stoichiometry was arbitrarily set to 1. Graph representing the repair index for all the substrates (bottom). The repair index corresponds to the ratio (full-length repair product)/(sum of all the intermediate products). Black bars show the repair index for the cerebellar stoichiometry, and gray bars show the repair index for the striatal stoichiometry.

that SN-BER is involved in the processing of nonrepetitive sequences (Figure S2 of the Supporting Information).

Repair of 5′-^{32}P-labeled CAG D1D1c substrates confirmed the radioincorporation experiments described above, showing that repair was less efficient at the striatal BER protein stoichiometry (Figure 3C). In particular, in the absence of ATP supplementation, more intermediate products, involving multinucleotide synthesis, and fewer full-length repair products were generated with the CAG substrate using striatal BER protein levels as compared to the cerebellar conditions. Taken together, the results indicate that the stoichiometry of BER proteins influences both repair efficiency and subpathway usage. The results also suggest that the DNA sequence affects BER subpathway selection, with CAG repeats and nonrepetitive sequences following preferentially a LP-BER and a SN-BER path, respectively.

Tissue-Specific Protein Extracts Affect the Repair of AP Site-Containing CAG Repeat Substrates. We next investigated repair of 5′-^{32}P-labeled CAG (D1D1c) and control (U16G16) substrates, after incubation with protein extracts from the striatum and cerebellum of wild-type or HD transgenic mice, with or without supplemented ATP (Figure 4). Supplementation with ATP led to the formation of full-length repair products with both substrates and extracts (Figure

dx.doi.org/10.1021/bi300410d | *Biochemistry* 2012, 51, 3919−3932

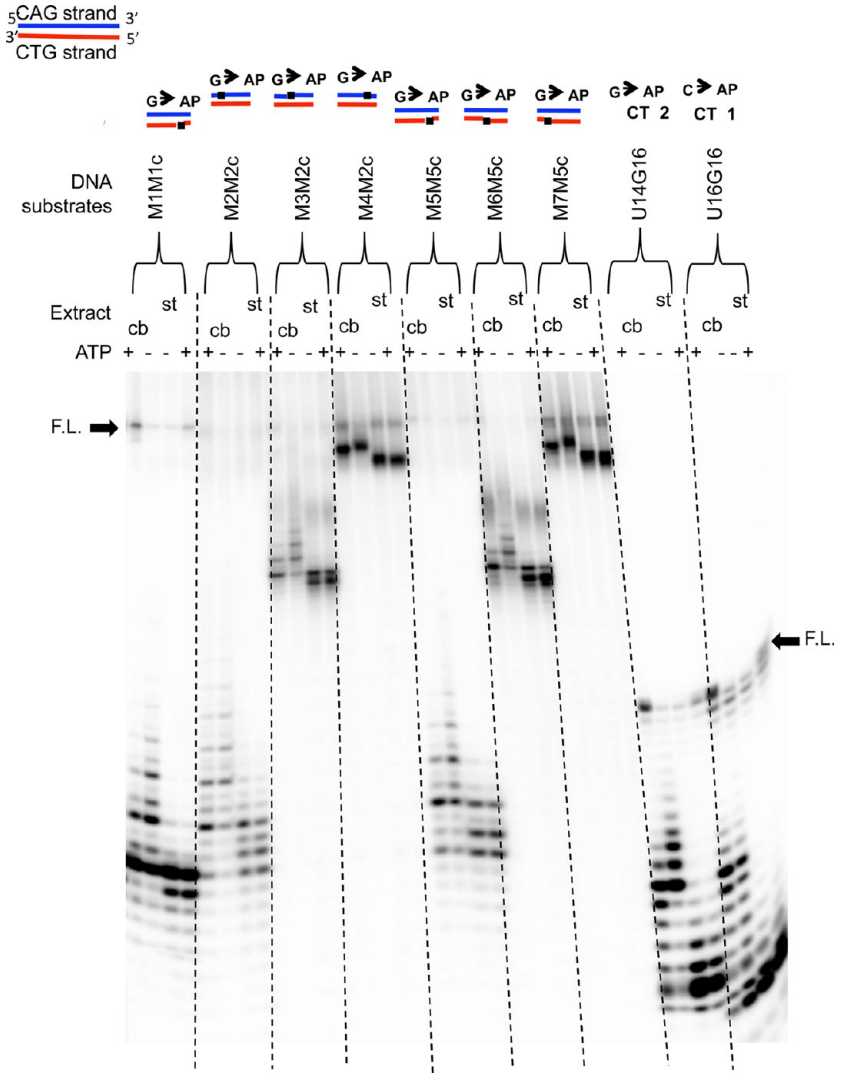


Figure 6. Position of the lesion within the CAG/CTG repeat stretch that modulates repair using tissue-based repair assays. Repair of the various 5′-labeled CAG or CTG substrates with or without 1 mM ATP and using cerebellar (cb) or striatal (st) protein extracts prepared from wild-type mice. The scheme represents the substrates. The CAG and CTG strands are colored blue and red, respectively. The dark square represents the AP lesion, and C or G indicates that the AP lesion replaces a cytosine or guanine, respectively. CT1 and CT2 correspond to the control substrates. Bold arrows (→) indicate the full-length (FL) repaired product.

