7856–7863 *Nucleic Acids Research, 2020, Vol. 48, No. 14* *Published online 3 July 2020*
doi: 10.1093/nar/gkaa573

A point mutation in the nuclease domain of MLH3 eliminates repeat expansions in a mouse stem cell model of the Fragile X-related disorders

Bruce E. Hayward[1], Peter J. Steinbach[2] and Karen Usdin [1,*]

[1]Section on Gene Structure and Disease Laboratory of Cell and Molecular Biology National Institute of Diabetes, Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, USA and [2]Center for Molecular Modeling, Center for Information Technology, National Institutes of Health, Bethesda, MD 20892, USA

Received March 16, 2020; Revised June 23, 2020; Editorial Decision June 23, 2020; Accepted July 02, 2020

ABSTRACT

The Fragile X-related disorders (FXDs) are Repeat Expansion Diseases, genetic disorders that result from the expansion of a disease-specific microsatellite. In those Repeat Expansion Disease models where it has been examined, expansion is dependent on functional mismatch repair (MMR) factors, including MutLγ, a heterodimer of MLH1/MLH3, one of the three MutL complexes found in mammals and a minor player in MMR. In contrast, MutLα, a much more abundant MutL complex that is the major contributor to MMR, is either not required for expansion or plays a limited role in expansion in many model systems. How MutLγ acts to generate expansions is unclear given its normal role in protecting against microsatellite instability and while MLH3 does have an associated endonuclease activity, whether that contributes to repeat expansion is uncertain. We show here, using a gene-editing approach, that a point mutation that eliminates the endonuclease activity of MLH3 eliminates expansions in an FXD mouse embryonic stem cell model. This restricts the number of possible models for repeat expansion and supports the idea that MutLγ may be a useful druggable target to reduce somatic expansion in those disorders where it contributes to disease pathology.

INTRODUCTION

To date more than 35 human diseases have been shown to result from an expansion in the size of a disease-specific short tandem array or microsatellite. This group of diseases, known as the Repeat Expansion Diseases (REDs) includes the Fragile X-related disorders (FXDs) that result from expansion of a CGG-repeat containing microsatellite located in exon 1 of the *FMR1* gene (1). While the mechanism of expansion is not fully understood, mismatch repair (MMR) complexes have been shown to be essential for expansion in a number of cell and mouse models of these disorders (2–8). Since functional MMR proteins normally protect against general microsatellite instability, the requirement for these proteins to generate expansions is perplexing. During classical MMR in mammals, MutS complexes, either MutSα, a heterodimer of MSH2/MSH6, or MutSβ (MSH2/MSH3), bind to mismatches in DNA. The MutS proteins then recruit one of three mammalian MutL complexes to process the lesion. In most instances the MutL complex recruited is MutLα (MLH1/PMS2), with the other MutL complexes, MutLβ (MLH1/PMS1) and MutLγ (MLH1/MLH3), playing a much smaller role in MMR (9,10). Of the four MMR complexes that have been tested to date, MutSα, MutSβ, MutLα and MutLγ, only MutSβ and MutLγ are required for expansion (2–8). The relevance of these proteins to human expansions is evidenced by Genome Wide Association Studies (GWAS) that implicate some of these proteins in the extent of somatic instability (11,12), the age at onset (11,13,14) and disease severity (12,15) of a number of REDs. For example, in a recent study of individuals with 40–50 CAG repeats in *HTT*, those with higher blood somatic expansion scores had worse HD outcomes, with the variation in expansion scores being associated with variants in MSH3 (positive false discovery rate (pFDR) = 0·009), MLH1 (pFDR = 0·004) and MLH3 (pFDR = $8{\cdot}0 \times 10^{-4}$) (14). However, understanding the role of these MMR proteins in the expansion process is complicated by the fact that some of these proteins also have roles outside of MMR, including in DNA damage signaling and meiosis, and that genes normally involved in other DNA repair pathways have also been implicated in the expansion process (11,13–19).

It is not clear why MLH3 is necessary for expansion in most REDs model systems that have been tested, while the related protein PMS2 protects against expansion in some cases, is not required for expansions in others, or is only required for a subset of expansions (20–22). It is also un-

*To whom correspondence should be addressed. Tel: +1 301 496 2189; Email: ku@helix.nih.gov

clear as to how MLH3 contributes to the generation of expansions. Several different expansion models have been proposed, some of which are consistent with a nuclease-independent role for the MutL proteins and some of which are not (16,23–27). While relatively little is known about MutLγ, it shares significant similarities with MutLα, a complex that is known to have both nuclease-dependent and nuclease-independent functions (28,29). A potential non-catalytic role for MLH3 is suggested by the fact that there are two splice isoforms of human MLH3, one that includes exon 7 where the endonuclease motif is located (isoform 1) and one that lacks exon 7 but nonetheless retains the normal reading frame (isoform 2) (30).

The potential role of the MLH3 endonuclease activity in repeat expansion was addressed previously in a cell culture model of GAA-repeat expansion in Friedreich ataxia. In this study, the authors reversed the relative levels of isoform 1 and 2 using a splice-switch oligonucleotide (20). Preferential production of isoform 2 resulted in a decrease in the rate of expansion in both the cell culture model system and patient cells, although the effect in patient cells did not reach statistical significance (20). While these data are suggestive of the fact that the MLH3 endonuclease activity is important for expansion, there are two caveats with such an interpretation. Firstly, the loss of 24 amino acids in isoform 2 could have adversely affected protein stability, overall protein conformation or some functionality other than endonuclease activity, all of which may be important contributors to the reduced expansion rate. For example, the MLH3 endonuclease motif overlaps with the C-terminal MLH1-interaction domain in the highly conserved yeast MLH3 homolog (31,32) and while human isoform 2 interacts with MLH1 in yeast 2-hybrid assays (33) and *in vitro* pull-down experiments with GST-tagged protein (34), in yeast even single point mutations in the highly conserved MLH3 endonuclease domain abolish the MLH1 interaction in 2-hybrid assays (31,32) and human isoform 2 does not bind MLH1 in a mammalian 2-hybrid assay (30). Thus, it is possible that the absence of exon 7 results in reduced levels of MutLγ, in addition to the loss of its nuclease activity. Secondly, because isoform 1 was only transiently and partially reduced in the experiments looking at its effect on repeat expansion in FRDA (20), it is unclear whether MLH3 and/or its nuclease activity is essential for all expansion events.

