Cell, Vol. 83, 539–545, November 17, 1995, Copyright © 1995 by Cell Press

The Stabilization of Repetitive Tracts of DNA by Variant Repeats Requires a Functional DNA Mismatch Repair System

Shaun M. Heale and Thomas D. Petes
Department of Biology
University of North Carolina at Chapel Hill
Chapel Hill, North Carolina 27599-3280

Summary

Simple repetitive tracts of DNA are unstable in all organisms thus far examined. In the yeast S. cerevisiae, we show that a 51 bp poly(GT) tract alters length at a rate of about 10^{-5} per cell division. Insertion of a single variant repeat (either AT or CT) into the middle of the poly(GT) tract results in 100-fold stabilization. This stabilization requires the DNA mismatch repair system. Alterations within tracts with variant repeats occur more frequently on one side of the interruption than on the other. The stabilizing effects of variant repeats and polarity of repeat alterations have also been observed in trinucleotide repeats associated with certain human diseases.

Introduction

The genomes of all eukaryotes thus far examined have regions in which a single base or small number of bases is repeated multiple times. A particularly common type of repeat consists of multiple copies of the dinucleotide GT, known as poly(GT). The human genome has about 10^5 poly(GT) tracts (20 or more base pairs in length), whereas the yeast Saccharomyces cerevisiae genome has about 50–100 such tracts (Hamada et al., 1982; Walmsley et al., 1983). In eukaryotes, poly(GT) tracts are quite unstable, altering at rates of about 10^{-4} to 10^{-5} per division in both yeast cells (Henderson and Petes, 1992) and mammalian cells (Farber et al., 1994). In Escherichia coli, poly(GT) tracts are even less stable, with length changes occurring at rates of 10^{-2} to 10^{-3} (Levinson and Gutman, 1987; Freund et al., 1989).

Two mechanisms have been proposed to explain the instability of simple repetitive DNA: unequal recombination (Smith, 1973) and DNA polymerase slippage (Streisinger et al., 1966). Unequal crossing over (Figure 1a) involves pairing between misaligned tandem repeats. An exchange would create one array with an increased number of repeats and a second array with the reciprocal decrease. In DNA polymerase slippage events (Figure 1b), the primer and template strands temporarily dissociate during DNA replication. When the strands reassociate, individual repeats are left unpaired. If the unpaired repeat is in the primer strand, continued DNA synthesis results in a larger tandem array. If the unpaired repeat is in the template strand, the same mechanism results in a deletion. Several arguments indicate that small changes in the numbers of repeats are likely to reflect DNA polymerase slippage rather than recombination. First, the frequency of tract alterations is unaffected by mutations that eliminate most types of recombination in both E. coli (Levinson and Gutman, 1987) and S. cerevisiae (Henderson and Petes, 1992). Second, mutations in genes affecting DNA mismatch repair dramatically elevate repeated sequence instability in E. coli (Levinson and Gutman, 1987) and yeast (Strand et al., 1993). In addition, several forms of familial colorectal cancer in humans are associated with defects in DNA mismatch repair and simple repeat instability (reviewed by Karran and Bignami, 1994). The simplest explanation of the destabilization of simple repeats by mutations eliminating DNA mismatch repair is that displaced repeats are frequently formed by DNA polymerase slippage, but the mismatch repair system corrects most such mismatches, restoring the wild-type tract length. Thus, when the correction system fails, destabilization is observed.