4A). In agreement with the BER protein reconstitution experiments described above (Figure 3), quantification of the full-length products showed that repair was more efficient (by ≈2-fold) using the cerebellar protein extracts as compared to the striatal extracts, regardless of the substrate (Figure 4B, top panel). The data also reveal that the CAG substrate was ≈5-fold less efficiently repaired than the control substrate, irrespective of the protein extract, indicating that CAG repeats are poor substrates (Figure 4B, top panel). Repair efficiencies on TNR substrates were similar for protein extracts from wild-type or HD transgenic animals (Figure 4C), consistent with our previous finding that the BER protein levels and activities are similar between the comparable tissues of HD and wild-type animals.[33]

Interestingly, intermediate repair products involving synthesis of one or more nucleotides were produced, regardless of the substrate or extract (Figure 4A). In agreement with the fully reconstituted repair assays, the amount of ≥2-nucleotide intermediate products relative to +1-nucleotide products was higher when using the cerebellar extract than when using the striatal extract, further indicating that LP-BER is more active in the cerebellum than in the striatum and is also likely responsible for the more efficient generation of full-length repaired product (Figure 4B, middle and bottom panels). Moreover, processing of the CAG substrate resulted in an increased proportion of ≥2-nucleotide intermediate products relative to +1-nucleotide products when compared to the control substrate, supporting the idea that the CAG sequence is preferentially processed via LP-BER.

Position of the AP Site within CAG or CTG Substrate Influences Repair Outcome. Given the data presented above, we conclude that the CAG substrate is processed primarily by LP-BER, suggesting that the propensity for forming hairpin structures influences subpathway choice. The stability of the hairpin structures formed at CAG/CTG repeats is greater on the CTG strand than on the CAG strand and increases with repeat length.[25,26,36,38] We reasoned that changing the location of the lesion within either a CTG or CAG repeat strand would influence repair outcome. In addition, the ability of the repeat sequence to form a hairpin