To directly address the role of the endonuclease motif in expansions of the CGG-repeat tract that causes the FXDs, we used embryonic stem cells (mESCs) derived from our Fragile X premutation mouse model (35). We have shown that these stem cells show repeat expansion in culture that, like expansion *in vivo*, is MSH2-dependent (36). We used these mESCs to generate cell lines with a base-edited version of *Mlh3* that has a single nucleotide G-to-A change that converts the aspartic acid (D) at amino acid 1185 in the highly conserved $DQHA(X)_2E(X)_4E$ endonuclease motif to asparagine (N). Yeast with the equivalent mutation are unable to carry out either MMR or meiosis (32) and mice with this mutation are unable to carry out crossing over during meiosis (37) consistent with the loss of MLH3 nuclease activity. However, the overall conformation and stability of the mouse D1185N mutant protein appears to be unchanged and the mutant protein retains its ability to interact normally with MLH1 (37). The equivalent MLH3 mutation in humans and other organisms also does not affect protein stability or the ability to interact with MLH1 (26,38–41). Furthermore, the D1185N mutant protein still facilitates the appropriate localization of MutLγ to the synaptonemal complex during pachynema and the proper loading of factors like CDK2 and HEI10, that are required for proper meiotic crossover resolution (37). This would be consistent with the idea that while the nuclease activity of the D1185N mutant is impaired or abolished, the overall stability and structure of the protein are preserved.

We show here that most, if not all, expansions are lost in mESC lines homozygous for the D1185N mutant protein and that a significant decrease in expansions is seen even in the heterozygous state. This would suggest that the nuclease activity of MLH3 is required for expansion. This has implications for the expansion mechanism. Inhibition of MSH3 has been suggested as a therapeutic approach in the treatment of Repeat Expansion Diseases in which somatic expansion contributes to a high early mortality (12,20). By the same token, modulation of MLH3 endonuclease may also be therapeutically useful, since MutLγ is a minor player in MMR and the MLH3 catalytic site may be a tractable target for drug development (42).

MATERIALS AND METHODS

Reagents and services

All reagents were from Sigma-Aldrich (St. Louis, MO, USA) unless otherwise specified. Primers were from Life Technologies (Grand Island, NY, USA). Capillary electrophoresis of fluorescently labeled PCR genotyping products was carried out by the Roy J Carver Biotechnology Center, University of Illinois (Urbana, IL, USA). The guide RNA expression plasmid MLM3636, a gift from Keith Joung (Addgene plasmid # 43860; http://n2t.net/addgene:43860; RRID:Addgene_43860) and the BE4-Gam base editor plasmid pLenti-BE4GamRA-P2A-Puro, a gift from Lukas Dow (Addgene plasmid # 112673; http://n2t.net/addgene:112673; RRID:Addgene_112673) (43), were obtained from Addgene (Watertown, MA, USA).

Generation of Mlh3 mutant cell lines

Mice were maintained in accordance with the guidelines of the NIDDK Animal Care and Use Committee and with the Guide for the Care and Use of Laboratory Animals (NIH publication no. 85-23, revised 1996). Mouse embryonic stem cells were derived from a mouse model of the Fragile X premutation (35) as described elsewhere (36). Briefly, embryos were isolated from superovulated C57BL/6 $Fmr1^{WT/KI}$ females (35) mated with C57BL/6 $Fmr1^{WT}$ males by flushing the uterine horns with M2 medium (MTI-GlobalStem, Gaithersburg, MD, USA) at post coitum day 3.5. After successive washes with M2, a 1:1 mixture of M2 and KSOM medium (MTI-GlobalStem) and KSOM, embryos were plated in individual 0.1% gelatin (Millipore) coated wells pre-equilibrated with KSOM medium. After hatching from the zona pellucida, N2B27 medium supplemented with 2i (3 μM CHIR99021 (Selleckchem, Hous-

ton, TX, USA), 1 μM PD0325901 (Selleckchem) and LIF (1000 unit/ml; Millipore, Burlington, MA, USA) (N2B27 2i/LIF medium) was used with daily medium changes until the emergence of ES-cell colonies. For routine propagation, cells were maintained on 0.1% gelatin coated wells in N2B27 2i/LIF media with daily medium changes and passaging every 2–3 days. Verification of the pluripotency markers Nanog (Novus Biologicals, Littleton, CO, USA), Oct4 (Santa Cruz Biotech, Dallas, TX, USA) and Sox2 (Millipore) was carried out using standard immunofluorence protocols (36).

A CRISPR–Cas9-based gene editing strategy was used to generate a D1185N point mutation in both *Mlh3* alleles as follows: The complementary oligonucleotides MmMlh3_D1185N-F (5′-ACACCGCTGGTCCACCAGGACTAACG-3′) and MmMlh3_D1185N-R (5′-AAAACGTTAGTCCTGGT GGACCAGCG-3′), specifying the guide RNA sequence (GCUGGUCCACCAGGACUAA), were annealed together, cloned into MLM3636 digested with *Bsm*BI and transfected together with pLenti-BE4GamRA-P2A-Puro into PM mESCs containing either ∼175 or ∼275 CGG repeats in the *Fmr1* 5′ UTR. Individual $Mlh3^{D/N}$ and $Mlh3^{N/N}$ clones were identified by direct sequencing of the PCR product as shown in Supplementary Figure S1B. The size of the repeat tract was then determined as previously described (2) and size matched cell lines chosen for further analysis.

