[Frontiers in Bioscience 8, s653-665, May 1, 2003]

DNA REPAIR AND TRINUCLEOTIDE REPEAT INSTABILITY

Robert S. Lahue and Danielle L. Slater

Eppley Institute for Research in Cancer and Allied Diseases, and Department of Pathology and Microbiology, University of Nebraska Medical Center, Omaha, Nebraska, USA

TABLE OF CONTENTS

1. ABSTRACT

Genes harboring certain trinucleotide repeat (TNR) sequences are at risk for high-frequency mutations that expand or contract the repeat tract. The triplet sequences CNG (where N = any nucleotide) and GAA are known to cause human disease when they expand by more than a few repeats in certain key genes. One of the crucial questions in the field is the mechanism (or, more likely, mechanisms) of triplet repeat expansions and contractions. The available evidence indicates that TNRs can change length as a result of aberrant DNA replication in proliferating cells. In addition, TNR instability can arise from gene conversion or by error-prone DNA repair whether the cell is dividing or not, since most cell types have recombination and repair activities. The latter of these three sources, DNA repair, is the subject of this review because of some recent provocative findings. Two non-mutually exclusive views of DNA repair and TNR instability predominate at this time. One idea is that aberrant DNA structure within TNRs blocks repair. Thus even cells with normal repair activities are inhibited from preventing expansions or contractions, due to local DNA structures formed by TNR sequences. A pernicious second model is that DNA repair actually contributes to TNR instability. This idea of pro-mutagenic DNA repair, although seemingly counterintuitive, has support from a number of studies. A simple explanation is that repair is triggered either by DNA damage in or near the TNR, or perhaps by the aberrant TNR-DNA structure itself. Subsequent excision of nucleotides is followed by error-prone repair synthesis. The idea that repair synthesis is a culprit in expansions or contractions ties into the established ideafz that DNA replication through TNRs gives rise to instability. Since DNA synthesis also occurs during gene conversion, a common source of TNR instability could well be the errors that arise when DNA polymerases attempt to synthesize the problematic triplet repeat sequence.

2. TRIPLET REPEAT MUTATIONS AND SOURCES OF INSTABILITY

2.1. The curious genetics of triplet repeat expansion diseases

Triplet repeat sequences display enigmatic inheritance patterns in human families affected by Huntington's disease, Fragile X syndrome, myotonic

dystrophy, or other diseases caused by TNR expansions. While each disease has its own genetic characteristics, commonalities underscore the unique nature of these mutational patterns in affected families. These common features include the high frequency of mutation and the relatively large size of many TNR mutations, compared to other DNA microsatellites which usually undergo low-frequency, small length changes. While the absolute frequency of mutation varies among TNR diseases, this review will consider TNR mutations as a single class of high-frequency events. TNR instability commonly shows a strong dependence on allele length (longer tracts are more unstable), and also sequence, since CNG and GAA triplets are known to expand frequently in human families. When the tract is imperfect, *i.e.* interrupted by base pair substitutions, expansions occur significantly less often. There is nearly always a strong proclivity towards expansions rather than contractions in humans (but not in most model systems). The locus-specific nature of TNR instability contrasts with the genome-wide instability associated with loss of DNA repair functions. In fact TNR instability occurs readily in cells and tissues where DNA repair is unaltered. Several recent reviews (1-7) elegantly describe replicative models of instability that will briefly be alluded to here before focusing discussion on DNA repair and TNR mutability. Several reviews (2, 4, 6) highlight structural details of TNRs, including the appropriate citations that must be largely curtailed in the next section of this review.

2.2. Structural features of TNRs

These fascinating genetic traits have encouraged many laboratories to investigate the causes of TNR instability. Part of the answer came from structural studies of TNR DNA *in vitro* (2, 4, 6). Single-stranded CAG, CTG, CGG, and CCG sequences readily form hairpins with substantial thermodynamic and kinetic stability. CGG hairpins can subsequently fold into a more complicated tetraplex structure, which further increases their energetic stability. Alternative structures can also form in the context of a duplex DNA. When linear duplexes containing CTG•CAG repeats plus flanking sequence are denatured and reannealed, slipped-strand and loopout structures are formed in which TNRs are extruded as single-stranded loops. Surprisingly, when one strand carries more repeats than the other, heteroduplex extrusion of excess repeats occurs at preferred locations within the microsatellite (8) rather than being randomly distributed. GAA repeats form triplex structures, although a very recent report suggests that GAA and TTC repeats can also form large hairpins during reiterative DNA synthesis *in vitro* (9). Evidence for alternative structures *in vivo*, such as replicational pausing at TNR tracts in bacteria and yeast (10, 11) and inhibition of repair (12) is indirect but fairly compelling. To our knowledge, no one has definitively proven the existence of TNR hairpins or triplexes *in vivo.* Nonetheless the correlation between the ability to assume unusual secondary structures and the high degree of genetic instability is a powerful argument that DNA structure plays a major role in TNR instability. It also helps explain the sequence-selective, *cis*-acting nature of TNR mutations, since local DNA structure at one gene would not be expected to affect other loci in the same cell. As described later, DNA structure is also thought to inhibit or alter DNA repair processes, thus explaining why TNR mutations can be frequent in cells with normal DNA repair function.

2.3. Biological approaches provide evidence for three sources of instability

In vivo studies of TNR instability have taken advantage of model genetic systems in *E. coli*, yeast, transgenic mice and cell lines derived from them, human cell lines, and human sperm samples (1-7). These studies provided significant evidence for three sources of TNR mutation: aberrant DNA replication, gene conversion, and error-prone DNA repair. It is important to recognize that TNR instability may well be a combination of all three. In other words, these mechanisms are not mutually exclusive. Most of the attention has focused on aberrations in DNA replication because this seems to be a major source of spontaneous TNR instability in proliferating cells. Some of the major evidence supporting replicational models include: stalling of polymerases *in vitro* (13-15) and replication forks *in vivo* (10, 11) at TNR tracts; the fact that TNR mutability often displays orientation effects with respect to the direction of DNA replication (16-21); and alterations in TNR instability in bacterial or yeast mutants with deficiencies in DNA replication (22-24). Although bypass synthesis of TNR hairpins has been frequently invoked to explain contractions, TNR mutations and DNA bypass polymerases have been investigated only once, for yeast pol eta and pol zeta (25) with negative results. In all of these studies, the TNR mutations were apparently spontaneous. In other words, there is limited experimentation on how DNA damage or cellular stress might induce TNR instability. This would seem to be a fruitful area for future research, and one publication (26) described in section 4.4, provides a useful start.

The major caveat to the replicational studies is the requirement for cell proliferation. In humans, expansions and contractions can occur in key non-proliferating tissues like brain. Thus additional sources of instability are likely. The structural idiosyncrasies of TNR DNA, such as slipped-strand and loopout structures, almost certainly contribute to the instability problem in both proliferating and non-proliferating cells (8). For example, slipped TNR homoduplexes arise in non-proliferating cells, while slipped-intermediate heteroduplexes can occur whether proliferation is occurring or not. Considering the impact of unorthodox DNA structures on TNR mutations, and the likelihood that at least some instability occurs in the absence of cell proliferation, there is considerable interest in how DNA metabolizing enzymes aside from DNA polymerases contribute to the mutational potential of TNR sequences. The rest of this review will focus on error-prone DNA repair as a source of triplet repeat mutations. As described later, there is considerable controversy over the possible connection between repair and TNR instability. There are ample opportunities for advancement of this field, especially in terms of developing detailed molecular mechanisms and introducing biochemical exploration of how DNA repair might become error-prone and thus contribute to TNR mutability.

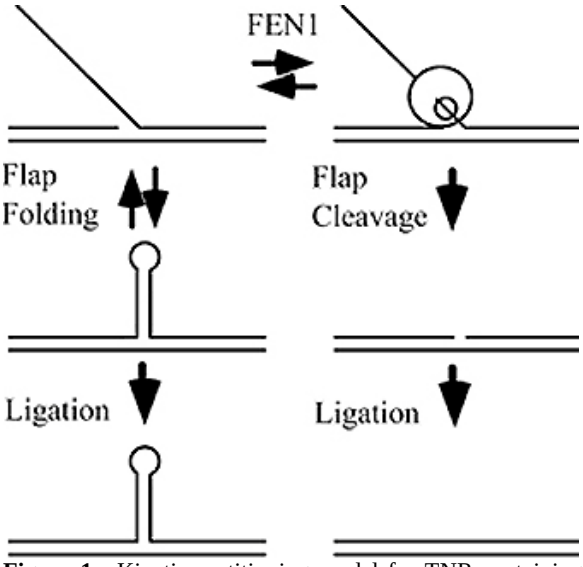

Figure 1. Kinetic partitioning model for TNR-containing DNA flaps, derived in part from (33-36). The upper left panel shows a 5' flap within a TNR-containing stretch of DNA. The flap is presumably formed by strand displacement by a DNA polymerase. If the flap folds quickly (left-hand column), FEN1 activity is precluded and ligation can occur. This leads to an expanded top strand, which can be converted to an expanded duplex in the next round of synthesis. An alternative possibility is that FEN1 binds to the flap prior to its folding and can cleave the ss-dsDNA junction (right panel). This enzymatic action would thereby prevent expansion. FEN1 might also gain access to a flap generated by enzymatic or thermal unwinding of a hairpin; in other words, an unligated hairpin from the left panel could conceivably be unfolded to allow processing by FEN1.