dx.doi.org/10.1021/bi300410d | *Biochemistry* 2012, 51, 3919−3932

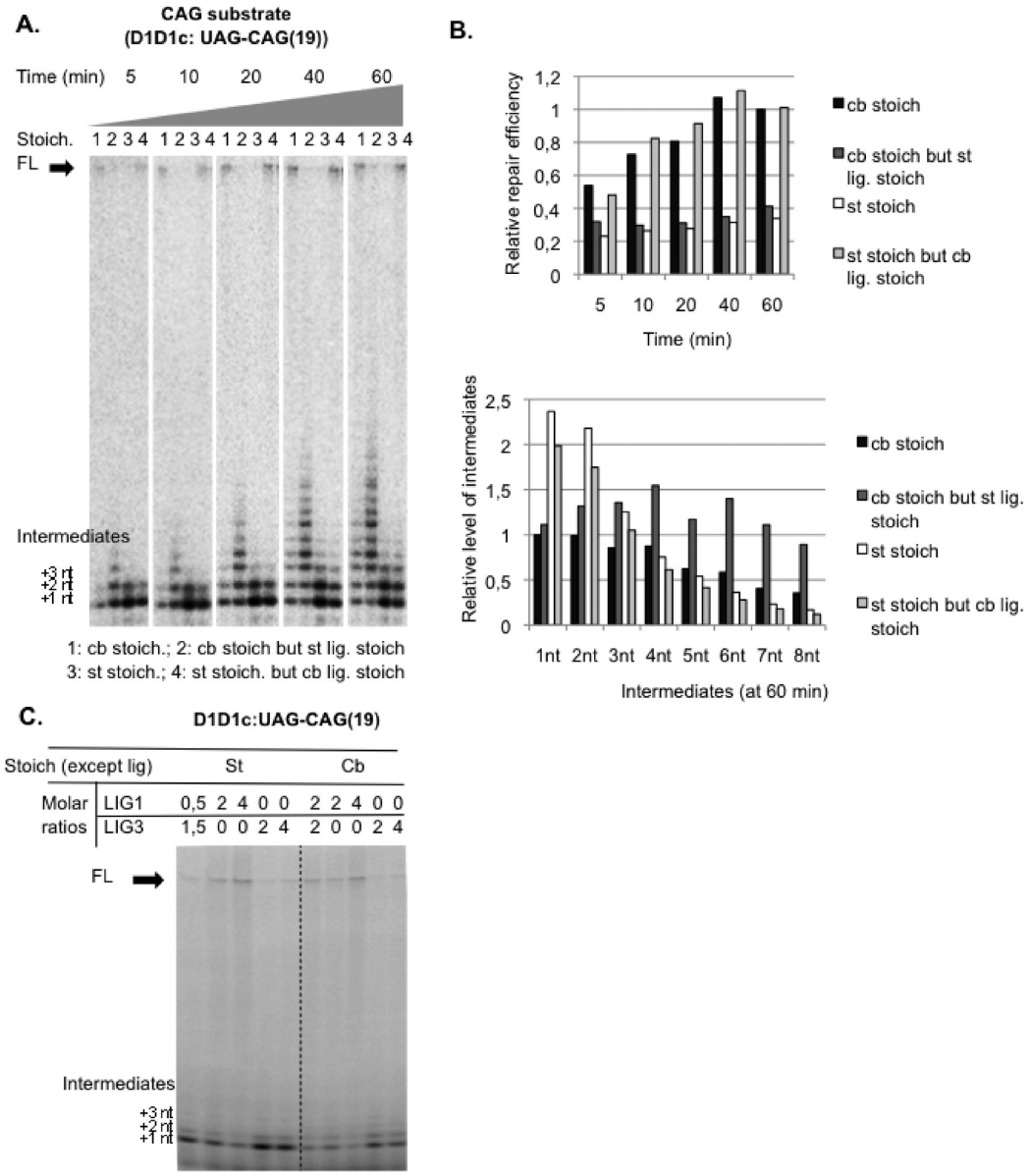


Figure 7. LIG1 concentration is critical in modulating repair of the CAG substrate. (A) Radioincorporation experiments showing the time course of repair of the CAG substrate D1D1c when the stoichiometries of LIG1 and LIG3 are being changed. The substrate was incubated with a mixture of BER proteins reflecting the stoichiometry in the cerebellum of HD mice, except that the concentrations of LIG1 and LIG3 were equal to that measured in the striatum (lanes 2), or, conversely, with a mixture of BER proteins reflecting the stoichiometry in the striatum, except that the concentrations of LIG1 and LIG3 were equal to that found in the cerebellum (lanes 4). As controls, the substrate was also incubated with mixtures of BER proteins reflecting the levels in the cerebellum (lanes 1) and striatum (lanes 3). (B) Graph representing repair efficiencies of the CAG D1D1c susbstrate at the striatal or cerebellar BER protein stoichiometries (top). Relative repair efficiencies correspond to the ratios (full-length repair product at a specific time point and BER stoichiometry)/(full-length repair product at 60 min at the striatal stoichiometry). Graph representing the relative levels of intermediate products resulting from the incorporation of 1−8 nucleotides at 60 min (bottom). The level of +1-nucleotide products at the cerebellar stoichiometry was arbitrarily set to 1. (C) Radioincorporation experiment showing repair of the CAG substrate D1D1c upon removal of either LIG1 or LIG3 from the reaction mixture. The substrates were incubated with a mixture of BER proteins according to the stoichiometry in the striatum or cerebellum, except that the LIG1:LIG3 molar ratio was varied (2:0, 4:0, 0:2, and 0:4). As controls, the striatal ratio (0.5:1.5) and the cerebellar ratio (2:2) were included. Full-length (FL) repair products are indicated with bold arrows (→). The experiments were performed without ATP supplementation.

during repair may be influenced by the starting position of the lesion. To test these hypotheses, we synthesized a series of DNA substrates designed to harbor an AP site that (i) replaces a cytosine (D1D1c and M1M1c) or a guanine (M2M2c and M5M5c) near the 5′ end of a CTG or CAG repeat sequence (5′-oriented), (ii) is embedded within the middle of the repeat stretch (M3M2c and M6M5c), or (iii) is located toward the 3′ end of the repeats (3′-oriented, M4M2c and M7M5c) (Figures 2 and 5A). Two AP site-containing random sequence substrates were used as controls (U16G16 and U14G14).

All DNA substrates tested yielded intermediate and full-length repair products following incubation at cerebellar or striatal BER protein stoichiometries (Figure 5A). However, the relative amount of radiolabeled full-length product and the pattern of intermediate products were different depending on the substrate sequence, the position of the lesion, and the BER

dx.doi.org/10.1021/bi300410d | *Biochemistry* 2012, 51, 3919−3932

protein ratio. The level of full-length repair products relative to intermediate products was higher under the cerebellar conditions for each substrate, as compared to the striatal conditions (Figure 5B, top and middle panels). A "repair index" corresponding to the relative ratio of the full-length product to intermediate products was determined (Figure 5B, bottom panel) and indicates that repair is more efficient for all substrates at the cerebellar BER protein stoichiometry. This finding is consistent with the results depicted in Figure 3.

Placing the AP site within the 3′ portion of the repeat tract led to a greater proportion of full-length repair product, as well as shorter intermediate products, indicative of an efficient SN-BER response, when compared to the lesion placed toward the 5′ end, which led to an accumulation of longer +*n*-nucleotide products (Figure 5A,B). The overall repair outcome for the 3′-oriented substrates (M4M2c and M7M5c) was more similar to that for the control substrates (Figure 5A,B). This effect of position was seen with both the CAG and CTG strand and at both tissue-specific stoichiometries, suggesting more effective repair when the lesion is 3′-oriented than 5′-oriented, likely because of the weakened propensity of the 3′ lesions to form a stable hairpin structure. Finally, although the CAG and CTG substrates exhibited a similar repair pattern when the damage was located in the same position (i.e., 5′, middle, or 3′), substrates with lesions in the CAG strand exhibited a higher repair index than lesions in the CTG strand [compare D1D1c, D2D2c, M2M2c, M3M2c, and M4M2c (CAG substrates) with M1M1c, M5M5c, M6M5c, and M7M5c (CTG substrates) (Figure 5B)].

The conclusions described above were essentially recapitulated by assessing the repair of 5′-^{32}P-labeled AP-DNA substrates upon incubation with protein extracts from the striatum or cerebellum of HD mice, in the presence or absence of supplemented ATP (Figure 6). Repair efficiency increased when the AP site was located 3′ to the repeat end, as compared to the 5′ end or middle position, and this improved efficiency correlated with the production of shorter intermediate products and an increased level of full-length repair product (Figure 6). Addition of ATP promoted full-length product formation. Taken together, our results show that both the strand (CAG vs CTG) and the location of the lesion within the repeat tract affect repair outcome, supporting the view that the propensity to form a hairpin structure influences BER subpathway selection and repair efficiency.

LIG1 Contributes to the Differential Repair Efficiency of Cerebellar and Striatal BER Stoichiometries. A key difference between the BER protein stoichiometry in the mouse striatum and cerebellum is the lower levels of LIG1 in the striatum, while that of LIG3 is similar in the two tissues (Figure 1). In particular, the level of LIG1 is 5-fold lower in the striatum than in the cerebellum. Thus, LIG1 levels may contribute to the differential repair between the striatal and cerebellar BER protein ratios. To examine this hypothesis, we incubated the AP site-containing CAG substrate (D1D1c) with a BER protein cocktail consisting of various ligase protein combinations (Figure 7A). AP-DNA repair was then assessed over time. Repair under the strict cerebellar (cb, lane 1 throughout) or striatal (st, lane 3 throughout) conditions was consistent with the results seen above (Figures 3−5). Interestingly, the repair efficiency as measured by full-length product formation was highest at the cerebellar ligase stoichiometry, regardless of the other BER protein concentrations (Figure 7B, top panel). The presence of intermediate products was generally more pronounced under conditions mimicking the striatal situation for ligases (Figure 7B, bottom panel). Most strikingly, reducing the ligase levels to that of the striatum, where all other BER proteins were at the cerebellar ratios, led to a markedly increased intensity for the >2-nucleotide intermediate products (Figure 7B, bottom panel). Similar results were obtained using the U16G16 control substrate, except that the difference in repair efficiency for the various BER protein cocktails was not as pronounced when the ligase concentration was being changed (Figure S3 of the Supporting Information). In total, the data are consistent with the suspicion that the lower ligase activity found in the striatum contributes to a reduced level of repair (Figure 3).