Analysis of repeat expansion

CRISPant cell lines were grown in parallel with size-matched but unedited cell lines as controls. Cells were harvested at regular intervals and the DNA isolated from 200 000–400 000 cells. The repeat number was determined as previously described using high resolution capillary electrophoresis to resolve the PCR products generated from 75 ng of genomic DNA, i.e. 25 000 genome equivalents (2). The resultant fsa files were analyzed using a previously described custom R script (44) that is available on request. The script identifies the size of the highest peak in the PCR profile and this is converted to the mean repeat size by subtracting the number of bases flanking the repeat in the PCR amplicon and dividing the result by three. Comparison of the expansion rates of $Mlh3^{D/D}$ and $Mlh3^{N/N}$ cells was carried out using analysis of covariance (ANCOVA) in the R package. For the ∼184 repeat lines, the data from Figure 3A were used. For the larger repeat lines, the data from the ∼275 repeat $Mlh3^{D/D}$ and the ∼272 repeat $Mlh3^{N/N}$ lines shown in Figure 3B were used together with the data from an additional $Mlh3^{D/D}$ cell line with ∼277 repeats (45).

Molecular modeling

A three-dimensional model was built for residues Leu 1154 to Glu 1374 of mouse MLH3, using the Prime software tools (Schrödinger, LLC) and the crystal structure of the C-terminal domain of the MutL homolog of *Neisseria gonorrhoeae* (3ncv.pdb, (39)) as the template. Of the 221 MLH3 residues modeled, 52 (24%) are identical to the nearest residue in the template structure. The model was rendered using the programs MolScript (46) and Raster3D (47).

RESULTS AND DISCUSSION

CRISPR-Cas9 base-editing of male mESCs carrying a FX PM allele was carried out to introduce a point mutation in the putative endonuclease domain of *Mlh3* as described in the Methods section. This mutation causes the aspartic acid at residue 1185 to be replaced with asparagine (D1185N). The mutant cell lines were sequence verified (Supplementary Figure S1B). Unfortunately, the endogenous levels of MLH3 in most cells are very low (48) and consistent with previous reports (20), our own testing of three different commercially available antibodies did not identify any that were suitable for western blotting of endogenous mouse MLH3. However, it has been shown that while this mutation abolishes the MMR activity and the ability to complete meiosis, MLH1–MLH3–D1185N complexes and MLH1–MLH3 complexes form heterodimers with similar stabilities and chromatographic properties (37). The MLH1–MLH3–D1185N complex also localizes on meiotic chromosomes normally and recruits the normal meiotic MutLγ-interacting partners (37). These observations suggest that the D1185N mutation has little, if any, effect on the stability of mouse MLH3 or its interactome. This conclusion is supported by the fact that the equivalent mutation in human MLH3, a protein with 84% identity and 90% similarity in the CTD (Supplementary Figure S2A), also has a stability indistinguishable from the native protein and an apparently normal interaction with MLH1 (26). The same is true of the corresponding mutations in yeast MLH3, in the *Neisseria gonohorroea* MutL protein (38,39) and in human PMS2 (38), a related MLH1-binding partner with sequence and functional similarities to MLH3.

To further assess the effects, if any, of the D1185N mutation on protein structure and stability, we built a homology model of the C-terminal domain of mouse MLH3 using the structure of the *Neisseria* MutL protein (39) as a template. Although the underlying sequence identity is limited, the model strongly suggests that the conserved $DQHA(X)_2E(X)_4E$ motif is positioned near the conserved CXHGRP motif important for Zn^{2+} binding (Supplementary Figure S2B). A similar clustering has been observed in molecular models of PMS2, a related protein that also binds MLH1, that were built using the *Escherichia coli* MutL protein as a template (38). Furthermore, our modeling places D1185 at the very end of a β-strand, on the protein surface, with its side chain extended into the solvent (Supplementary Figure S2B and C), as do the corresponding side chains in the other MutL proteins studied (38). Therefore, it is reasonable to expect that mouse MLH3 can accommodate the conservative replacement of this aspartate with asparagine without significantly compromising protein structure or stability. The equivalent aspartate residues are similarly situated on the surfaces of other related proteins that are not destabilized by the D-to-N mutation, supporting this contention (38). Thus, considerable evidence points to very similar properties for the native and mutant mouse MLH3 proteins except for the catalytic activity, consistent with the findings of Toledo *et al.* (37).

We initially examined expansion of the repeat in two independently derived $Mlh3^{D/D}$ ($Mlh3^{+/+}$, WT) lines with