One drawback in the field of TNR instability is that there is no standardized definition of instability in terms of the frequency or size of TNR mutations that qualifies as unstable. Individual laboratories (including ours) tend to use their own, somewhat descriptive, definitions of instability that are often determined by the sensitivity limits of the assays being used. Another limitation of model systems is that contractions almost always predominate. It has been difficult to recapitulate the proclivity towards expansions seen in affected human families. However, expansion bias has been observed in reports for certain model systems (21, 27-31), most of which specify *cis*- and *trans*-acting factors that are important in determing mutation bias. Thus the reader should be aware of differences between systems when evaluating what factors do or do not contribute to instability.

3. FEN1, FLAP ENDONUCLEASE

3.1. FEN1 as an anti-expansion factor

A key finding in 1997 helped trigger the idea that FEN1, a flap endonuclease involved in DNA replication and repair, might provide a key to understanding TNR instability. Kolodner and colleagues (32) showed that yeast mutants lacking the homologous protein Rad27 give rise to many duplication mutations. They hypothesized that flap structures in the mutant could reanneal to the template DNA at regions of microhomology and then be ligated into the newly replicated chain. This would lead to duplication of the intervening sequence in the next round of replication. By inference, the flap processing activity of Rad27 in wild type yeast cells is a powerful mutation prevention system, and FEN1 in higher organisms would provide a similar function. This finding stimulated Gordenin and colleagues (33) to propose that flaps containing triplet repeat sequences fold into hairpins, thus blocking the cleavage activity of FEN1/Rad27, which works optimally on single-stranded flaps. By this model, the creation and folding of flaps within a TNR undergoing DNA synthesis is a major source of expansions. The model does not directly address contractions.

3.2. Biochemical and genetic interactions of FEN1/Rad27 and TNRs

A simple model to describe FEN1/Rad27 activity at TNR-containing flaps is shown in figure 1. Flaps become kinetically partitioned into either folded hairpins that escape nuclease activity or are captured by the enzyme in a single-stranded form and cleaved. This idea is derived largely from elegant *in vitro* analyses of FEN1 (34-36) showing that oligonucleotide flaps with structured TNRs inhibited, but did not completely abolish, cleavage by the enzyme (34, 35, 37). One study (37) found that FEN1 cleaved a CTG repeat flap better than the corresponding CAG repeat substrate, although another report (34) showed similarly reduced levels of cleavage by both sequences. Experiments that included DNA ligase I with FEN1 suggested a competition between the two enzymes (36) in accordance with figure 1. However the appearance of expanded and non-expanded DNA products were irrespective of enzyme concentrations, suggesting that, before enzyme addition, DNA structural intermediates were present in the substrate oligonucleotides which helped dictate the outcome (36). A very recent publication (38) takes this work one step further and suggests that flaps and bubble structures can equilibrate, and that FEN1 uses endonuclease activity to prevent expansions. The important inference from these studies is that expansion can occur even in cells where FEN1/Rad27 are functional, due to the folding propensity of TNR-containing flaps.

In vivo tests of the model have occurred primarily in *rad27* null mutants of *S. cerevisiae* and the results generally support the model as shown. While TNRs show some instability in wild type cells, in *rad27* mutants instability is increased by up to several hundred fold (34, 39-42). Loss of Rad27 function results in approximately equal numbers of expansions and contractions, primarily due to increases in expansion rates (34, 39). Some of the expansions in the nuclease-deficient background are substantially longer than the original TNR tract (34). It is interesting to note that expansions in the *rad27* mutant are not limited to hairpin-forming TNR sequences; expansions of $(CTA)_{25}$ for example showed a similar fold increase as $(CTG)_{25}$, although the absolute rate of instability was significantly higher for CTG repeats in both wild type and

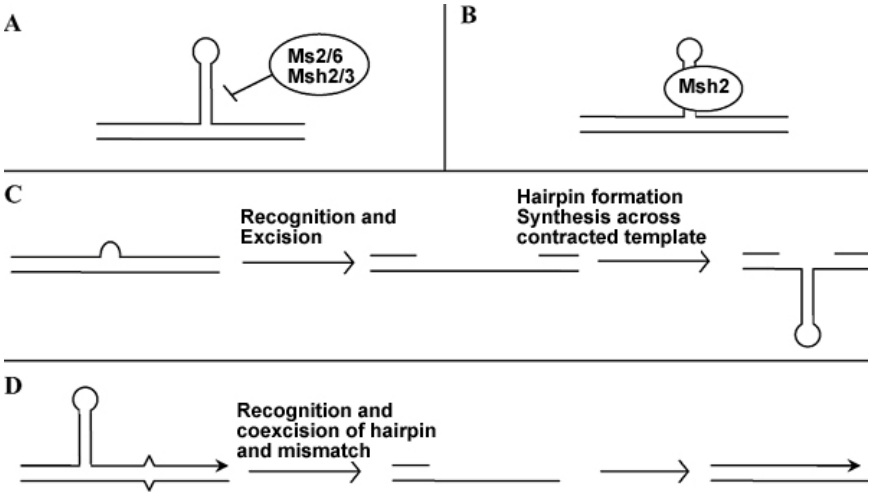

Figure 2. Models describing the role of MR in large TNR stability. A. A TNR hairpin forms during DNA synthesis. The hairpins are refractory to the MR machinery and contractions and expansions occur, unchecked by MR. B. Msh2/Msh3 complex recognizes and binds mismatches within the TNR hairpin. Binding stabilizes the hairpin and the lesion persists; thus expansions and contractions are proposed to occur in an Msh2-dependent manner. C. MR recognizes and excises small, slipped-register loops formed by polymerase slippage. Following excision, hairpins can form within the single stranded TNR region. Subsequent repair synthesis across the shortened template results in a MR-induced contraction. D. The MR system recognizes mismatched bases near a TNR hairpin. MR proteins load onto free DNA termini and initiate coexcision of the mismatch and hairpin, thus reducing TNR instability by removing expansion/contraction precursors.

mutant strains (34). White *et al* (41) observed a 10-fold elevation in CGG tract expansion frequency in a *rad27* strain compared to wild type. Thus expansions of CNG repeats in yeast are partially blocked when Rad27 is present. Interestingly, TNR contractions are also elevated in *rad27* mutants (34, 39-42), although the relative rate increase is not as large as for expansions. This point is notable since contractions probably occur via hairpin formation on the template strand; it is not obvious how the presence or absence of Rad27 processing of flaps on the newly synthesized strand would affect contractions. This is one indication that the situation might be more complex than is currently appreciated.

3.3. Complexities and unanswered questions

While these studies provide compelling evidence that FEN1/Rad27 is a powerful anti-expansion activity, several complexities remain. For example, FEN1 is modulated by interaction with factors such as PCNA (see for example (43, 44), plus FEN1 is covalently modified by acetylation (45). How does modulation affect the ability of FEN1 to act at TNR sequences *in vivo*? Another complexity is that additional or alternative functions besides FEN1 may influence the outcome of flap folding versus flap cleavage. If so, these other functions require identification and characterization. Yeast Dna2 may be a good candidate based on its enzymatic properties (46-48), although *dna2-1* mutants did not affect CAG repeat instability (24). Another candidate is DNA polymerase delta; yeast experiments suggest that the 3' to 5' exonuclease activity of polymerase delta is redundant with Rad27/FEN1 in some aspects of Okazaki fragment processing (49). A third complexity is the pleiotropic nature of yeast with *rad27* null mutations. For example, these mutants are hypermutators at TNRs (34, 39-42), but they also cause frequent mutations at other microsatellite repeats (50, 51) and in non-repeating sequences (32, 52). Most alleles of *rad27* are hyperrecombinant (52), presumably due to alternative flap processing that leads to double-strand breaks (53). *rad27* null strains also show altered processing of telomeres (54), and so on. This pleiotropy suggests that *rad27* effects at TNRs might be indirect as well as direct, making interpretation of data more complicated. For example it is thought that *rad27* mutations lead to strand breaks which are subsequently processed by recombinational repair (33, 55). Thus recombination, not the lack of Rad27 *per se*, might be the mutagenic event. A final complexity in the FEN1 story is that mutations in FEN1 were not observed in Huntington's disease patients, suggesting that loss of FEN1 was not the explanation for CAG tract expansions in these individuals (56). For these reasons, a welcome addition to the field will be more refined experiments in yeast and in mammalian cells, perhaps with more subtle *rad27* or FEN1 mutants, plus additional biochemical characterization to better determine the role of this interesting protein in TNR mutability.