The LIG1:LIG3 ratio and the total amount of DNA ligase are both increased in the cerebellum in comparison to the striatum (Figure 1). The LIG1:LIG3 ratio is 4 in the cerebellum and 1 in the striatum, and the total amount of ligase is 2-fold higher in the cerebellum than in the striatum. To clarify whether the total amount of ligase protein or the relative ratio of the two ligases determines repair efficacy, we incubated the AP site-containing CAG substrate (D1D1c) with different LIG1:LIG3 molar ratios, while maintaining the other BER proteins at either the cerebellar or striatum stoichiometry (Figure 7C). Repair efficiency was influenced mainly by the LIG1 levels, whereas the amount of LIG3 had a weaker effect on repair outcome. These results indicate that the lower LIG1 level primarily contributes to the reduced repair efficiency seen with the striatal protein stoichiometry.

DISCUSSION

Several studies support a role for BER in disease-associated CAG/CTG instability, yet the underlying mechanisms remain elusive. It has been hypothesized that LP-BER is specifically involved in TNR instability. However, it is unclear whether LP-BER is necessary for processing of a lesion at CAG/CTG repeats and contributes to the tissue selectivity of TNR instability. Our results demonstrate that tissue-specific BER protein stoichiometry and DNA damage location along the repeat tract (5′ vs 3′) influence BER subpathway choice and affect repair efficiency. In repair assays involving either reconstituted protein mixtures or tissue extracts, repair of AP substrates was less efficient at the striatal BER protein stoichiometry, the tissue showing the highest CAG/CTG instability, as compared to the stoichiometry of the cerebellum, a tissue exhibiting minimal TNR instability. This reduced level of repair was observed irrespective of the substrate (e.g., CAG/CTG or control substrate), although repair of the CAG/CTG substrates was less efficient than that of control substrates and primarily involved multinucleotide LP-BER, particularly when the lesion was 5′-oriented within the repeat sequence, presumably because of the increased propensity to form a hairpin structure. Interestingly, the cerebellar BER protein stoichiometry promoted LP-BER, likely because of the high levels of the two LP-BER enzymes FEN1 and LIG1. Together, these results indicate that the inefficient processing of a lesion at CAG/CTG repeats results from inefficient LP-BER activity, as found in the striatum. We suggest that inefficient BER at CAG/CTG repeats, resulting from both the tissue-specific BER protein stoichiometry and the position of the lesion within the repeat tract, contributes to TNR instability.

How BER is regulated at a cell- or tissue-specific level throughout life and how this impacts disease remain largely unknown. Studies have shown that the pattern of expression of

BER genes is regulated throughout development and postnatal life in a tissue-specific manner.[51] Furthermore, tissue- and age-dependent variations of BER protein levels and activities have been reported.[51−54] For example, the levels of different DNA glycosylases, including OGG1, were assessed in different mouse tissues and revealed variations between tissues.[52] In addition, the relative abundance of other BER proteins, such as APE1, POLβ, XRCC1, LIG1, and LIG3, was found to differ between the mouse liver and brain.[54] BER activity was also impaired in terminally differentiated cells[55] and was significantly lower in skeletal muscle than in liver or kidney.[53] We determined the molar ratio of several major BER proteins, including APE1, POLβ, XRCC1, LIG1, LIG3, and PCNA, in mouse striatum and cerebellum and found a significant difference between brain regions, consistent with the idea that BER activity is highly tissue-specific.

Several BER proteins have previously been implicated in affecting the stability of CAG/CTG repeats. For instance, the DNA glycosylase OGG1, an enzyme that initiates BER, is necessary for increasing the somatic instability of CAG/CTG repeats in HD mice,[28] although excision of an 8-oxoG lesion within a CAG hairpin is rate-limiting.[33,39,40] In addition, yeast studies support a role for FEN1 and LIG1 in the instability of CAG/CTG tracts,[3,29,30] and LIG1 is implicated in human cell and transgenic mouse studies.[3] Furthermore, FEN1 cleaves 5′-flap-bearing structures formed by CTG repeats less efficiently than unstructured flaps[56] and is inhibited by the secondary structures formed at these repeats in a length-dependent manner.[57] Reconstitution experiments suggest that coordination between POLβ and FEN1 modulates CAG/CTG repeat expansion during LP-BER.[33,34] These results suggest that lesions at CAG/CTG repeats are poorly processed by BER, likely because of structural impediments, and optimal coordination of the BER enzymatic steps is essential to ensure correct repair and prevent TNR instability.