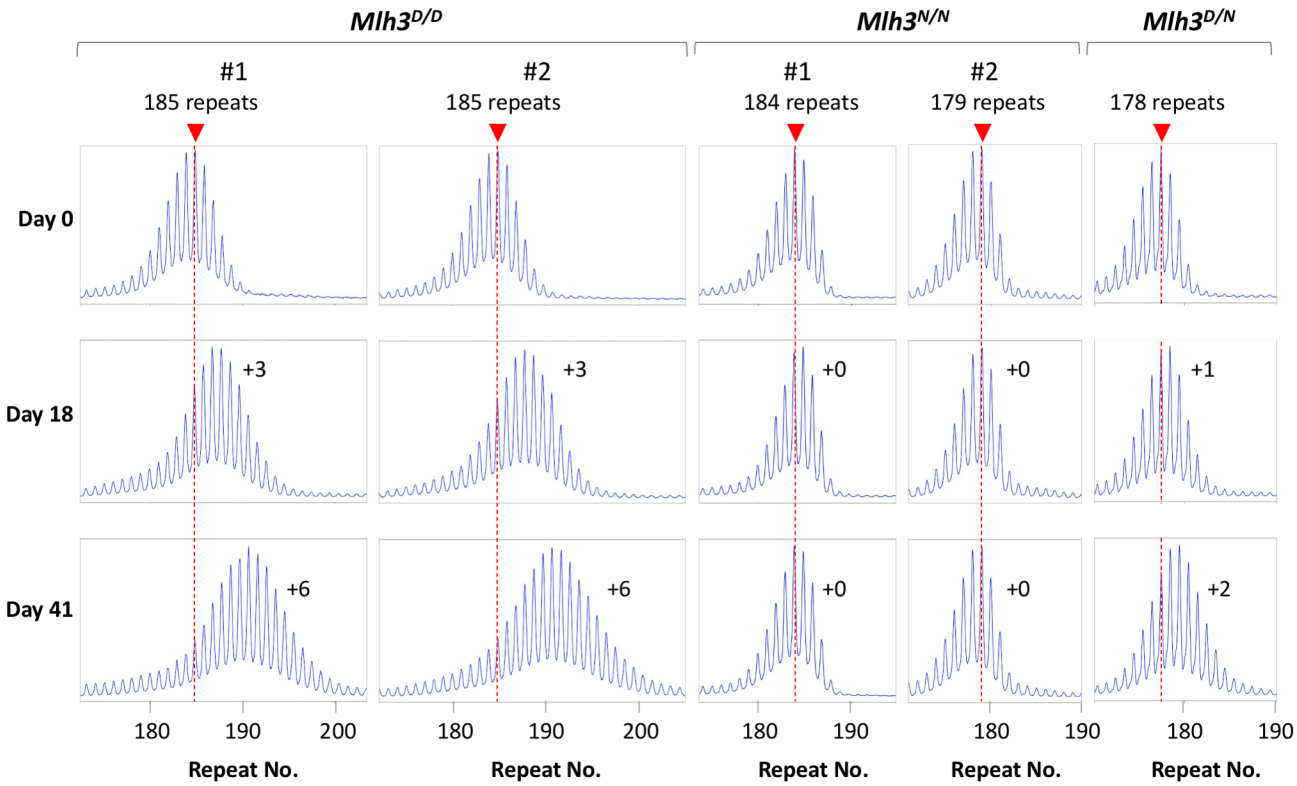

Figure 1. The effect of the point mutation in the nuclease domain on repeat expansion of alleles with ∼180 CGG-repeats. Mouse ESC lines with the indicated number of repeats that were homozygous for the D (WT) *Mlh3* allele ($Mlh3^{D/D}$), homozygous for the N (nuclease-dead) allele ($Mlh3^{N/N}$) or heterozygous for the D and the N allele ($Mlh3^{D/N}$) were propagated for a total of 41 days in culture. The repeat number in the cells in culture was then determined using total genomic DNA isolated from these cultures and compared to the repeat number in the genomic DNA from the original cultures. The red dotted line indicates the position of the original allele in the repeat PCR profile.

∼185 repeats, two independently derived $Mlh3^{N/N}$ lines with ∼179 and 184 repeats and a $Mlh3^{D/N}$ line with 178 repeats. All cells had similar proliferation rates. We have previously shown that during propagation of WT mESCs all cells in a population undergo a high frequency of small expansions with 1–2 repeats being added with each expansion event (36), consistent with what is observed in the FXD mice (49). As a consequence, the average repeat size of the population increases with time as can be seen in the PCR profiles generated from bulk genomic DNA isolated from large populations of cells analyzed at successive time points. In these profiles, the individual peaks reflect the abundance in the population of cells with alleles of the indicated repeat number. As can be seen in Figure 1 both of the $Mlh3^{D/D}$ lines expand at a very similar rate such that after 41 days in culture, the PCR profile produced from bulk genomic DNA from both cell lines shows that the average repeat size of the alleles in both cell populations is now 6 repeats larger than it was in the starting population. In contrast, neither of the $Mlh3^{N/N}$ lines gained any repeats over the 41 days in culture. As can be seen in Figure 3A, this results in expansion rates that are significantly different ($P = 5 \times 10^{-10}$), with the $Mlh3^{D/D}$ lines showing an almost constant expansion rate with most, if not all, cells in the population gaining one repeat every ∼7 days, while the populations of $Mlh3^{N/N}$ cells showed no change in repeat number. We also evaluated a $Mlh3^{D/N}$ line and found that it gained only two repeats over the same period of time, suggesting that even the loss of a single allele with a functional nuclease was enough to significantly reduce the expansion frequency.

To test the effect of the inactivating point mutation more stringently, we wanted to examine the stability of the repeat in cell lines with larger repeat numbers tested over a longer period of time. Since the *Fmr1* gene in these cells is not transcriptionally silenced, as similarly sized human alleles are, $Mlh3^{D/D}$ lines generated from male mouse embryos with ∼250 repeats continue to expand and since the expansion rate is proportional to repeat number, they expand at a faster rate than $Mlh3^{D/D}$ lines with 175 repeats. For example, as can be seen in Figure 2, a $Mlh3^{D/D}$ cell line with 240 repeats showed a gain of 6 repeats at day 31, 8 repeats at day 43 and 12 repeats at day 59. As can be seen in Figure 3B, this corresponds to an almost linear expansion rate, with expansion events occurring in almost every cell in the population every 5 days. Similarly, a $Mlh3^{D/D}$ cell line with 275 repeats showed a gain of 11, 15 and 21 repeats over the same period of time for an average of an expansion event every 3 days. For both cell lines, no alleles containing the original repeat number could be seen above background in the PCR profile at 59 days. Thus, each cell in the population has undergone expansion at least once, with most having undergone multiple expansion events. In contrast, in a $Mlh3^{N/N}$ cell line with 272 repeats that was generated by gene-editing of those mESCs, no gain of repeats was seen in