4. MISMATCH REPAIR

4.1. Small and large TNR tract alterations arise differently

The precise role of mismatch repair (MR) in TNR instability is perplexing and controversial. One source of debate is the definition of mutation sizes that define TNR instability. TNR mutations are characterized by both large and small changes in tract size. Small TNR mutations are defined for the purposes of this review as additions/deletions of one to three repeats (three to nine base pairs). Typically, these small mutations occur due to

polymerase slippage events during DNA replication, whereas large mutations are generated by alternatively structured intermediates. In some cases, a small TNR change may be relevant to human disease. For example, expansion of $(CAG)_{35}$ to $(CAG)_{36}$ at the HD locus might mean the difference between normal and a disease life to a Huntington's family member. In general, human TNR diseases are the consequence of the larger TNR mutations; thus large TNR mutations comprise a more biologically relevant form of TNR instability. Nonetheless both size classes are considered here to help delineate the controversy over the role of MR on TNR instability.

Small TNR mutations are normally prevented by the MR system. As with other microsatellites, MR plays a corrective role in rectifying small TNR loops that arise during replicative synthesis. An increase in small expansions and contractions in response to MR defects is a well-documented phenomenon for virtually all microsatellites, including TNRs. In yeast and *E. coli*, TNRs display an increased frequency of small expansions and contractions when MR is inactive (57-61). Thus the role for MR in preventing small TNR insertions and deletions seems clear-cut.

4.2. Evidence that secondary structure evades MR in yeast

Complications arise regarding MR's influence on the frequency of large TNR mutations, as three very different sets of conclusions have been reached from studies in yeast, *E. coli*, and mice. These conclusions are 1) MR does not affect large TNR expansions or contractions in yeast; 2) MR accelerates TNR contractions in bacteria; 3) MR is required for expansions in mice. The yeast data will be considered first. A simple model summarizing this data (figure 2A) presumes that hairpins form during DNA synthesis and these expansion and contraction precursors persist in the cell because the MR machinery does not repair them. Whether the hairpin evades MR at the recognition event or at some downstream step is a question that remains to be answered clearly. The Petes group illustrated that TNRs capable of forming hairpins are inefficiently repaired during meiosis, whereas repair was seen for sequences incapable of hairpin formation (12). Since MR is active during meiosis, the inference is that MR failed to repair the structured loops. Moreover, additional yeast studies with mitotically-dividing cells showed that the MR system does not affect the frequency of large contractions (19, 61) or large expansions (20, 41, 61, 62). Presumably hydrogen bonding within TNR hairpins allow the lesions to escape repair (figure 2A). Although to our knowledge the blocking effect of TNR hairpins on MR proteins has not been examined biochemically, MR is known to be ineffective at correcting hairpins formed by palindromic sequences during yeast recombination (63, 64). However placement of a hairpin-forming palindrome near a well-repaired mismatch leads to efficient correction of both alleles in yeast (65). This observation indicates that a hairpin is repairable if MR can be initiated nearby. Evidence for MR-mediated co-excision of a mismatch and a TNR hairpin are described in section 4.5. In agreement with the yeast data, studies from human cell culture (66) also failed to show an MR effect on TNR instability. Mutations in *hMLH1* or *hMSH2* did not result in significant expansion, within the limits of detection, of the TNRs at the endogenous myotonic dystrophy or Fragile X loci. However the TNR tract lengths used in this study were relatively short, so the possibility remains that human MR might affect instability of long alleles.

4.3. Evidence for Msh2-dependent TNR instability: a gap repair model

Since TNR instability in humans has been most extensively documented during transmission from parent to child, it raises a pertinent question; does TNR instability occur during meiosis or are most mutations generated by other mechanisms prior to or after meiosis? This point is relevant in helping distinguish between replication, recombination, and repair as major sources of instability. Evidence from human genetics suggests that most instability associated with transmission between generations may not occur during meiosis *per se*, but during prezygotic mitotic divisions (67, 68) or postzygotically (69). Additional evidence in mice supports this assertion. Kovtun and McMurray (70) investigated instability of a $(CAG)_{117}$ tract at the HD locus of transgenic mice. Using PCR-mediated amplification and tract size analysis, they observed that most male germline expansions are limited to the post-meiotic haploid germ cell. This evidence suggests neither replication nor recombination is a causative factor for murine germline expansions; alternatively another synthetic event such as gap repair must be responsible. In a similar vein, Kaytor *et al* (71) concluded that recombinational repair, rather than replication or unequal sister chromatid exchange, best explains SCA1 CAG repeat instability in oocytes of transgenic mice.

Within the mouse system MR appears to be a contributing factor in TNR instability. Data from three studies suggest the provocative conclusion that the Msh2/Msh3 heterodimer appears to be an active contributor to repeat expansion in mice, contrary to the normal role of MR in preventing mutations. Two groups have illustrated via *Msh2*$^{-/-}$ knockout strains that TNR expansions in mice, both germline and somatic, are dependent on Msh2 protein (70, 72). In addition, the Wieringa group examined the role of each of the MR heterodimers, Msh2/Msh3 and Msh2/Msh6, using *Msh3*$^{-/-}$ and *Msh6*$^{-/-}$ knockouts (73). They observed that somatic expansions of a $(CTG)_{84}$ allele were eliminated in the *Msh3*$^{-/-}$ strain, consistent with the hypothesis that the Msh2/Msh3 dimer is contributory to TNR expansions. In contrast, an increase in expansion frequency was seen in the *Msh6*$^{-/-}$ background, suggesting that the mouse Msh2/Msh6 complex somehow helps prevent expansions. One caveat to the idea of Msh2-dependent repair in late stages of spermatogenesis is the report that the mRNA and protein levels of Msh2 and Pms2 are severely reduced in postmeiotic spermatids (74), although Msh3 mRNA levels remain high late into spermatogenesis. Clearly the timing of the putative contribution of MR to TNR instability remains to be resolved.

The biochemical mechanism of Msh2-dependent TNR instability is still unclear. The various groups (70, 72, 73) advocate a model (figure 2B) where an Msh2-dependent complex binds to and stabilizes TNR hairpins. The results of van den Broek *et al* (73) indicate that the Msh2/Msh3 heterodimer binds the hairpins. Perhaps the base-base mismatches that occur at every third position within the stem of TNR hairpins are the binding target (70). This binding is proposed to stabilize the hairpin, allowing the lesion to persist and then be utilized as an expanded template during subsequent DNA synthesis. This model seems most appealing for relatively small slipouts that might not have sufficient length to survive cellular nucleases without protection from a binding protein. The anti-expansion function of the Msh2/Msh6 heterodimer was suggested (73) to arise through competitive binding of Msh2 by Msh6, thus reducing the availability of the pro-mutagenic Msh2/Msh3 complex. An important caveat of this model is that the TNR hairpin binding affinities of the Msh2/6 and Msh2/3 heterodimers have not yet been determined. The sole paper that examined TNR hairpin binding (75) utilized Msh2 protein without either of its Msh3 or Msh6 partners. Additional TNR hairpin binding studies using the Msh2/6 and Msh2/3 complexes would be a worthwhile addition to the field.

It is not yet clear how much of the total expansion capacity in mice is msh2-dependent, nor is it known whether this effect is part of a mismatch repair response *per se.* Alternatively, the results might be due to a specialized function of msh2 and its heterodimeric partner msh3 that does not require other mr proteins. For example, the mcmurray group speculates that msh2 binding and hairpin stabilization occurs during gap repair of single-strand or double-strand breaks (70). To help distinguish mr and gap repair models, mouse studies with some of the mutl homolog knockouts such as *mlh1* or *pms2* should help establish which model better explains the apparent triggering of expansions. Another key question is whether human *msh2* provides a similar pro-expansion function as indicated by the mouse studies, but the appropriate human cell culture studies with long tnr tracts in *msh2*$^{+/+}$ and *msh2*$^{-/-}$ backgrounds have yet to be reported. Specific mr assays with hairpin-containing substrates would also be informative.