What the optimal coordination is for BER is less clear. On the one hand, we showed that low levels of FEN1 relative to POLβ correlate with high CAG instability in HD mouse tissues, suggesting that inefficient LP-BER would promote TNR instability.[33,34] On the other hand, using in vitro repair assays, it has been reported that high levels of FEN1 or LP-BER cofactors such as HMGB1 drive CAG instability.[33,34] Our results support the first scenario, showing that repair of CAG/CTG substrates was inefficient when using the BER protein stoichiometry reflecting the situation in the striatum, in comparison with the cerebellar stoichiometry (Figures 3 and 4). The low repair efficiency under the striatal BER stoichiometries correlated with reduced levels of APE1, FEN1, and LIG1 relative to POLβ, as compared to the ratio in the cerebellum, suggesting that poor coordination of DNA synthesis by POLβ with the upstream and/or downstream BER enzymatic steps leads to inefficient repair. In accordance, decreasing the level of DNA ligase, and in particular LIG1, led to less efficient repair of CAG/CTG substrates and a concomitant accumulation of repair intermediates (Figure 7), consistent with previous studies showing that LIG1 controls repair patch length.[3,58,59] In HD and other CAG/CTG repeat-associated diseases, somatic instability is elevated in the striatum and is minimal in the cerebellum, suggesting that CAG/CTG instability is associated with suboptimal tissue-specific BER.

FEN1 and LIG1 have been implicated in CAG/CTG instability, suggesting that LP-BER is specifically required for processing of oxidative DNA damage at CAG/CTG repeats.[3,29,30,33,34] Our data indicate that repair of CAG/CTG substrates is strongly dependent upon LIG1 and not LIG3 (Figure 7). In addition, our results support the view that an AP lesion located in a CAG/CTG tract is preferentially processed by LP-BER, as revealed by the high proportion of ≥2-nucleotide intermediate products, whereas SN-BER is more frequently involved in repairing a lesion within a random DNA sequence (Figures 3 and 4 and Figure S1 of the Supporting Information). LP-BER was executed on CAG/CTG substrates at both the striatal and cerebellar BER protein stoichiometries, though to different extents (Figure 4B). The high levels of FEN1 and LIG1 (relative to that of POLβ) likely contribute to the increased level of involvement of multinucleotide LP-BER in the cerebellum and presumably are necessary for successful execution of complete repair. These results indicate that selection of LP-BER at CAG/CTG repeats is dependent upon both the DNA sequence and the BER protein stoichiometry.

A 5′-oriented AP site resulted in the production of ≥2-nucleotide intermediate products, whereas +1-nucleotide intermediate products predominated when the lesion was 3′-oriented (Figures 5 and 6). This finding suggests that the level of involvement of LP-BER at CAG/CTG repeats increases when the AP site is located 5′ within the CAG or CTG repeat tract, presumably because of the increased propensity to form a downstream secondary structure via strand displacement.[60] In addition, the production of longer +n intermediates was associated with a decrease in repair efficiency, as less full-length repair product was observed with lesions located at the 5′ end of the repeat tract relative to those at the 3′ end (Figures 5 and 6). Interestingly, locating an AP site upstream of or downstream from a CAG/CTG stem−loop structure had dramatic effects upon its ability to form slipped DNAs and to undergo strand exchange.[61] Our observation of an increased level of repair intermediates depending upon lesion position may be due to damage location-dependent alterations in the ability to form slipped DNAs. Thus, it appears that LP-BER operates more frequently when processing a lesion within a repetitive DNA sequence prone to forming a structural impediment, such as a hairpin.

SN-BER plays a prominent role in cells,[62] yet the factors contributing to LP-BER selection in vivo remain poorly defined. Several studies have provided evidence that multinucleotide LP-BER is functional in vertebrate cells, including brain cells where it is catalyzed by POLβ.[49,63,64] It has been reported that reduction or oxidation of AP sites, as well as ATP cellular concentration, controls BER subpathway selection.[48,65] In agreement, our data support the view that a low ATP concentration promotes extended DNA synthesis, likely because of impaired ligation (Figures 3 and 4). In addition, we find that the DNA sequence that surrounds the lesion and the BER protein stoichiometry also influence LP-BER selection. CAG/CTG sequences are poor repair substrates as compared to control random sequences (Figures 4 and 6), suggesting that LP-BER selection at CAG/CTG repeats may act as a backup pathway to facilitate processing of lesions at sequences refractory to repair because of their propensity to form secondary structure. Most notably, the suboptimal striatal LP-BER activity results in poor repair of CAG/CTG substrates and an increased level of formation of incomplete intermediate repair products as compared to the efficient cerebellar LP-BER activity (Figures 3−5). Because CAG/CTG instability is high in the striatum and minimal in the cerebellum, this observation

suggests that LP-BER selection at CAG/CTG repeats is necessary, but suboptimal LP-BER, as found in striatal extracts, is detrimental, resulting in stalled repair progression and the formation of persistent intermediate products. While we do not exclude the possibility that several mechanisms might account for BER-induced TNR instability, we speculate that persistent intermediate products might provide an entry site for MMR, increasing the risk of repeat instability. This situation could explain the moderate reduction in the somatic instability in HD mice deficient for *Ogg1*, in contrast to the severe effect seen in HD and DM1 mice deficient for *Msh2* or *Msh3*.[22,28]

■ ASSOCIATED CONTENT

Supporting Information

Removal of the uracil by UNG (Figure S1), reconstitution repair assay using the U16G16 control substrate (Figure S2), and reconstitution repair assay using the U16G16 control substrate and varying the DNA ligase stoichiometry (Figure S3). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: merienne@igbmc.fr. Phone: +33 3 88 65 34 06.