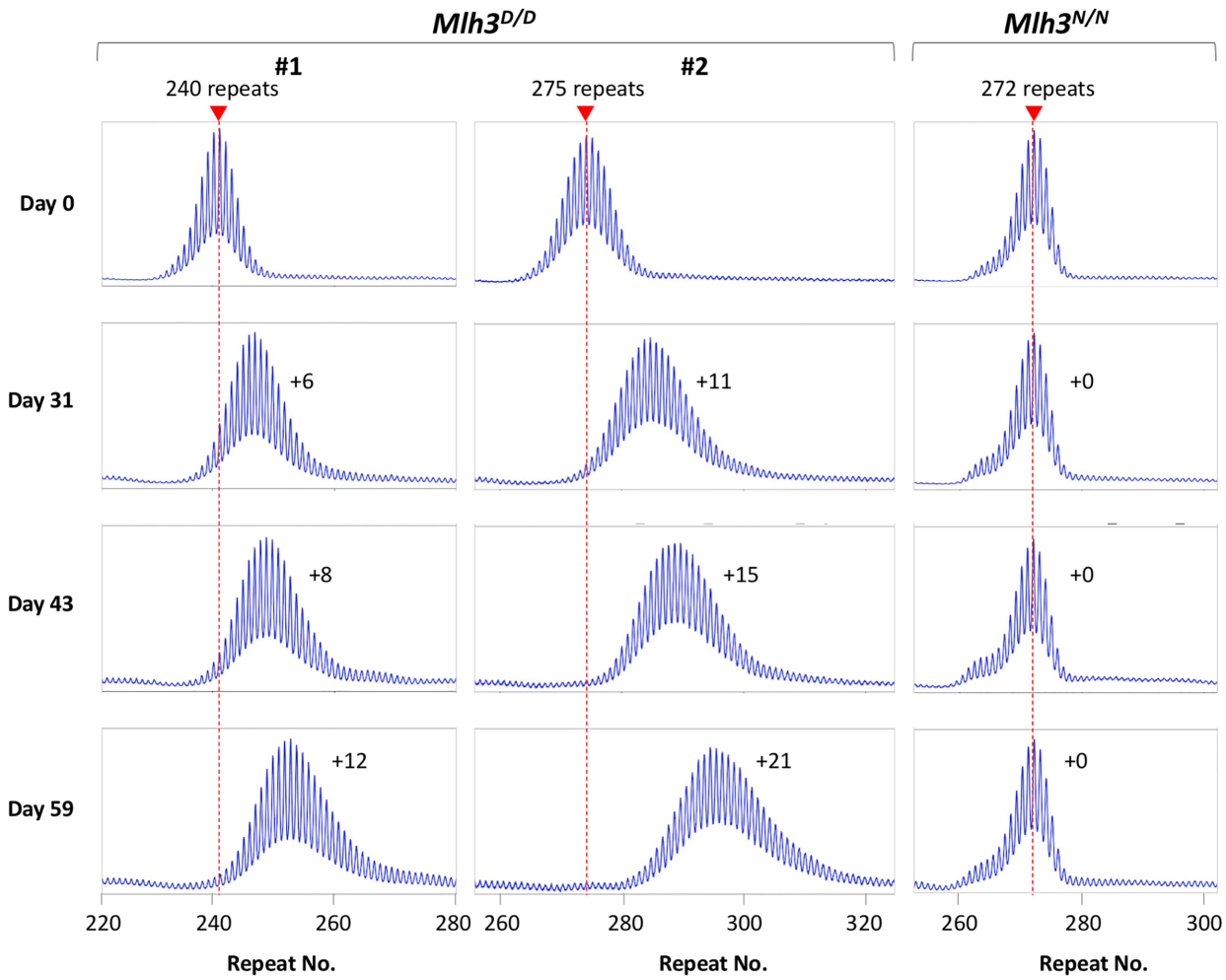

Figure 2. The effect of the point mutation in the nuclease domain on repeat expansion of alleles with ∼250 CGG-repeats. Mouse ESC lines with the indicated repeat numbers that were homozygous for the D (WT) *Mlh3* allele (*Mlh3*$^{D/D}$) or homozygous for the N (nuclease-dead) allele (*Mlh3*$^{N/N}$) were propagated for a total of 59 days in culture. The repeat number in the cells in culture was then determined using total genomic DNA isolated from these cultures and compared to the repeat number in the genomic DNA from the original cultures. The red dotted line indicates the position of the original allele in the repeat PCR profile.

this time period and thus the PCR profiles of the day 0 and day 59 samples are indistinguishable (Figure 2). Thus, the expansion rates of *Mlh3*$^{D/D}$ and *Mlh3*$^{N/N}$ cells with ∼275 repeats, like those with ∼180 repeats are also significantly different ($P = 3.7 \times 10^{-12}$).

Thus, four out of four cell lines with WT *Mlh3* alleles showed the gain of 1–2 repeats every 3–7 days depending on the starting allele size. This is consistent with our previous demonstrations that nine other WT lines with 174–294 repeats also consistently gain multiple repeats over time (36,45). In contrast, three out of three homozygous mutant cell lines showed no detectable expansions at all over an extended time in culture. Furthermore, the single heterozygous line tested gained fewer repeats than comparably sized WT lines, consistent with our previous demonstration that while mice heterozygous for a null allele of *Mlh3* do show expansions, they gain significantly fewer repeats than WT mice (2).

Given the fact that the D1185N homozygous mutant lines lack expansions altogether, one would have to posit that the conservative D-to-N substitution results in a protein that is completely unstable for this to adequately explain the absence of expansions. All of the available data, including that from molecular modeling (this manuscript), along with data from a mouse model (37) and from studies on homologous proteins in humans and other organisms (26,38,39), suggest that this explanation is unlikely. Thus, our data lend support to the idea that the nuclease activity of MLH3 is required for expansion in FX mESCs. This has ramifications for models of repeat expansion. For example, it would preclude those models in which the MMR proteins simply act by stabilizing the expansion substrates for processing by other DNA repair enzymes (16,23,24) or those that invoke loop incorporation initiated by repair of a damaged base by enzymes such as the MUTYH glycosylase on the opposite strand (25). The differences between MutLα and MutLγ cleavage may provide useful insights into what the actual mechanism may be. MutLγ has recently been shown to cleave preferentially on DNA strands opposite loop-outs (26). In principle, this could result in the loop-out