Expansions of trinucleotide sequences have been recently identified as the molecular basis for several human neurological disorders (Nelson and Warren, 1993). Spinocerebellar ataxia type I (SCA1) is caused by expansion of a poly(CAG) tract located near the *SCA1* target gene (Orr et al., 1993), and the fragile X syndrome is caused by expansion of a poly(CGG) tract near the *FMR1* target gene (Kremer et al., 1991). Both *SCA1* and *FMR1* have stable and unstable alleles. Stable *SCA1* alleles are usually interrupted by two CAT elements and have the configuration $(CAG)_nCATCAGCAT(CAG)_n$ with n = 10–20 (Chung et al., 1993). Expanded alleles do not have the CAT interruptions, suggesting that these interruptions may contribute to stability of the poly(CAG) tract. This possibility was supported by the finding that poly(CAG) tracts with interruptions are more stable in mitotic human cells than uninterrupted tracts of approximately the same length (Chong et al., 1995). Similarly, stable *FMR1* alleles often have the configuration $(AGG)(CGG)_9$ as the basic repeat unit with two to five copies of this repeat (Eichler et al., 1994; Hirst et al., 1994; Kunst and Warren, 1994; Snow et al., 1994). Haplotypes associated with the longest uninterrupted stretch of poly(CGG) for normal alleles represent the most common fragile X haplotype. Stabilization of a poly(CCG) tract by variant repeats is also associated with the *FRA16A* fragile site (Nancarrow et al., 1995).

In alleles of *FMR1* in which the uninterrupted stretch of poly(CGG) is expanded, the expansions occur on the 3′ side of the array (Eichler et al., 1994; Hirst et al., 1994; Kunst and Warren, 1994; Snow et al., 1994). Polarity of alterations has also been observed within the human minisatellite *MS32*, which has a 29 bp repeat (Jeffreys et al., 1994). These alterations appear to reflect gene conversion events initiated in one of the flanking sequences.

Below, we report that poly(GT) tracts in yeast are dramatically stabilized by a single variant repeat in the middle of the tract and that this stabilization requires a functional DNA mismatch repair system. We also find that alterations in poly(GT) tracts that contain a variant repeat preferentially occur to one side of the variant repeat.

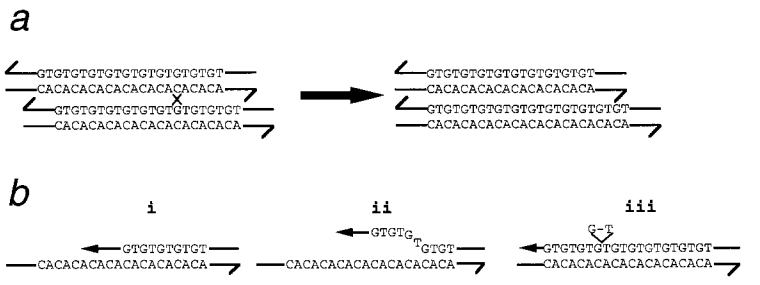

Figure 1. Mechanisms of Altering the Length of Repetitive Tracts

(a) Unequal crossing over. A crossover between misaligned repeats would result in a deletion of repeats from one array and an addition of repeats to the other. A similar nonreciprocal recombination event (gene conversion) could alter the number of repeats in one array without a change in the other.

(b) DNA polymerase slippage. During DNA replication (i), the primer and template strand dissociate (ii). During reassociation, one or more repeats remain unpaired (iii). If the unpaired repeat is in the primer strand (as shown), continued synthesis would lead to an increase in the length of the array. If the unpaired repeat is in the template strand, continued synthesis would decrease the length of the array.

Results

Rationale

To examine the effect of variant repeats on the stability of repetitive DNA sequences, we used a plasmid system in which alterations in the length of repetitive tract could be monitored by using selective media (Henderson and Petes, 1992; Strand et al., 1993). These plasmids contained an in-frame insertion of a poly(GT) tract within the coding sequence of a gene encoding a fusion protein with wild-type *URA3* activity (Figure 2). Since strains with the wild-type *URA3* gene cannot grow on media containing 5-fluororotic acid (5-FOA) (Boeke et al., 1984), strains with plasmids with in-frame insertions cannot grow on 5-FOA-containing media. Cells containing the plasmid that are capable of growth in such media usually contain a poly(GT) tract of altered length resulting in an out-of-frame insertion or a mutation elsewhere in the plasmid-borne *URA3* gene (Henderson and Petes, 1992).

We examined the stability of poly(GT) tracts of four types (plasmids containing the tracts indicated in parentheses): a 51 bp tract with no variant repeats (pTA4), a 51 bp tract with a variant AT repeat in the middle (pTA1), a 51 bp tract with a variant CT repeat in the middle (pTA3), and a 69 bp tract with a variant AT repeat in the middle (pTA2). The sequences of the insertions in each plasmid are shown in Table 1, and the yeast strains containing the plasmids are described in Table 2. By measuring the rate of tract instability in cells containing these plasmids and by DNA sequence analysis of the altered plasmids, we determined the effect of variant repeats on tract stability.