4.4. Error-prone repair synthesis during MR

Interesting observations in *E. coli* also suggest that MR contributes to TNR instability, but primarily through facilitating contractions. Several studies showed a higher frequency of large TNR contractions in wild type as compared to MR-deficient strains (57-60, 76). Furthermore, Parniewski *et al* (57) determined that the frequency of the MR-dependent large contractions is directly proportional to the length of the TNR tract. These results implicate a model (figure 2C) where TNR contractions arise during normal MR. With this model, MR recognizes and induces excision of one repeat (three base) TNR loopouts that typically arise during replication due to polymerase slippage. Excision results in a stretch of single stranded DNA and the repeats within are capable of folding into a stable hairpin. Repair synthesis across this shortened template generates a contractive lesion and subsequent rounds of replication result in a contraction. This model, however, does not provide a simple explanation for expansions due to MR action, as has been inferred from the mouse studies described earlier.

Collectively, the evidence from *E. coli* suggests that the action of MR is a significant route for the formation of contractive hairpins. An important implication of this error-prone repair synthesis model is that any MR-mediated excision event resulting in a single stranded TNR region could increase TNR instability. Thus one could envision that DNA damage or mismatched bases within or near a TNR tract could elicit an error-prone response by MR. Hashem and Sinden (26) observed that treatment of *E. coli* with DNA damaging agents or with anticancer agents induced deletion of expanded repeats. Induction of deletions required heavy dosing of the cells such that survival was only 10-50%. Presumably deletion occurs through repair of DNA damage within TNR; however the role of DNA repair in this system was not specifically addressed. The concept of repair-induced TNR instability would be further validated with evidence illustrating that TNR instability is increased when MR is specifically elicited to repair the TNR or a nearby region.

4.5. MR-dependent stabilization of interrupted TNRs: a coexcision model

Observations from human genetics noted that interrupted alleles containing one to three base pair substitutions are genetically stable, whereas expanded alleles show the loss of one or more interruptions. This finding indicates that interruptions stabilize repeats, but no direct tests had examined the stabilization mechanism. There is now evidence that MR is the major stabilizing factor. Experiments utilizing interrupted TNR alleles in yeast illustrate that significantly more expansions occur in the absence of MR (62). Also, stabilization of interrupted alleles was mismatch specific, as predicted by a MR-dependent model, since poorly recognized C-C mispairs failed to stabilize the interrupted TNR even in wild type strains. In addition, by utilizing a specific yeast *msh2* allele which encodes a protein proficient in mismatch binding but deficient in the downstream functions, the Lahue group illustrated that TNR stabilization requires the downstream effects of Msh2 and not merely Msh2 binding (62). While MR clearly provides most of the stabilization in these experiments (about 20-fold), there was also a minor (two to five-fold) stabilizing influence that was MR-independent. The interruptions likely reduce the hairpin stability by increased mispairing in the stem (77).

The molecular model for stabilization of interrupted repeats (figure 2D) entails that MR recognizes mismatches which occur due to the noncomplementarity of the interrupted bases. Upon mismatch recognition, the MR machinery excises both the mismatch and the hairpin, using the nearby DNA terminus as a recognition site for the newly synthesized strand. In the absence of MR, excision does not occur and the expansion precursors persist. This model predicts that contractions would not be as well-stabilized by interruptions or by MR, because free termini

are not normally present on the template strand; therefore MR cannot load onto and repair the hairpins that lead to contractions. Corroborating contraction data were also reported for interrupted TNRs in yeast (62). One appeal of this model is the parallel with evidence from palindromic hairpins (65), which indicates that a hairpin is repairable if MR can be initiated nearby. A slight variation of this idea (62) is that the mismatch might reside within the TNR hairpin, rather than in the duplex; however the mismatch still stimulates co-excision of the hairpin, leading to its removal. Another significant inference from the model in figure 2D is the interesting idea that MR helps prevent neurodegenerative diseases by inhibiting TNR expansions from interrupted alleles, in addition to MR's well-known anticancer activity. This model remains to be tested in human cells, to see if the principle applies to a higher system than yeast.

5. DOUBLE-STRAND BREAK REPAIR AND GENETIC RECOMBINATION

Double-strand break repair (DSBR) proteins are active in processing broken chromosomes. In addition, many of these proteins are also active during genetic recombination, especially in meiotic cells, where exchanges are generally accepted to be instigated by double-strand breaks. Therefore we will evaluate together the potential roles of DSBR and recombination on triplet repeat mutagenesis. Some relevant break repair studies were described in section 4.3.

It is important to recognize that in human families affected by TNR instability, there is little or no evidence for crossing-over as a source of instability. For example, flanking markers are rarely if ever exchanged in conjunction with expansions. Also the length of the shorter, non-disease causing allele is almost always left unaltered when the long allele undergoes expansion. These observations argue quite strongly against a reciprocal crossing-over mode of instability. Therefore gene conversion events are the most frequent target of experimental studies because gene conversion, defined as the non-reciprocal exchange of sequence information between alleles, provides a satisfactory explanation of the genetic observations.

5.1. Gene conversion in *E. coli* and in mitotic yeast cells

Work with *E. coli* promotes the idea that, at least for CTG•CAG sequences, TNR mutations can arise from gene conversion events. In experiments where cells carried two different CTG•CAG plasmids, cointegrates were observed that were associated with frequent expansions (78, 79). Instability in this system was tract length-dependent, required the recombination proteins RecA and RecBCD, was biased towards expansions, and was reduced by interruptions in the repeat tracts. Curiously there was no evidence of recombination between like plasmids, even in high copy; it would seem that if the CTG•CAG repeats are the major stimulus for recombination, then repeats on two plasmid molecules in the same cell would also be prone to instability, regardless of the nature of the vector sequences. Little or no instability was observed for repeats of CGG•CCG, GAA•TTC, or GTC•GAC (78), even though these sequences are capable of forming unusual secondary structures (2, 4, 6) and they are unstable in humans (1-7) and yeast (80, 81).

Assuming that these gene conversion experiments in *E. coli* recapitulate at least some of the instability in human tissues, the bacterial results provide a mechanism of instability in non-proliferating human cells, since nearly all cells harbor recombination activities whether the cells are dividing or not. In contrast, experiments in yeast, for example, suggest that spontaneous mitotic recombination plays a limited role in TNR instability. Two labs reported that loss of *RAD52* function, which catalyzes the great majority of mitotic gene conversion in yeast, has no detectable effect on CTG•CAG tract instability (19, 20, 40). If CTG•CAG expansions or contractions occur via recombination in yeast, the *rad52* strains would be expected to show a decrease in instability, but this was not seen. There is one important note regarding this issue for the expansion data of Zakian and colleagues (40, 80). Their system involves expansion-associated chromosome fragility in yeast. Expansions lead to a higher likelihood of fragility in or near the TNR tract. The expansion step of this process, assayed by Southern blotting, was reported to not involve *RAD1*, *RAD50*, *RAD51*, or *RAD52* for CTG•CAG tracts (40). Subsequent recombinational processing of the DSB leads to deletions of the TNR tract plus several kilobase pairs of flanking DNA. The DSB processing was significantly reduced in strains harboring mutations in these recombination genes (40, 80). Thus molecular processes leading to expansions are genetically distinct from the recombinational repair of Fragile sites.

5.2. Meiotic recombination in yeast

Since meiotic recombination occurs with high frequency, recombination enzymes that are error-prone at TNRs would be expected to generate substantial levels of expansions or contractions. In part to test this idea, three studies compared the frequency of meiotic and mitotic TNR alterations in yeast cells (41, 82, 83). There were either no significant changes in CGG repeat instability compared to mitotically growing cells for tracts up to 74 repeats long (41) or there were 4- to 10-fold increases in TNR mutation frequency (with contractions comprising the majority of events) for CAG•CTG tracts up to 79 repeats (82, 83). An additional study with $(CAG)_{78}$ repeats (84) did not directly compare mitotic and meiotic instability, but data from an earlier mitotic study from the same laboratory (17) suggest similar levels of instability in both types of experiments. Since genetic recombination in yeast is elevated about 1,000-fold in meiosis compared to mitosis (85), these findings indicate that while TNR instability can increase during meiosis, overall the levels of instability do not correlate well with induction of meiotic recombination in yeast. This conclusion is rather striking since long CAG repeats can stimulate frequent double-strand breaks that map to the repeat allele (83, 86), although Schweitzer *et al* (84) describe findings that did not support frequent DSBs at CAG tracts. If meiotic recombination events were error-prone, it would seem that the DSBs in the repetitive

sequence would provide an excellent opportunity for a high frequency of expansions or contractions, yet this prediction is not borne out by available evidence in meiotic yeast cells.