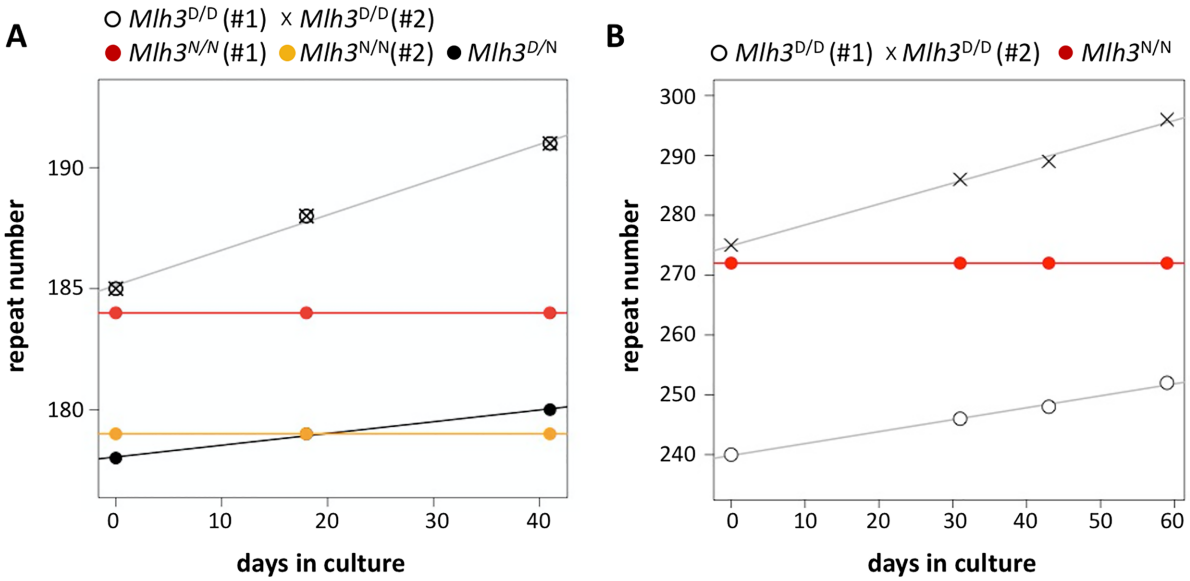

Figure 3. Expansion in $Mlh3^{D/D}$, $Mlh3^{D/N}$ and $Mlh3^{N/N}$ cell lines with different starting repeat numbers. (**A**) Increase in repeat number with time in culture for $Mlh3^{D/D}$, $Mlh3^{D/N}$ and $Mlh3^{N/N}$ cell lines with <200 repeats. Note that the data points for the two $Mlh3^{D/D}$ lines coincide since they have exactly the same starting repeat number and expand at the same rate. (**B**) Increase in repeat number with time in culture for $Mlh3^{D/D}$ and $Mlh3^{N/N}$ cell lines with >200 repeats.

bases being incorporated into the repaired strand, i.e., expansions, as suggested elsewhere (26). In contrast, MutLα has no particular strand cleavage preference (50). This could in principle result in the generation of equal numbers of expanded alleles and alleles that are repaired so as to regenerate the original allele. If so, fewer expansions might be generated in the absence of MutLγ, but they would not be expected to disappear completely.

It is possible that the specific requirement for MutLγ is related to the fact that, unlike cleavage by MutLα, MutLγ cleavage does not depend on proliferating cell nuclear antigen (PCNA) and replication factor C (RFC) (26,50). PCNA is a polymerase processivity factor that is normally loaded by RFC at primer-template junctions during DNA replication and repair synthesis (51). While small loop-outs have been shown to support PCNA loading on unnicked templates, this occurs at a much lower rate than for a typical nicked substrate *in vitro* and its efficiency varies with the size and sequence of the loop-out (50). It may be that most expansion substrates are generated during transcription rather than replication when PCNA/RFC is not ready to be loaded. Delays in loading may be exacerbated by the specific properties of the loop-outs. It may also be that the specific requirement for MLH3 reflects the role of MLH3-specific interacting partners in the MutLγ-mediated cleavage process.

The fact that the MutLγ nuclease activity is required for expansion suggests that targeting MLH3, as has been suggested for MSH3 (6,12), may be useful therapeutically. Since MLH3, like MSH3, is just a minor player in MMR and enzyme binding pockets are intrinsically suited to bind drug-like molecules, MutLγ may be a good druggable target (42). This may allow somatic expansions to be reduced in those REDs where it affects age at onset and/or disease severity.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors want to thank the members of the Usdin group for their suggestions and advice.

FUNDING

Intramural Program of the NIDDK [DK057808 to K.U.]; CIT [CT000265 to P.J.S.]. Funding for open access charge: Intramural Program of the NIDDK [DK057808-07 to K.U.].
Conflict of interest statement. None declared.

REFERENCES

1. Lozano,R., Rosero,C.A. and Hagerman,R.J. (2014) Fragile X spectrum disorders. *Intractable Rare Dis. Res.*, **3**, 134–146.
2. Zhao,X., Zhang,Y., Wilkins,K., Edelmann,W. and Usdin,K. (2018) MutLgamma promotes repeat expansion in a Fragile X mouse model while EXO1 is protective. *PLoS Genet.*, **14**, e1007719.
3. Pinto,R.M., Dragileva,E., Kirby,A., Lloret,A., Lopez,E., St Claire,J., Panigrahi,G.B., Hou,C., Holloway,K., Gillis,T. *et al.* (2013) Mismatch repair genes Mlh1 and Mlh3 modify CAG instability in Huntington's disease mice: genome-wide and candidate approaches. *PLoS Genet.*, **9**, e1003930.
4. Zhao,X.N., Kumari,D., Gupta,S., Wu,D., Evanitsky,M., Yang,W. and Usdin,K. (2015) Mutsbeta generates both expansions and contractions in a mouse model of the Fragile X-associated disorders. *Hum. Mol. Genet.*, **24**, 7087–7096.
5. Du,J., Campau,E., Soragni,E., Jespersen,C. and Gottesfeld,J.M. (2013) Length-dependent CTG.CAG triplet-repeat expansion in myotonic dystrophy patient-derived induced pluripotent stem cells. *Hum. Mol. Genet.*, **22**, 5276–5287.
6. Halabi,A., Ditch,S., Wang,J. and Grabczyk,E. (2012) DNA mismatch repair complex MutSbeta promotes GAA.TTC repeat expansion in human cells. *J. Biol. Chem.*, **287**, 29958–29967.