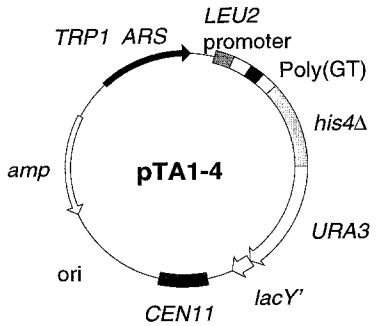

Figure 2. Plasmids Used to Monitor Alterations in Repetitive Tracts

In these plasmids (pTA1 to pTA4), an in-frame poly(GT) tract (with or without a variant repeat) is located within a fusion protein that has wild-type *URA3* activity. Alterations in tract length are detected by plating cells containing the assay plasmid on 5-FOA, which kills Ura$^+$ cells. The sequences of oligonucleotides used for the plasmids pTA1 to pTA4 are shown in Table 1.

Variant Repeats Stabilize Simple Repetitive DNA Tracts in Wild-Type Yeast Cells

The rate of tract alterations in a wild-type strain (SMH6) containing a 51 bp uninterrupted poly(GT) tract was 1×10^{-5} per cell division (Table 3). Sequence analysis of plasmids rescued from 31 independent 5-FOA-resistant (5-FOA^R) derivatives of SMH6 showed 28 plasmids with alterations involving one GT repeat (22 deletions of 2 bp and six additions of 2 bp) and three plasmids with alterations involving two repeats (two deletions of 4 bp and one addition of 4 bp). These results confirm previous observations that the most common types of tract alteration involve deletion or addition of only one or two repeats (Henderson and Petes, 1992).

Strains containing plasmids with interrupted poly(GT) tracts had a lower rate of 5-FOA^R colonies: 3×10^{-6} for SMH3 (AT repeat interrupting a 51 bp tract), 3×10^{-6} for SMH5 (CT repeat interrupting a 51 bp tract), and 4×10^{-6} for SMH4 (AT repeat interrupting a 69 bp tract). When plasmids were rescued from 5-FOA^R derivatives of these strains and sequenced, 0 of 18 plasmids from SMH3, 0 of 15 from SMH5, and only 1 of 18 from SMH4 had tract alterations. Consequently, most of the 5-FOA^R colonies derived from these strains must reflect an alteration elsewhere within the *URA3* fusion protein rather than an alter-

Table 1. Simple Repetitive Sequences within Plasmids

Plasmid Name	Sequence of Insertion[a]
pTA1	TCGA$(GT)_{12}AT(GT)_{12}G$ $(CA)_{12}TA(CA)_{12}$CCTAG
pTA2	TCGA$(GT)_{16}AT(GT)_{17}G$ $(CA)_{16}TA(CA)_{17}$CCTAG
pTA3	TCGA$(GT)_{12}CT(GT)_{12}G$ $(CA)_{12}GA(CA)_{12}$CCTAG
pTA4	TCGA$(GT)_{25}G$ $(CA)_{25}$CCTAG

[a]The orientation of the sequence left to right is 5′ to 3′ with respect to the *URA3* coding sequence.