5.3. Double-Strand break repair In yeast

Haber and colleagues (42, 87) described a system in which dsbs were induced by ho endonuclease at a specific chromosomal site such that repair by homologous recombination required the use of a tnr-containing plasmid template. For cag•ctg tracts of 5-40 repeats, the recipient locus (suffering the dsb) showed contractions in about 20% of the events, but no significant level of expansions were seen (42). When the tnr length in the plasmid donor sequence was increased to 97-98 repeats and dsbs were induced, expansions of +11 to +42 repeats comprised about 30% of all tnr repeat changes (87). Thus long cag•ctg sequences were readily expanded during dna synthesis associated with repair of the break. Control experiments without induction of the endonuclease produced only contractions. Another important observation was that no expansions were seen for a $(caa)_{87}$ control sequence (87). Since caa repeats do not readily form hairpins (77), this evidence suggested that structure formation by cag or ctg sequences was important for instability. In support of this idea, cell survival after dsb induction (which requires traversal of the dsb by dna repair synthesis) was lower for the cag•ctg tracts than for the caa•ttg control. Richard *et al* (87) went on to test the hypothesis that the yeast mre11-rad50-xrs2 nuclease complex normally processes cag or ctg hairpins and thus prevents at least some expansions. Mutation of either *mre11* or *rad50* led to a higher proportion of expansions compared with contractions. Overexpression of mre11 or rad50 did not alter the ratio of expansions to contractions, but these conditions increased the average size of expansions to about twice that seen in wild type controls. Together these results suggest that conditions that stabilize the mre11-rad50-xrs2 complex facilitate removal of cag or ctg hairpins that occur during dna synthesis associated with dsbr/gene conversion of these repeats. Thus repair synthesis associated with dsbr at tnrs can be mutagenic. Since mre11 shows homology to *e. Coli* sbcc and rad50 to sbcd, these findings are consistent with the observations of sarkar *et al* (88) that *sbcc* mutants promote expansions of cag repeats in bacteria (see section 7.2). In light of the results described earlier in *e. Coli* and different yeast systems, clearly there are some interesting issues to be resolved for gene conversion and dsbr as avenues of tnr instability.

6. NUCLEOTIDE EXCISION REPAIR

Interaction of nucleotide excision repair (NER) proteins with TNRs has been examined in *E. coli* (58, 89, 90). The premise of this work is that an NER damage recognition protein ($UvrA_2$ in this bacterium), which is known to have a broad range of substrates, might recognize aberrant TNR structure. Perhaps $UvrA_2$ binds the TNR structure and stimulates excision repair by NER or by an alternative pathway. Experiments examining TNR sequences in extensively propagated plasmids (89) showed a substantial increase in $(CTG)_{175}$ large contraction frequency for *uvrA* mutants compared to wild type controls, provided that active transcription occurred through the TNR sequences. This prompted the proposal (89) that duplex opening as a result of transcription facilitates hairpin formation, which is then destabilized by $UvrA_2$ to permit accurate replication. Since the behavior of the CTG repeats was significantly different in the *uvrA* mutant from the *uvrB* mutant (89), these results suggest that NER *per se* is not involved, but may indicate an alternative function of $UvrA_2$. In support of this idea, no effect of UV irradiation on instability was detected in wild type cells (89). A subsequent study (90) used genetic assays, a heteroduplex transformation assay, and DNA binding experiments to investigate the role of UvrA. This study came to the opposite conclusion, namely that deletions were less frequent in the *uvrA* mutant for CTG•CAG tracts of 25 to 79 repeats. The authors (90) suggested that differences in tract lengths might explain the discordant results between their work and the earlier study (89). The study by Oussatcheva *et al* (90) supported the idea that $UvrA_2$ binds to TNR-containing loops and initiates repair, but that the repair synthesis is error-prone and leads to deletions. The *uvrB* mutant behaved differently than the *uvrA* strain, and the DNA binding studies indicated that only $UvrA_2$ bound to CAG-loop containing substrates. In NER, loading of the UvrB helicase is the next step; however for the triplet repeat experiments (90) no evidence was seen for a UvrB•DNA complex, even when $UvrA_2B$ was provided to potentially load UvrB onto the DNA. Expansions were not described in either of these studies. In summary, the available evidence from *E. coli* (89, 90) suggests that $UvrA_2$ acts through a non-canonical pathway to affect CAG•CTG deletions, but that the outcome of UvrA action is dependent on the initial TNR length. Tests of eukaryotic NER for a role in TNR instability have lagged behind. The sole report for eukaryotic NER and TNR instability (40) showed no increase in large expansions or contractions of CTG repeats in a *rad1* (nuclease) mutant compared to wild type. If eukaryotic NER contributes to expansions or contractions, additional experiments will be needed to look more closely at the DNA damage recognition factors which best parallel $UvrA_2$ in *E. coli*.

7. OTHER REPAIR/PROCESSING PATHWAYS

7.1. Repair of large loops in yeast is inhibited by structure-forming TNRs

A 1999 paper by Petes and colleagues (12) was influential in establishing the idea that TNR-mediated structure escapes DNA repair in meiotic yeast cells. The experiment was to monitor recombination between two alleles of the *his4* gene, where one carried an insert of 10 repeats. The repeats tested were either competent for secondary structure when single-stranded (CTG or a palindromic sequence) or control sequences without structure. The presumed recombination heteroduplex, formed by pairing of the Watson strand from one allele with the Crick strand of the second allele, would contain a single-stranded loop of 30 bases. If the loop was unstructured, repair was efficient, as judged by a normal inheritance pattern of the haploid progeny. If the loop was capable of forming a hairpin, repair became inefficient.

The paper went on to show that structure was the inhibitory element, not just the sequence of the loop. Although hairpin structures were deduced, not directly demonstrated, the Moore *et al* paper (12) helped establish the idea of structure-mediated inhibition of DNA repair. Since MR is active during meiotic recombination, clearly this is one repair activity that is blunted by TNR-mediated structure. A second pathway capable of repairing large loops during meiosis has only been partially characterized (91, 92), and it does not seem to correspond precisely with the standard repair pathways. Clearly however this meiotic loop repair activity is inhibited by TNR-mediated secondary structure (12). A second yeast loop-correcting activity called large loop repair (LLR) has been observed in logarithmically growing cells and in cell-free extracts (93, 94). LLR is genetically distinct from the activity seen in meiosis; however LLR is also sensitive to secondary structure within the loop (94). As few as eight looped-out CTG repeats were sufficient to inhibit large loop repair activity, as measured either in cell-free extracts or *in vivo* by transformation with heteroduplex plasmid substrates. Together these studies (12, 94) provide good evidence that single-stranded TNR sequences adopt conformations that inhibit MR and two loop repair pathways.

7.2. Nucleases that process TNR-containing structures

Although the unusual secondary structures adopted by TNRs are generally inhibitory to nucleases, there is some evidence for processing of slipped-strand structures. These nucleases were first identified for their activity in other functions on DNA, such as processing of recombination intermediates, but they may also function as nucleases when TNR sequences are present. The SbcCD processing function in *E. coli* may be one example of an enzyme capable of cleaving DNA containing TNR loops (88). SbcCD is a heterodimeric complex with nuclease activity that is important in certain recombination pathways in *E. coli.* SbcCD shares structural and functional aspects with yeast Rad50 and Mre11 proteins, respectively (95). Some *E. coli* strains lacking SbcCD activity, due to an *sbcC* mutation, lead to a striking degree of CTG repeat instability. Lower growth temperature (25° C) favored expansions whereas the normal 37° C growth temperature resulted in a high frequency of CTG tract deletion. The reason for this striking temperature dependence is not clear. Introduction of a wild type *sbcC* gene into the *sbcC* mutant strains reversed the tendency towards large expansions. Since its publication in 1998, however, this report (88) has been followed up with only a single other paper (90) that examined *sbcC* and TNRs, and the later study examined only deletions.

There is some indirect evidence from *in vitro* studies that favors the idea of structure-specific activity on aberrant TNR conformations. Pearson and colleagues (8) showed that T7 endonuclease I could cleave the looped-out strand of CAG- and CTG-containing repeats, and that this enzyme also cleaved the junction of the loop with the non-repeating duplex region. Furthermore the CAG repeat-containing loop (but not its more structured CTG counterpart) was preferentially bound by single-strand binding protein. If the human proteins with similar activities also act upon these putative mutagenic intermediates, the direct processing of slipped-structures may be involved in repeat instability. It will be very interesting to see if the human counterparts of T7 endo I and single-strand binding protein do indeed recognize these structures and somehow channel them into an error-prone repair pathway.