Funding

This research was supported by the Centre National de la Recherche Scientifique (CNRS), the Institut National de la Santé et de le Recherche Médicale (INSERM), and the University of Strasbourg (K.M.), by a grant from the French Agence Nationale de la Recherche (ANR-2011-JSV6-003-01 to K.M.), by the Intramural Research Program of the National Institutes of Health (NIH), National Institute on Aging (D.M.W.), by the Muscular Dystrophy Association Canada (C.E.P.), the Canadian Institutes of Health Research (MOP-94966 to C.E.P.), and the Paul Wellstone Muscular Dystrophy Cooperative Research Center (C.E.P.), and by grants from the NIH (U54NS48843 to C.E.P. and GM57479 and ES012512 to A.E.T.). A.-V.G. was supported by the French Ministry of Research and the Fondation de la Recherche Medicale (FRM).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank J. L. Mandel and H. Puccio for constant support, C. Weber for technical assistance, and F. Klein for discussions. We are grateful to P. J. McKinnon (St. Jude Children's Research Hospital) for the kind gift of the antibody to XRCC1.

■ REFERENCES

(1) 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.

(2) Tome, S., Panigrahi, G. B., Lopez Castel, A., Foiry, L., Melton, D. W., et al. (2011) Maternal germline-specific effect of DNA ligase I on CTG/CAG instability. *Hum. Mol. Genet. 20*, 2131−2143.

(3) Lopez Castel, A., Tomkinson, A. E., and Pearson, C. E. (2009) CTG/CAG repeat instability is modulated by the levels of human DNA ligase I and its interaction with proliferating cell nuclear antigen: A distinction between replication and slipped-DNA repair. *J. Biol. Chem. 284*, 26631−26645.

(4) Telenius, H., Kremer, B., Goldberg, Y. P., Theilmann, J., Andrew, S. E., et al. (1994) Somatic and gonadal mosaicism of the Huntington disease gene CAG repeat in brain and sperm. *Nat. Genet. 6*, 409−414.

(5) Shelbourne, P. F., Keller-McGandy, C., Bi, W. L., Yoon, S. R., Dubeau, L., et al. (2007) Triplet repeat mutation length gains correlate with cell-type specific vulnerability in Huntington disease brain. *Hum. Mol. Genet. 16*, 1133−1142.

(6) Chong, S. S., McCall, A. E., Cota, J., Subramony, S. H., Orr, H. T., et al. (1995) Gametic and somatic tissue-specific heterogeneity of the expanded SCA1 CAG repeat in spinocerebellar ataxia type 1. *Nat. Genet. 10*, 344−350.

(7) Hashida, H., Goto, J., Suzuki, T., Jeong, S., Masuda, N., et al. (2001) Single cell analysis of CAG repeat in brains of dentatorubral-pallidoluysian atrophy (DRPLA). *J. Neurol. Sci. 190*, 87−93.

(8) Lopes-Cendes, I., Maciel, P., Kish, S., Gaspar, C., Robitaille, Y., et al. (1996) Somatic mosaicism in the central nervous system in spinocerebellar ataxia type 1 and Machado-Joseph disease. *Ann. Neurol. 40*, 199−206.

(9) Kennedy, L., Evans, E., Chen, C. M., Craven, L., Detloff, P. J., et al. (2003) Dramatic tissue-specific mutation length increases are an early molecular event in Huntington disease pathogenesis. *Hum. Mol. Genet. 12*, 3359−3367.

(10) Ishii, S., Nishio, T., Sunohara, N., Yoshihara, T., Takemura, K., et al. (1996) Small increase in triplet repeat length of cerebellum from patients with myotonic dystrophy. *Hum. Genet. 98*, 138−140.

(11) Wong, L. J., Ashizawa, T., Monckton, D. G., Caskey, C. T., and Richards, C. S. (1995) Somatic heterogeneity of the CTG repeat in myotonic dystrophy is age and size dependent. *Am. J. Hum. Genet. 56*, 114−122.

(12) Thornton, C. A., Johnson, K., and Moxley, R. T., III (1994) Myotonic dystrophy patients have larger CTG expansions in skeletal muscle than in leukocytes. *Ann. Neurol. 35*, 104−107.

(13) Anvret, M., Ahlberg, G., Grandell, U., Hedberg, B., Johnson, K., et al. (1993) Larger expansions of the CTG repeat in muscle compared to lymphocytes from patients with myotonic dystrophy. *Hum. Mol. Genet. 2*, 1397−1400.

(14) Lopez Castel, A., Nakamori, M., Tome, S., Chitayat, D., Gourdon, G., et al. (2011) Expanded CTG repeat demarcates a boundary for abnormal CpG methylation in myotonic dystrophy patient tissues. *Hum. Mol. Genet. 20*, 1−15.

(15) Groh, W. J., Groh, M. R., Shen, C., Monckton, D. G., Bodkin, C. L., et al. (2011) Survival and CTG repeat expansion in adults with myotonic dystrophy type 1. *Muscle Nerve 43*, 648−651.

(16) Groh, W. J., Lowe, M. R., and Zipes, D. P. (2002) Severity of cardiac conduction involvement and arrhythmias in myotonic dystrophy type 1 correlates with age and CTG repeat length. *J. Cardiovasc. Electrophysiol. 13*, 444−448.

(17) Swami, M., Hendricks, A. E., Gillis, T., Massood, T., Mysore, J., et al. (2009) Somatic expansion of the Huntington's disease CAG repeat in the brain is associated with an earlier age of disease onset. *Hum. Mol. Genet. 18*, 3039−3047.

(18) Cleary, J. D., Tome, S., Lopez Castel, A., Panigrahi, G. B., Foiry, L., et al. (2010) Tissue- and age-specific DNA replication patterns at the CTG/CAG-expanded human myotonic dystrophy type 1 locus. *Nat. Struct. Mol. Biol. 17*, 1079−1087.

(19) van den Broek, W. J., Nelen, M. R., Wansink, D. G., Coerwinkel, M. M., te Riele, H., et al. (2002) Somatic expansion behaviour of the (CTG)n repeat in myotonic dystrophy knock-in mice is differentially affected by Msh3 and Msh6 mismatch-repair proteins. *Hum. Mol. Genet. 11*, 191−198.