7. Foiry,L., Dong,L., Savouret,C., Hubert,L., te Riele,H., Junien,C. and Gourdon,G. (2006) Msh3 is a limiting factor in the formation of intergenerational CTG expansions in DM1 transgenic mice. *Hum. Genet.*, **119**, 520–526.
8. van den Broek,W.J., Nelen,M.R., Wansink,D.G., Coerwinkel,M.M., te Riele,H., Groenen,P.J. and Wieringa,B. (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.
9. Hegan,D.C., Narayanan,L., Jirik,F.R., Edelmann,W., Liskay,R.M. and Glazer,P.M. (2006) Differing patterns of genetic instability in mice deficient in the mismatch repair genes Pms2, Mlh1, Msh2, Msh3 and Msh6. *Carcinogenesis*, **27**, 2402–2408.
10. Prolla,T.A., Baker,S.M., Harris,A.C., Tsao,J.L., Yao,X., Bronner,C.E., Zheng,B., Gordon,M., Reneker,J., Arnheim,N. *et al.* (1998) Tumour susceptibility and spontaneous mutation in mice deficient in Mlh1, Pms1 and Pms2 DNA mismatch repair. *Nat. Genet.*, **18**, 276–279.
11. Genetic Modifiers of Huntington's Disease Consortium. (2019) CAG repeat not polyglutamine length determines timing of Huntington's disease onset. *Cell*, **178**, 887–900.
12. Flower,M., Lomeikaite,V., Ciosi,M., Cumming,S., Morales,F., Lo,K., Hensman Moss,D., Jones,L., Holmans,P., Investigators,T.-H. *et al.* (2019) MSH3 modifies somatic instability and disease severity in Huntington's and myotonic dystrophy type 1. *Brain*, **142**, 1876–1886.
13. Bettencourt,C., Hensman-Moss,D., Flower,M., Wiethoff,S., Brice,A., Goizet,C., Stevanin,G., Koutsis,G., Karadima,G., Panas,M *et al.* (2016) DNA repair pathways underlie a common genetic mechanism modulating onset in polyglutamine diseases. *Ann. Neurol.*, **79**, 983–990.
14. Ciosi,M., Maxwell,A., Cumming,S.A., Hensman Moss,D.J., Alshammari,A.M., Flower,M.D., Durr,A., Leavitt,B.R., Roos,R.A.C., team,T.-H. *et al.* (2019) A genetic association study of glutamine-encoding DNA sequence structures, somatic CAG expansion, and DNA repair gene variants, with Huntington disease clinical outcomes. *EBioMedicine*, **48**, 568–580.
15. Hensman Moss,D.J., Pardinas,A.F., Langbehn,D., Lo,K., Leavitt,B.R., Roos,R., Durr,A., Mead,S., investigators,T.-H., investigators,R. *et al.* (2017) Identification of genetic variants associated with Huntington's disease progression: a genome-wide association study. *Lancet Neurol.*, **16**, 701–711.
16. Hubert,L. Jr., Lin,Y., Dion,V. and Wilson,J.H. (2011) Xpa deficiency reduces CAG trinucleotide repeat instability in neuronal tissues in a mouse model of SCA1. *Hum. Mol. Genet.*, **20**, 4822–4830.
17. Kovtun,I.V., Liu,Y., Bjoras,M., Klungland,A., Wilson,S.H. and McMurray,C.T. (2007) OGG1 initiates age-dependent CAG trinucleotide expansion in somatic cells. *Nature*, **447**, 447–452.
18. Møllersen,L., Rowe,A.D., Illuzzi,J.L., Hildrestrand,G.A., Gerhold,K.J., Tveteras,L., Bjolgerud,A., Wilson,D.M. 3rd, Bjoras,M. and Klungland,A. (2012) Neil1 is a genetic modifier of somatic and germline CAG trinucleotide repeat instability in R6/1 mice. *Hum. Mol. Genet.*, **21**, 4939–4947.
19. Lokanga,R.A., Senejani,A.G., Sweasy,J.B. and Usdin,K. (2015) Heterozygosity for a hypomorphic Polbeta mutation reduces the expansion frequency in a mouse model of the Fragile X-related disorders. *PLoS Genet*, **11**, e1005181.
20. Halabi,A., Fuselier,K.T.B. and Grabczyk,E. (2018) GAA*TTC repeat expansion in human cells is mediated by mismatch repair complex MutLgamma and depends upon the endonuclease domain in MLH3 isoform one. *Nucleic Acids Res.*, **46**, 4022–4032.
21. Bourn,R.L., De Biase,I., Pinto,R.M., Sandi,C., Al-Mahdawi,S., Pook,M.A. and Bidichandani,S.I. (2012) Pms2 suppresses large expansions of the (GAA.TTC)n sequence in neuronal tissues. *PLoS One*, **7**, e47085.
22. Gomes-Pereira,M., Fortune,M.T., Ingram,L., McAbney,J.P. and Monckton,D.G. (2004) Pms2 is a genetic enhancer of trinucleotide CAG.CTG repeat somatic mosaicism: implications for the mechanism of triplet repeat expansion. *Hum. Mol. Genet.*, **13**, 1815–1825.
23. Usdin,K., House,N.C. and Freudenreich,C.H. (2015) Repeat instability during DNA repair: Insights from model systems. *Crit. Rev. Biochem. Mol. Biol.*, **50**, 142–167.
24. McMurray,C.T. (2010) Mechanisms of trinucleotide repeat instability during human development. *Nat. Rev. Genet.*, **11**, 786–799.
25. Cilli,P., Ventura,I., Minoprio,A., Meccia,E., Martire,A., Wilson,S.H., Bignami,M. and Mazzei,F. (2016) Oxidized dNTPs and the OGG1 and MUTYH DNA glycosylases combine to induce CAG/CTG repeat instability. *Nucleic Acids Res.*, **44**, 5190–5203.
26. Kadyrova,L.Y., Gujar,V., Burdett,V., Modrich,P. and Kadyrov,F.A. (2020) Human MutLγ, the MLH1-MLH3 heterodimer, has a novel endonuclease activity that promotes DNA expansion. *Proc. Natl. Acad. Sci. U.S.A.*, **117**, 3535–3542.
27. Iyer,R.R., Pluciennik,A., Napierala,M. and Wells,R.D. (2015) DNA triplet repeat expansion and mismatch repair. *Annu. Rev. Biochem.*, **84**, 199–226.
28. Fischer,J.M., Dudley,S., Miller,A.J. and Liskay,R.M. (2016) An intact Pms2 ATPase domain is not essential for male fertility. *DNA Repair (Amst)*, **39**, 46–51.
29. van Oers,J.M., Roa,S., Werling,U., Liu,Y., Genschel,J., Hou,H. Jr., Sellers,R.S., Modrich,P., Scharff,M.D. and Edelmann,W. (2010) PMS2 endonuclease activity has distinct biological functions and is essential for genome maintenance. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 13384–13389.
30. Lipkin,S.M., Wang,V., Jacoby,R., Banerjee-Basu,S., Baxevanis,A.D., Lynch,H.T., Elliott,R.M. and Collins,F.S. (2000) MLH3: a DNA mismatch repair gene associated with mammalian microsatellite instability. *Nat. Genet.*, **24**, 27–35.
31. Al-Sweel,N., Raghavan,V., Dutta,A., Ajith,V.P., Di Vietro,L., Khondakar,N., Manhart,C.M., Surtees,J.A., Nishant,K.T. and Alani,E. (2017) mlh3 mutations in baker's yeast alter meiotic recombination outcomes by increasing noncrossover events genome-wide. *PLoS Genet.*, **13**, e1006974.
32. Nishant,K.T., Plys,A.J. and Alani,E. (2008) A mutation in the putative MLH3 endonuclease domain confers a defect in both mismatch repair and meiosis in Saccharomyces cerevisiae. *Genetics*, **179**, 747–755.
33. Kondo,E., Horii,A. and Fukushige,S. (2001) The interacting domains of three MutL heterodimers in man: hMLH1 interacts with 36 homologous amino acid residues within hMLH3, hPMS1 and hPMS2. *Nucleic Acids Res.*, **29**, 1695–1702.
34. Santucci-Darmanin,S., Neyton,S., Lespinasse,F., Saunieres,A., Gaudray,P. and Paquis-Flucklinger,V. (2002) The DNA mismatch-repair MLH3 protein interacts with MSH4 in meiotic cells, supporting a role for this MutL homolog in mammalian meiotic recombination. *Hum Mol. Genet.*, **11**, 1697–1706.
35. Entezam,A., Biacsi,R., Orrison,B., Saha,T., Hoffman,G.E., Grabczyk,E., Nussbaum,R.L. and Usdin,K. (2007) Regional FMRP deficits and large repeat expansions into the full mutation range in a new Fragile X premutation mouse model. *Gene*, **395**, 125–134.
36. Gazy,I., Miller,C., Kim,G. and Usdin,K. (2020) CGG repeat expansion, elevated Fmr1 transcription and mitochondrial copy number in a new Fragile X PM mouse embryonic stem cell model. *Front. Cell Dev. Biol.*, doi:10.3389/fcell.2020.00482.
37. Toledo,M., Sun,X., Brieno-Enriquez,M.A., Raghavan,V., Gray,S., Pea,J., Milano,C.R., Venkatesh,A., Patel,L., Borst,P.L *et al.* (2019) A mutation in the endonuclease domain of mouse MLH3 reveals novel roles for MutLgamma during crossover formation in meiotic prophase I. *PLoS Genet.*, **15**, e1008177.
38. Kosinski,J., Plotz,G., Guarne,A., Bujnicki,J.M. and Friedhoff,P. (2008) The PMS2 subunit of human MutLalpha contains a metal ion binding domain of the iron-dependent repressor protein family. *J. Mol. Biol.*, **382**, 610–627.
39. Namadurai,S., Jain,D., Kulkarni,D.S., Tabib,C.R., Friedhoff,P., Rao,D.N. and Nair,D.T. (2010) The C-terminal domain of the MutL homolog from Neisseria gonorrhoeae forms an inverted homodimer. *PLoS One*, **5**, e13726.
40. Rogacheva,M.V., Manhart,C.M., Chen,C., Guarne,A., Surtees,J. and Alani,E. (2014) Mlh1-Mlh3, a meiotic crossover and DNA mismatch repair factor, is a Msh2-Msh3-stimulated endonuclease. *J. Biol. Chem.*, **289**, 5664–5673.
41. Ranjha,L., Anand,R. and Cejka,P. (2014) The Saccharomyces cerevisiae Mlh1-Mlh3 heterodimer is an endonuclease that preferentially binds to Holliday junctions. *J. Biol. Chem.*, **289**, 5674–5686.
42. Copeland,R.A., Harpel,M.R. and Tummino,P.J. (2007) Targeting enzyme inhibitors in drug discovery. *Expert. Opin. Ther. Targets*, **11**, 967–978.