8. CONCLUSIONS AND PERSPECTIVES

The available evidence is fairly compelling that DNA repair pathways respond to TNRs in one of two ways. In some cases the aberrant structure formed by the DNA repeat inhibits repair functions. Examples include FEN1/Rad27, mismatch repair in yeast and human cells, and two different large loop repair pathways. According to this idea, expansions and contractions are frequent even in normal cells because repair pathways are defeated by TNR DNA structures. Therefore TNR instability does not require loss of DNA repair capacity, in distinct contrast to hereditary cancer syndromes like XP and HNPCC that are caused by inherited deficiencies in repair. A second, more dangerous role for DNA repair is that it promotes TNR instability either through error-prone repair synthesis and/or by stabilizing mutational intermediates. Error-prone repair is supported by studies of MR and NER in *E. coli*, MR in mice, and gene conversion/break repair in yeast and mice. An interesting question regarding the error-prone model is whether repair proteins are responding directly to aberrant TNR structures or whether DNA damage in or near the repeat triggers excision of bases followed by error-prone repair synthesis. If DNA damage is the initiating factor, can damage be somehow controlled to either minimize expansions or perhaps to stimulate contractions? Another attractive area for future research is to utilize biochemical systems to better understand the interplay between repair and instability. The FEN1 studies have lead the way in this regard and they encourage more experimentation using defined systems. A last feature that has so far baffled the field is why humans show a propensity towards expansions once a TNR tract exceeds the key threshold length, whereas model systems fail to recapitulate the prevalence of expansions. Perhaps differences in DNA repair between humans and other organisms are the key to understanding the fascinating properties of triplet repeat instability.

9. ACKNOWLEDGMENTS

This work was supported by grants GM61961 and GM65508 from the National Institutes of Health (to R.S.L) and by a graduate fellowship from the University of Nebraska Medical Center (to D.L.S.).

10. REFERENCES

1. Cummings, C. J. & H. Y. Zoghbi: Fourteen and counting: unraveling trinucleotide repeat diseases. *Hum. Mol. Genet.* 9, 909-916 (2000)

2. Usdin, K. & E. Grabczyk: DNA repeat expansions and human disease. *Cell. Mol. Life Sci.* 57, 914-931 (2000)

3. Richards, R. I.: Dynamic mutations: a decade of unstable expanded repeats in human genetic disease. *Hum. Mol. Genet.* 10, 2187-2194 (2001)

4. Kovtun, I., G. Goellner, & C. T. McMurray: Structural features of trinucleotide repeats associated with DNA expansion. *Biochem. Cell Biol.* 79, 325-336 (2001)

5. Bowater, R. P. & R. D. Wells: The intrinsically unstable life of DNA triplet repeats associated with human hereditary disorders. *Prog. Nucleic Acid. Res. Mol. Biol.* 66, 159-202 (2001)

6. Sinden, R. R., V. N. Potaman, E. A. Oussatcheva, C. E. Pearson, Y. L. Lyubchenko, & L. S. Shlyakhtenko: Triplet repeat DNA structures and human genetic disease: dynamic mutations from dynamic DNA. *J. Biosci.* 27, 53-65 (2002)

7. Lenzmeier, B. A. & C. H. Freudenreich: Trinucleotide Repeat Instability: A Hairpin Curve at the Crossroads of Replication, Recombination, and Repair. *Cytogenet. Genome Res.* in press, (2003)

8. Pearson, C. E., M. Tam, Y.-H. Wang, S. E. Montgomery, A. C. Dar, J. D. Cleary, & K. Nichol: Slipped-strand DNAs formed by long (CAG)•(CTG) repeats: slipped-out repeats and slip-out junctions. *Nucleic Acids Res.* 30, 4534-4547 (2002)

9. Heidenfelder, B. L., A. M. Makhov, & M. D. Topal: Hairpin formation in Friedreich's ataxia triplet repeat expansion. *J. Biol. Chem.* 278, 2425-2431 (2003)

10. Samadashwily, G. M., G. Raca, & S. M. Mirkin: Trinucleotide repeats affect DNA replication *in vivo*. *Nature Genet.* 17, 298-304 (1997)

11. Pelletier, R., A. S. Krasilnikova, G. M. Samadishwily, R. S. Lahue, & S. M. Mirkin: Replication and expansion of trinucleotide repeats in yeast. *Mol. Cell. Biol.* in press, (2003)

12. Moore, H., P. W. Greenwell, C.-P. Liu, N. Arnheim, & T. D. Petes: Triplet repeats form secondary structures that escape DNA repair in yeast. *Proc. Natl. Acad. Sci. USA* 96, 1504-1509 (1999)

13. Kang, S., K. Ohshima, M. Shimizu, S. Amirhaeri, & R. D. Wells: Pausing of DNA synthesis *in vitro* at specific loci in CTG and CGG triplet repeats from human hereditary disease genes. *J. Biol. Chem.* 270, 27014-27021 (1995)

14. Usdin, K. & K. J. Woodford: CGG repeats associated with DNA instability and chromosome fragility form structures that block DNA synthesis *in vitro*. *Nucleic Acids Res.* 23, 4202-4209 (1995)

15. Gacy, A. M. & C. T. McMurray: Influence of hairpins on template reannealing at trinucleotide repeat duplexes: a model for slipped DNA. *Biochemistry* 37, 9426-9434 (1998)

16. Kang, S., A. Jaworski, K. Ohshima, & R. D. Wells: Expansion and deletion of CTG repeats from human disease genes are determined by the direction of replication in *E. coli*. *Nat. Genet.* 10, 213-218 (1995)

17. Maurer, D. J., B. L. O'Callaghan, & D. M. Livingston: Orientation dependence of trinucleotide CAG repeat instability in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 16, 6617-6622 (1996)

18. Freudenreich, C. H., J. B. Stavenhagen, & V. A. Zakian: Stability of a CTG/CAG trinucleotide repeat in yeast is dependent on its orientation in the genome. *Mol. Cell. Biol.* 17, 2090-2098 (1997)

19. Miret, J. J., L. Pessoa-Brandao, & R. S. Lahue: Instability of CAG and CTG trinucleotide repeats in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 17, 3382-3387 (1997)

20. Miret, J. J., L. Pessoa-Brandao, & R. S. Lahue: Orientation-dependent and sequence-specific expansions of CTG/CAG trinucleotide repeats in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* 95, 12438-12443 (1998)

21. Cleary, J. D., K. Nichol, Y.-H. Wang, & C. E. Pearson: Evidence of cis-acting factors in replication-mediated trinucleotide repeat instability in primate cells. *Nat. Genet.* 31, 37-46 (2002)

22. Iyer, R. R., A. Pluciennik, W. A. Rosche, R. R. Sinden, & R. D. Wells: DNA polymerase III proofreading mutants enhance the expansion and deletion of triplet repeat sequences in *Escherichia coli*. *J. Biol. Chem.* 275, 2174-2184 (2000)

23. Schweitzer, J. K. & D. M. Livingston: The effect of DNA replication mutations on CAG tract stability in yeast. *Genetics* 152, 953-963 (1999)

24. Ireland, M. J., S. S. Reinke, & D. M. Livingston: The impact of lagging strand replication mutations on the stability of CAG repeat tracts in yeast. *Genetics* 155, 1657-1665 (2000)

25. Dixon, M. J. & R. S. Lahue: Examining the potential role of DNA polymerases h and z in triplet repeat instability in yeast. *DNA Repair* 1, 763-770 (2002)

26. Hashem, V. I. & R. R. Sinden: Chemotherpeutically induced deletion of expanded triplet repeats. *Mutat. Res.* 508, 107-119 (2002)

27. Seznec, H., A.-S. Lia-Baldini, C. Duros, C. Fouquet, C. Lacroix, H. Hofmann-Radvanyi, C. Junien, & G. Gourdon: Transgenic mice carrying large human genomic sequences with expanded CTG repeat mimic closely the DM CTG repeat intergenerational and somatic instability. *Hum. Mol. Genet.* 9, 1185-1194 (2000)

28. Zhang, Y., D. G. Monckton, M. J. Siciliano, T. H. Connor, & M. L. Meistrich: Age and insertion site dependence of repeat number instability of a human DM1 transgene in individual mouse sperm. *Hum. Mol. Genet.* 11, 791-798 (2002)

29. Panigrahi, G. B., J. D. Cleary, & C. E. Pearson: *In vitro* (CTG)/(CTG) expansions and deletions by human cell extracts. *J. Biol. Chem.* 277, 13926-13934 (2002)

30. Bhattacharyya, S., M. L. Rolfsmeier, M. J. Dixon, K. Wagoner, & R. S. Lahue: Identification of *RTG2* as a modifier gene for CAG•CTG repeat instability in *Saccharomyces cerevisiae*. *Genetics* 162, 579-589 (2002)