(20) Tome, S., Holt, I., Edelmann, W., Morris, G. E., Munnich, A., et al. (2009) MSH2 ATPase domain mutation affects CTG*CAG repeat instability in transgenic mice. *PLoS Genet. 5*, e1000482.

(21) Savouret, C., Garcia-Cordier, C., Megret, J., te Riele, H., Junien, C., et al. (2004) MSH2-dependent germinal CTG repeat expansions are produced continuously in spermatogonia from DM1 transgenic mice. *Mol. Cell. Biol. 24*, 629−637.

(22) Dragileva, E., Hendricks, A., Teed, A., Gillis, T., Lopez, E. T., et al. (2009) Intergenerational and striatal CAG repeat instability in Huntington's disease knock-in mice involve different DNA repair genes. *Neurobiol. Dis. 33*, 37−47.

(23) 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.
(24) Wheeler, V. C., Lebel, L. A., Vrbanac, V., Teed, A., te Riele, H., et al. (2003) Mismatch repair gene Msh2 modifies the timing of early disease in Hdh(Q111) striatum. *Hum. Mol. Genet. 12*, 273−281.
(25) Pearson, C. E., and Sinden, R. R. (1996) Alternative structures in duplex DNA formed within the trinucleotide repeats of the myotonic dystrophy and fragile X loci. *Biochemistry 35*, 5041−5053.
(26) 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.
(27) Seriola, A., Spits, C., Simard, J. P., Hilven, P., Haentjens, P., et al. (2011) Huntington's and myotonic dystrophy hESCs: Down-regulated trinucleotide repeat instability and mismatch repair machinery expression upon differentiation. *Hum. Mol. Genet. 20*, 176−185.
(28) Kovtun, I. V., Liu, Y., Bjoras, M., Klungland, A., Wilson, S. H., et al. (2007) OGG1 initiates age-dependent CAG trinucleotide expansion in somatic cells. *Nature 447*, 447−452.
(29) Subramanian, J., Vijayakumar, S., Tomkinson, A. E., and Arnheim, N. (2005) Genetic instability induced by overexpression of DNA ligase I in budding yeast. *Genetics 171*, 427−441.
(30) Freudenreich, C. H., Kantrow, S. M., and Zakian, V. A. (1998) Expansion and length-dependent fragility of CTG repeats in yeast. *Science 279*, 853−856.
(31) Robertson, A. B., Klungland, A., Rognes, T., and Leiros, I. (2009) DNA repair in mammalian cells: Base excision repair: The long and short of it. *Cell. Mol. Life Sci. 66*, 981−993.
(32) Fortini, P., and Dogliotti, E. (2007) Base damage and single-strand break repair: Mechanisms and functional significance of short- and long-patch repair subpathways. *DNA Repair 6*, 398−409.
(33) Goula, A. V., Berquist, B. R., Wilson, D. M., III, Wheeler, V. C., Trottier, Y., et al. (2009) Stoichiometry of base excision repair proteins correlates with increased somatic CAG instability in striatum over cerebellum in Huntington's disease transgenic mice. *PLoS Genet. 5*, e1000749.
(34) Liu, Y., Prasad, R., Beard, W. A., Hou, E. W., Horton, J. K., et al. (2009) Coordination between polymerase β and FEN1 can modulate CAG repeat expansion. *J. Biol. Chem. 284*, 28352−28366.
(35) Hou, C., Zhang, T., Tian, L., Huang, J., Gu, L., et al. (2011) The Role of XPG in Processing (CAG)n/(CTG)n DNA Hairpins. *Cell Biosci. 1*, 11.
(36) 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.
(37) Panigrahi, G. B., Slean, M. M., Simard, J. P., Gileadi, O., and Pearson, C. E. (2010) Isolated short CTG/CAG DNA slip-outs are repaired efficiently by hMutSβ, but clustered slip-outs are poorly repaired. *Proc. Natl. Acad. Sci. U.S.A. 107*, 12593−12598.
(38) Panigrahi, G. B., Lau, R., Montgomery, S. E., Leonard, M. R., and Pearson, C. E. (2005) Slipped (CTG)*(CAG) repeats can be correctly repaired, escape repair or undergo error-prone repair. *Nat. Struct. Mol. Biol. 12*, 654−662.
(39) Jarem, D. A., Wilson, N. R., and Delaney, S (2009) Structure-dependent DNA damage and repair in a trinucleotide repeat sequence. *Biochemistry 48*, 6655−6663.
(40) Jarem, D. A., Wilson, N. R., Schermerhorn, K. M., and Delaney, S. (2011) Incidence and persistence of 8-oxo-7,8-dihydroguanine within a hairpin intermediate exacerbates a toxic oxidation cycle associated with trinucleotide repeat expansion. *DNA Repair 10*, 887−896.
(41) Lee, B. I., and Wilson, D. M., III (1999) The RAD2 domain of human exonuclease 1 exhibits 5′ to 3′ exonuclease and flap structure-specific endonuclease activities. *J. Biol. Chem. 274*, 37763−37769.
(42) Nguyen, L. H., Barsky, D., Erzberger, J. P., and Wilson, D. M., III (2000) Mapping the protein-DNA interface and the metal-binding site of the major human apurinic/apyrimidinic endonuclease. *J. Mol. Biol. 298*, 447−459.
(43) Erzberger, J. P., Barsky, D., Scharer, O. D., Colvin, M. E., and Wilson, D. M., III (1998) Elements in abasic site recognition by the major human and *Escherichia coli* apurinic/apyrimidinic endonucleases. *Nucleic Acids Res. 26*, 2771−2778.