43. Zafra,M.P., Schatoff,E.M., Katti,A., Foronda,M., Breinig,M., Schweitzer,A.Y., Simon,A., Han,T., Goswami,S., Montgomery,E. *et al.* (2018) Optimized base editors enable efficient editing in cells, organoids and mice. *Nat. Biotechnol.*, **36**, 888–893.
44. Hayward,B.E., Zhou,Y., Kumari,D. and Usdin,K. (2016) A set of assays for the comprehensive analysis of FMR1 alleles in the Fragile X-related disorders. *J. Mol. Diagn.*, **18**, 762–774.
45. Miller,C., Kim,G.-Y., Zhao,X. and Usdin,K. (2020) All three mammalian MutL complexes are required for repeat expansion in a mouse cell model of the Fragile X-related disorders. *PLoS Genet.*, **16**, e1008902.
46. Kraulis,P.J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr*, **24**, 946–950.
47. Merritt,E.A. and Bacon,D.J. (1997) Raster3D: photorealistic molecular graphics. *Methods Enzymol.*, **277**, 505–524.
48. Cannavo,E., Marra,G., Sabates-Bellver,J., Menigatti,M., Lipkin,S.M., Fischer,F., Cejka,P. and Jiricny,J. (2005) Expression of the MutL homologue hMLH3 in human cells and its role in DNA mismatch repair. *Cancer Res.*, **65**, 10759–10766.
49. Zhao,X., Gazy,I., Hayward,B., Pintado,E., Hwang,Y.H., Tassone,F. and Usdin,K. (2019) Repeat Instability in the Fragile X-related disorders: lessons from a mouse model. *Brain Sci.*, **9**, E52.
50. Pluciennik,A., Burdett,V., Baitinger,C., Iyer,R.R., Shi,K. and Modrich,P. (2013) Extrahelical (CAG)/(CTG) triplet repeat elements support proliferating cell nuclear antigen loading and MutLalpha endonuclease activation. *Proc. Natl. Acad. Sci. U. S. A.*, **110**, 12277–12282.
51. Choe,K.N. and Moldovan,G.L. (2017) Forging ahead through darkness: PCNA, still the principal conductor at the replication fork. *Mol. Cell*, **65**, 380–392.