Table 2. Yeast Strains

Strain Name	Relevant Genotype
AMY125	α *ade5-1 his7-2 leu2-3,112 trp1-289 ura3-52*
MS71	*LEU2* derivative of AMY125
SMH1	Isogenic with MS71 with *msh2::Tn10Luk*
GCY140	Isogenic with MS71 with *msh3::hisG*
AMY101	Isogenic with AMY125 with *pms1::LEU2*
SMH3	MS71 with pTA1, *MSH2 MSH3 PMS1*
SMH4	MS71 with pTA2, *MSH2 MSH3 PMS1*
SMH5	MS71 with pTA3, *MSH2 MSH3 PMS1*
SMH6	MS71 with pTA4, *MSH2 MSH3 PMS1*
SMH7	SMH1 with pTA1, *msh2 MSH3 PMS1*
SMH8	SMH1 with pTA2, *msh2 MSH3 PMS1*
SMH9	SMH1 with pTA3, *msh2 MSH3 PMS1*
SMH10	SMH1 with pTA4, *msh2 MSH3 PMS1*
SMH11	AMY101 with pTA1, *MSH2 MSH3 pms1*
SMH12	AMY101 with pTA2, *MSH2 MSH3 pms1*
SMH13	AMY101 with pTA3, *MSH2 MSH3 pms1*
SMH14	AMY101 with pTA4, *MSH2 MSH3 pms1*
SMH20	GCY140 with pTA1, *MSH2 msh3 PMS1*
SMH21	GCY140 with pTA4, *MSH2 msh3 PMS1*

ation of the repetitive tract. Thus, the rates of tract instability shown in Table 3 reflect both the rate of formation of 5-FOA^R colonies and the fraction of the plasmids derived from these colonies with altered tracts (details discussed in Experimental Procedures). For strains SMH3 and SMH5, the rates calculated were maximum rates at which tract alterations occurred, since no altered tracts were detected. The data summarized in Table 3 indicate that a single variant repeat stabilizes a poly(GT) tract by at least a factor of 50. Variant AT and variant CT repeats had approximately the same stabilizing effects.

The 5-FOA^R colonies derived from SMH3, SMH4, and SMH5 that do not have plasmids with altered tracts could reflect mutations elsewhere within the *URA3* coding sequence or gene conversion events between the *URA3* gene on the plasmid and the mutant *ura3* gene on the chromosome. Since the *ura3* gene on the chromosome contained an insertion of the transposable element Ty, a conversion event in which this mutation was introduced into the plasmid could be readily detected by restriction analysis of the plasmid-borne *URA3* gene. Of plasmids derived from nine independent 5-FOA^R colonies of SMH3, none contained a Ty insertion. Consequently, it is likely that most of the 5-FOA^R derivatives of SMH3, SMH4, and SMH5 represent point mutations in the *URA3* gene rather than gene conversion events.

The Stabilizing Effect of Variant Repeats Is Dependent on the Mismatch Repair Functions *MSH2*, *PMS1*, and *MSH3*

In E. coli (Levinson and Gutman, 1987), yeast (Strand et al., 1993), and humans (reviewed by Karran and Bignami, 1994), mutations in genes involved in DNA mismatch repair greatly destabilize uninterrupted tracts of simple repetitive DNA sequences. To determine the effects of such mutations on the stability conferred by variant repeats, we transformed yeast strains containing mutations in *msh2*, *pms1*, or *msh3* with the assay plasmids described above. The *MSH2* gene encodes a homolog to the MutS protein of E. coli, and mutations in this gene result in high levels of spontaneous mutation, postmeiotic segregation, and instability of simple repetitive DNA sequences (Reenan and Kolodner, 1992a, 1992b; Strand et al., 1993). *PMS1* encodes a homolog to the MutL protein of E. coli, and mutations in this gene result in phenotypes similar to those observed in *msh2* strains (Williamson et al., 1985; Kramer et al., 1989; Strand et al., 1993). Unlike *msh2* or *pms1* strains, strains with a mutation in the *msh3* gene (a MutS homolog) have wild-type levels of spontaneous mutations in a forward mutation assay and only a small increase in the level of postmeiotic segregation (New et al., 1993). The *msh3* mutation elevates simple repeat instability about 40-fold (compared with the 100- to 200-fold elevations observed in *msh2* and *pms1* strains) and has a larger bias toward deletion of repeats than observed in *msh2* and *pms1* strains (Strand et al., 1995). *MSH2* and *MSH3* have additive effects in preventing recombination between diverged repeats (Selva et al., 1995).