(44) Chen, X., Pascal, J., Vijayakumar, S., Wilson, G. M., Ellenberger, T., et al. (2006) Human DNA ligases I, III, and IV-purification and new specific assays for these enzymes. *Methods Enzymol. 409*, 39−52.
(45) Della-Maria, J., Zhou, Y., Tsai, M. S., Kuhnlein, J., Carney, J. P., et al. (2011) Human Mre11/Human Rad50/Nbs1 and DNA Ligase IIIα/XRCC1 Protein Complexes Act Together in an Alternative Nonhomologous End Joining Pathway. *J. Biol. Chem. 286*, 33845−33853.
(46) Levin, D. S., Bai, W., Yao, N., O'Donnell, M., and Tomkinson, A. E. (1997) An interaction between DNA ligase I and proliferating cell nuclear antigen: Implications for Okazaki fragment synthesis and joining. *Proc. Natl. Acad. Sci. U.S.A. 94*, 12863−12868.
(47) Mangiarini, L., Sathasivam, K., Seller, M., Cozens, B., Harper, A., et al. (1996) Exon 1 of the *HD* gene with an expanded CAG repeat is sufficient to cause a progressive neurological phenotype in transgenic mice. *Cell 87*, 493−506.
(48) Petermann, E., Ziegler, M., and Oei, S. L. (2003) ATP-dependent selection between single nucleotide and long patch base excision repair. *DNA Repair 2*, 1101−1114.
(49) Wei, W., and Englander, E. W. (2008) DNA polymerase β-catalyzed-PCNA independent long patch base excision repair synthesis: A mechanism for repair of oxidatively damaged DNA ends in post-mitotic brain. *J. Neurochem. 107*, 734−744.
(50) Rao, K. S., Annapurna, V. V., and Raji, N. S. (2001) DNA polymerase-β may be the main player for defective DNA repair in aging rat neurons. *Ann. N.Y. Acad. Sci. 928*, 113−120.
(51) Wilson, D. M., III, and McNeill, D. R. (2007) Base excision repair and the central nervous system. *Neuroscience 145*, 1187−1200.
(52) Karahalil, B., Hogue, B. A., de Souza-Pinto, N. C., and Bohr, V. A. (2002) Base excision repair capacity in mitochondria and nuclei: Tissue-specific variations. *FASEB J. 16*, 1895−1902.
(53) Szczesny, B., Tann, A. W., and Mitra, S. (2010) Age- and tissue-specific changes in mitochondrial and nuclear DNA base excision repair activity in mice: Susceptibility of skeletal muscles to oxidative injury. *Mech. Ageing Dev. 131*, 330−337.
(54) Intano, G. W., Cho, E. J., McMahan, C. A., and Walter, C. A. (2003) Age-related base excision repair activity in mouse brain and liver nuclear extracts. *J. Gerontol., Ser. A 58*, 205−211.
(55) Narciso, L., Fortini, P., Pajalunga, D., Franchitto, A., Liu, P., et al. (2007) Terminally differentiated muscle cells are defective in base excision DNA repair and hypersensitive to oxygen injury. *Proc. Natl. Acad. Sci. U.S.A. 104*, 17010−17015.
(56) Vallur, A. C., and Maizels, N. (2010) Complementary roles for exonuclease 1 and Flap endonuclease 1 in maintenance of triplet repeats. *J. Biol. Chem. 285*, 28514−28519.
(57) Spiro, C., Pelletier, R., Rolfsmeier, M. L., Dixon, M. J., Lahue, R. S., et al. (1999) Inhibition of FEN-1 processing by DNA secondary structure at trinucleotide repeats. *Mol. Cell 4*, 1079−1085.
(58) Pascucci, B., Stucki, M., Jonsson, Z. O., Dogliotti, E., and Hubscher, U. (1999) Long patch base excision repair with purified human proteins. DNA ligase I as patch size mediator for DNA polymerases δ and ε. *J. Biol. Chem. 274*, 33696−33702.
(59) Levin, D. S., McKenna, A. E., Motycka, T. A., Matsumoto, Y., and Tomkinson, A. E. (2000) Interaction between PCNA and DNA ligase I is critical for joining of Okazaki fragments and long-patch base-excision repair. *Curr. Biol. 10*, 919−922.
(60) Hartenstine, M. J., Goodman, M. F., and Petruska, J. (2002) Weak strand displacement activity enables human DNA polymerase β to expand CAG/CTG triplet repeats at strand breaks. *J. Biol. Chem. 277*, 41379−41389.
(61) Volker, J., Plum, G. E., Klump, H. H., and Breslauer, K. J. (2009) DNA repair and DNA triplet repeat expansion: The impact of abasic lesions on triplet repeat DNA energetics. *J. Am. Chem. Soc. 131*, 9354−9360.

(62) Sobol, R. W., Horton, J. K., Kuhn, R., Gu, H., Singhal, R. K., et al. (1996) Requirement of mammalian DNA polymerase-β in base-excision repair. *Nature 379*, 183−186.
(63) Asagoshi, K., Liu, Y., Masaoka, A., Lan, L., Prasad, R., et al. (2010) DNA polymerase β-dependent long patch base excision repair in living cells. *DNA Repair 9*, 109−119.
(64) Asagoshi, K., Tano, K., Chastain, P. D., II, Adachi, N., Sonoda, E., et al. (2010) FEN1 functions in long patch base excision repair under conditions of oxidative stress in vertebrate cells. *Mol. Cancer Res. 8*, 204−215.
(65) Klungland, A., and Lindahl, T. (1997) Second pathway for completion of human DNA base excision-repair: Reconstitution with purified proteins and requirement for DNase IV (FEN1). *EMBO J. 16*, 3341−3348.