31. Libby, R. T., D. G. Monckton, Y.-H. Fu, R. A. Martinez, J. P. McAbney, R. Lau, D. D. Einum, K. Nichol, C. B. Ware, L. J. Ptacek, C. E. Pearson, & A. R. La Spada: Genomic context drives SCA7 CAG repeat instability, while expressed SCA7 cDNAs are intergenerationally and somatically stable in transgenic mice. *Hum. Mol. Genet.* 12, 41-50 (2003)

32. Tishkoff, D. X., N. Filosi, G. M. Gaida, & R. D. Kolodner: A novel mutation avoidance mechanism dependent on *S. cerevisiae RAD27* is distinct from DNA mismatch repair. *Cell* 88, 253-263 (1997)

33. Gordenin, D. A., T. A. Kunkel, & M. A. Resnick: Repeat expansion--all in a flap? *Nat. Genet.* 16, 116-118 (1997)

34. Spiro, C., R. Pelletier, M. L. Rolfsmeier, M. J. Dixon, R. S. Lahue, G. Gupta, M. S. Park, X. Chen, S. V. S. Mariappan, & C. T. McMurray: Inhibition of FEN-1 processing by DNA secondary structure at trinucleotide repeats. *Mol. Cell* 4, 1079-1085 (1999)

35. Henricksen, L. A., S. Tom, Y. Liu, & R. A. Bambara: Inhibition of flap endonuclease 1 by flap secondary structure and relevance to repeat sequence expansion. *J. Biol. Chem.* 275, 16420-16427 (2000)

36. Henricksen, L. A., J. Veeraraghavan, D. R. Chafin, & R. A. Bambara: DNA ligase I competes with FEN1 to expand repetitive DNA sequences *in vitro*. *J. Biol. Chem.* 277, 22361-22369 (2002)

37. Lee, S. & M. S. Park: Human FEN-1 can process the 5'-flap DNA of CTG/CAG triplet repeat derived from human genetic diseases by length and sequence dependent manner. *Exp. Mol. Med.* 34, 313-317 (2002)

38. Liu, Y. & R. A. Bambara: Analysis of human flap endonuclease 1 mutants reveals a mechanism to prevent triplet repeat expansion. *J. Biol. Chem.* M212061200 (2003)

39. Schweitzer, J. K. & D. M. Livingston: Expansions of CAG repeat tracts are frequent in a yeast mutant defective in Okazaki fragment maturation. *Hum. Mol. Genet.* 7, 69-74 (1998)

40. Freudenreich, C. H., S. M. Kantrow, & V. A. Zakian: Expansion and length-dependent fragility of CTG repeats in yeast. *Science* 279, 853-856 (1998)

41. White, P. J., R. H. Borts, & M. C. Hirst: Stability of the human Fragile X (CGG)n triplet repeat array in *Saccharomyces cerevisiae* deficient in aspects of DNA metabolism. *Mol. Cell. Biol.* 19, 5675-5684 (1999)

42. Richard, G.-F., B. Dujon, & J. E. Haber: Double-strand break repair can lead to high frequencies of deletions within short CAG/CTG trinucleotide repeats. *Mol. Gen. Genet.* 261, 871-882 (1999)

43. Tom, S., L. A. Henricksen, & R. A. Bambara: Mechanism whereby proliferating cell nuclear antigen stimulates flap endonuclease 1. *J. Biol. Chem.* 275, 10498-10505 (2000)

44. Gomes, X. V. & P. M. Burgers: Two modes of FEN1 binding to PCNA regulated by DNA. *EMBO J.* 19, 3811-3821 (2000)

45. Hasan, S., M. Stucki, P. O. Hassa, R. Imhof, P. Gehrig, P. Hunziker, U. Hubscher, & M. O. Hottiger: Regulation of human flap endonuclease-1 activity by acetylation through the transcriptional coactivator p300. *Mol. Cell* 7, 1221-1231 (2001)

46. MacNeill, S. A.: DNA replication: Partners in the Okazaki two-step. *Curr. Biol.* 11, R842-844 (2001)

47. Bae, S.-H., D. W. Kim, J. Kim, J.-H. Kim, D.-H. Kim, H.-D. Kim, H.-Y. Kang, & Y.-S. Seo: Coupling of DNA helicase and endonuclease activities of yeast Dna2 facilitates Okazaki fragment processing. *J. Biol. Chem.* 277, 26632-26641 (2002)

48. Ayyagari, R., X. V. Gomes, D. A. Gordenin, & P. M. J. Burgers: Okazaki fragment maturation in yeast: I. Distribution of functions between FEN1 and DNA2. *J. Biol. Chem.* 276, 1618-1625 (2003)

49. Jin, Y. H., R. Obert, P. M. J. Burgers, T. A. Kunkel, M. A. Resnick, & D. A. Gordenin: The 3' to 5' exonuclease of DNA polymerase delta can substitute for the 5' flap endonuclease Rad27/FEN1 in processing Okazaki fragments and preventing genome instability. *Proc. Natl. Acad. Sci. USA* 98, 5122-5127 (2001)

50. Johnson, R. E., G. K. Kovvali, L. Prakash, & S. Prakash: Requirement of the yeast *RTH1* 5' to 3' exonuclease for the stability of simple repetitive DNA. *Science* 269, 238-240 (1995)

51. Kokoska, R. J., L. Stefanovic, H. T. Tran, M. A. Resnick, D. A. Gordenin, & T. D. Petes: Destabilization of yeast micro- and minisatellite DNA sequences by mutations affecting a nuclease involved in Okazaki fragment processing (*rad27*) and DNA polymerase delta (*pol3-t*). *Mol. Cell. Biol.* 18, 2779-2788 (1998)

52. Vallen, E. A. & F. R. Cross: Mutations in *RAD27* define a potential link between G1 cyclins and DNA replication. *Mol. Cell. Biol.* 15, 4291-4302 (1995)

53. Symington, L. S.: Homologous recombination is required for the viability of *rad27* mutants. *Nucleic Acids Res.* 26, 5589-5595 (1998)

54. Parenteau, J. & R. J. Wellinger: Accumulation of single-stranded DNA and destabilization of telomeric repeats in yeast mutant strains carrying a deletion of *RAD27*. *Mol. Cell. Biol.* 19, 4143-4152 (1999)

55. Tishkoff, D. X., A. L. Boerger, P. Bertrand, N. Filosi, G. M. Gaida, M. F. Kane, & R. D. Kolodner: Identification and characterization of *Saccharomyces cerevisiae EXOI*, a gene encoding an exonuclease that interacts with MSH2. *Proc. Natl. Acad. Sci. USA* 94, 7487-7492 (1997)

56. Otto, C. J., E. Almqvist, M. R. Hayden, & S. E. Andrew: The 'flap' endonuclease gene FEN1 is excluded as a candidate gene implicated in the CAG repeat expansion underlying Huntington disease. *Clin. Genet.* 59, 122-127 (2001)

57. Parniewski, P., A. Jaworski, R. D. Wells, & R. P. Bowater: Length of CTG/CAG repeats determines the influence of mismatch repair on genetic instability. *J. Mol. Biol.* 299, 865-874 (2000)

58. Wells, R. D., P. Parniewski, A. Pluciennik, A. Bacolla, R. Gellibolian, & A. Jaworski: Small slipped register genetic instabilities in *Escherichia coli* in triplet repeat sequences associated with hereditary neurological diseases. *J. Biol. Chem.* 273, 19532-19541 (1998)

59. Schmidt, K. H., C. M. Abott, & D. R. F. Leach: Two opposing effects of mismatch repair on CTG repeat instability in *Escherichia coli. Mol. Micro.* 35, 463-471 (2000)

60. Hashem, V. I., W. A. Rosche, & R. R. Sinden: Genetic assays for measuring rates of (CAG)•(CTG) repeat instability in *Escherichia coli. Mutat. Res.* 502, 25-37 (2002)

61. Schweitzer, J. K. & D. M. Livingston: Destabilization of CAG trinucleotide repeat tracts by mismatch repair mutations in yeast. *Hum. Mol. Genet.* 6, 349-355 (1997)

62. Rolfsmeier, M. L., M. J. Dixon, & R. S. Lahue: Mismatch repair blocks expansions of interrupted trinucleotide repeats in yeast. *Mol. Cell* 6, 1501-1507 (2000)

63. Nag, D. K., M. A. White, & T. D. Petes: Palindromic sequences in heteroduplex DNA inhibit mismatch repair in yeast. *Nature* 340, 318-320 (1989)

64. Nag, D. K. & T. D. Petes: Seven-base-pair inverted repeats in DNA form stable hairpins *in vivo* in *Saccharomyces cerevisiae*. *Genetics* 129, 669-673 (1991)

65. Detloff, P. & T. D. Petes: Measurements of excision repair tracts formed during meiotic recombination in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 12, 1805-14 (1992)

66. Kramer, P. R., C. E. Pearson, & R. R. Sinden: Stability of triplet repeats of myotonic dystrophy and Fragile X loci in human mismatch repair cell lines. *Hum. Genet.* 98, 151-157 (1996)

67. Leeflang, E. P., L. Zhang, S. Tavare, R. Hubert, J. Srinidhi, M. E. MacDonald, R. H. Myers, M. de Young, N. S. Wexler, J. F. Gusella, & N. Arnheim: Single sperm analysis of the trinucleotide repeats in the Huntington's disease gene: quantification of the mutation frequency spectrum. *Hum. Mol. Genet.* 4, 1519-1526 (1995)

68. Moutou, C., M.-C. Vincent, V. Biancalana, & J.-L. Mandel: Transition from premutation to full mutation in Fragile X syndrome is likely to be prezygotic. *Hum. Mol. Genet.* 6, 971-979 (1997)

69. Kruyer, H., M. Mila, G. Glover, P. Carbonell, F. Ballesta, & X. Estivill: Fragile X syndrome and the (CGG)n mutation: two families with discordant MX twins. *Am. J. Hum. Genet.* 54, 437-442 (1994)

70. Kovtun, I. V. & C. T. McMurray: Trinucleotide expansion in haploid germ cells by gap repair. *Nature Genet.* 27, 407-411 (2001)

71. Kaytor, M. D., E. N. Burright, L. A. Duvick, H. Y. Zoghbi, & H. T. Orr: Increased trinucleotide repeat instability with advanced maternal age. *Hum. Mol. Genet.* 6, 2135-2139 (1997)

72. Manley, K., T. L. Shirley, L. Flaherty, & A. Messer: *Msh2* deficiency prevents *in vivo* somatic instability of the CAG repeat in Huntington disease transgenic mice. *Nat. Genet.* 23, 471-473 (1999)

73. van den Broek, W. J. A. A., M. R. Nelen, D. G. Wansink, M. M. Coerwinkel, H. te Riele, P. J. T. A. Groenen, & B. Wieringa: 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 (2002)

74. Richardson, L. L., C. Pedigo, & M. A. Handel: Expression of deoxyribonucleic acid repair enzymes during spermatogenesis in mice. *Biol. Reprod.* 62, 789-796 (2000)

75. Pearson, C. E., A. Ewel, S. Acharya, R. A. Fishel, & R. R. Sinden: Human MSH2 binds to trinucleotide repeat DNA structures associated with neurodegenerative diseases. *Hum. Mol. Genet.* 6, 1117-1123 (1997)

76. Jaworski, A., W. A. Rosche, R. Gellibolian, S. Kang, M. Shimizu, R. P. Bowater, R. R. Sinden, & R. D. Wells: Mismatch repair in *Escherichia coli* enhances instability of $(CTG)_n$ triplet repeats from human hereditary diseases. *Proc. Natl. Acad. Sci. USA* 92, 11019-11023 (1995)

77. Gacy, A. M., G. Goellner, N. Juranic, S. Macura, & C. T. McMurray: Trinucleotide repeats that expand in human disease form hairpin structure *in vitro*. *Cell* 81, 533-540 (1995)

78. Jakupciak, J. P. & R. D. Wells: Genetic instabilities in (CTG-CAG) repeats occur by recombination. *J. Biol. Chem.* 274, 23468-23479 (1999)

79. Jakupciak, J. P. & R. D. Wells: Gene conversion (recombination) mediates expansions of CTG•CAG repeats. *J. Biol. Chem.* 275, 40003-40013 (2000)

80. Balakumaran, B. S., C. H. Freudenreich, & V. A. Zakian: CGG/CCG repeats exhibit orientation-dependent instability and orientation-independent fragility in *Saccharomyces cerevisiae*. *Hum. Mol. Genet.* 9, 93-100 (2000)

81. Rolfsmeier, M. L., M. J. Dixon, L. Pessoa-Brandao, R. Pelletier, J. J. Miret, & R. S. Lahue: Cis-elements governing trinucleotide repeat instability in *Saccharomyces cerevisiae*. *Genetics* 157, 1569-1579 (2001)

82. Cohen, H., D. D. Sears, D. Zenvirth, P. Hieter, & G. Simchen: Increased instability of human CTG repeat tracts on yeast artificial chromosomes during gametogenesis. *Mol. Cell. Biol.* 19, 4153-4158 (1999)

83. Jankowski, C., F. Nasar, & D. K. Nag: Meiotic instability of CAG repeat tracts occurs by double-strand break repair in yeast. *Proc. Natl. Acad. Sci. USA* 97, 2134-2139 (2000)

84. Schweitzer, J. K., S. S. Reinke, & D. M. Livingston: Meiotic alterations in CAG repeat tracts. *Genetics* 159, 1861-1865 (2001)

85. Petes, T. D., R. E. Malone, & L. S. Symington, *Recombination in yeast*, in *The Molecular and Cellular Biology of the Yeast Saccharomyces*, J.R. Broach, J.R. Pringle, and E.W. Jones, Editors. 1991, Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY. p. 407-521.

86. Jankowski, C. & D. K. Nag: Most meiotic CAG repeat tract-length alterations in yeast are *SPO11* dependent. *Mol. Genet. Genomics* 267, 64-70 (2002)

87. Richard, G.-F., G. M. Goellner, C. T. McMurray, & J. E. Haber: Recombination-induced CAG trinucleotide repeat expansions in yeast involve the *MRE11-RAD50-XRS2* complex. *EMBO J.* 19, 2381-2390 (2000)

88. Sarkar, P. S., H.-C. Chang, F. B. Boudi, & S. Reddy: CTG repeats show bimodal amplification in E. coli. *Cell* 95, 531-540 (1998)

89. Parniewski, P., A. Bacolla, A. Jaworski, & R. D. Wells: Nucleotide excision repair affects the stability of long transcribed (CTG*CAG) tracts in an orientation-dependent manner in *Escherichia coli*. *Nucleic Acids Res.* 27, 616-623 (1999)

90. Oussatcheva, E. A., V. I. Hashem, Y. Zou, R. R. Sinden, & V. N. Potaman: Involvement of the nucleotide excision repair protein UvrA in instability off CAG•CTG repeat sequences in *Escherichia coli. J. Biol. Chem.* 276, 30878-30884 (2001)

91. Kirkpatrick, D. T. & T. D. Petes: Repair of DNA loops involves DNA-mismatch and nucleotide-excision repair proteins. *Nature* 387, 929-931 (1997)

92. Kearney, H. M., D. T. Kirkpatrick, J. L. Gerton, & T. D. Petes: Meiotic recombination involving heterozygous large insertions in *Saccharomyces cerevisiae*; formation and repair of large, unpaired DNA loops. *Genetics* 158, 1457-1476 (2001)

93. Corrette-Bennett, S. E., B. O. Parker, N. L. Mohlman, & R. S. Lahue: Correction of large mispaired DNA loops by extracts of *Saccharomyces cerevisiae*. *J. Biol. Chem.* 274, 17605-17611 (1999)

94. Corrette-Bennett, S. E., N. L. Mohlman, Z. Rosado, J. J. Miret, P. M. Hess, B. O. Parker, & R. S. Lahue: Efficient repair of large DNA loops in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 29, 4134-4143 (2001)

95. Sharples, G. J. & D. R. Leach: Structural and functional similarities between the SbcCD proteins of Escherichia coli and the RAD50 and MRE11 (RAD32) recombination and repair proteins of yeast. *Mol. Microbiol.* 17, 1215-1217 (1995)

Abbreviations: TNR, trinucleotide repeat; N, any nucleotide; MR, mismatch repair; HD, Huntington's disease; DSBR, double-strand break repair; DSB, double-strand break; NER, nucleotide excision repair; LLR, large loop repair; XP, xeroderma pigmentosum; HNPCC, hereditary non-polyposis colorectal cancer.

Key Words: Trinucleotide Repeat, Microsatellite Instability, Neurodegenerative Diseases, Expansion, Contraction, Hairpin, Mismatch Repair, Flap Endonuclease, Double-strand Break Repair, Nucleotide Excision Repair, Review

Send correspondence to: Robert Lahue, Eppley Institute, Box 986805, University of Nebraska Medical Center, Omaha, NE 68198-6805, USA, Tel: 402-559-4619, Fax: 402-559-8270, E-mail: rlahue@unmc.edu