We observed approximately the same rates of appearance of 5-FOA^R colonies in the *msh2* and *pms1* strains containing plasmids with no variant repeat (SMH10 and SMH14) and in *msh2* and *pms1* strains with plasmids with an interrupting repeat (strains SMH7 to SMH9 and SMH11 to SMH13). For all strains, the rates were about 1×10^{-3} to 3×10^{-3} per cell division. We sequenced plasmids isolated from about 15–20 5-FOA^R derivatives of each strain. Of 106 plasmids with a variant repeat sequenced, with one exception (a plasmid derived from strain SMH11), all had poly(GT) tracts of altered lengths. In summary, the mismatch repair mutants greatly destabilize simple repetitive sequences, and this effect is insensitive to variant repeats within the repetitive sequence. Thus, the stabilizing effects of the variant repeats require a functional DNA mismatch repair system.

Table 3. Rates of Alterations of Interrupted and Uninterrupted Poly(GT) Tracts in Wild-Type and Mismatch Repair–Deficient Yeast Strains

	Rate of Tract Alterations per Cell Division[a]		
Tract	Wild Type (Strain)	*msh2* (Strain)	*pms1* (Strain)
$(GT)_{25}G$	1×10^{-5} (SMH6)	1.4×10^{-3} (SMH10)	2×10^{-3} (SMH14)
$(GT)_{12}AT(GT)_{12}G$	$<1.7 \times 10^{-7}$ (SMH3)	9.7×10^{-4} (SMH7)	1.6×10^{-3} (SMH11)
$(GT)_{12}CT(GT)_{12}G$	$<2 \times 10^{-7}$ (SMH5)	1.4×10^{-3} (SMH9)	2.1×10^{-3} (SMH13)
$(GT)_{16}AT(GT)_{17}G$	2×10^{-7} (SMH14)	2.7×10^{-3} (SMH8)	2.5×10^{-3} (SMH12)

[a]Calculated as described in Experimental Procedures.

Table 4. Sequence Alterations of Poly(GT) Tracts in Plasmid Rescued from 5-FOA^R Derivatives of Strains with Mutations in *msh2, msh3,* or *pms1*

Strain (Mutation)	Tract	Number of Plasmids with Alterations[a]			
		5′ +2	5′ −2	3′ +2	3′ −2
SMH7 (*msh2*)	$(GT)_{12}AT(GT)_{12}G$	1	1	4	10
SMH11 (*pms1*)	$(GT)_{12}AT(GT)_{12}G$	1	2	7	9
SMH9 (*msh2*)	$(GT)_{12}AT(GT)_{12}G$	0	4	4	12
SMH13 (*pms1*)	$(GT)_{12}AT(GT)_{12}G$	2	2	4	8
SMH8 (*msh2*)	$(GT)_{16}AT(GT)_{17}G$	0	4	0	10
SMH12 (*pms1*)	$(GT)_{16}AT(GT)_{17}G$	2	3	2	13
SMH20 (*msh3*)	$(GT)_{12}AT(GT)_{12}G$	0	0	4	12
Total		6	16	25	74

[a]Plasmids were rescued from 5-FOA^R derivatives of strains with *msh2*, *msh3*, or *pms1* mutations into E. coli. Of 124 plasmids rescued, 121 had deletions or additions of one repeat; one plasmid from SMH11 and two from SMH20 had unaltered tracts. The columns represent the number of plasmids with additions or deletions of one repeat 5′ (5′ +2 or 5′ −2) or 3′ (3′ +2 or 3′ −2) of the variant base.

The sequence alterations detected in the *msh2* and *pms1* strains involved additions or deletions of single GT pairs (Table 4), as observed previously for a 33 bp poly(GT) tract (Strand et al., 1993). We found no plasmids in which the interrupting repeat was lost. An unexpected result was that the alterations were polar, preferentially occurring to one side of the interrupting repeat. The preferred side of the tandem array was 3′ with respect to the orientation of the fusion gene. The bias was 3- to 4-fold for both additions and deletions. Deletions were about 3-fold more common than insertions.

We sequenced plasmids derived from 5-FOA^R colonies of strains SMH10 (*msh2*) and SMH14 (*pms1*); these strains contained pTA4, which lacks a variant repeat. All plasmids had additions or deletions of one repeat: 14 deletions and six additions (SMH10) and 15 deletions and five additions (SMH14).

We did a more restricted analysis of the effects of *msh3* on simple repeat instability. The rate of tract alterations in SMH21 (*msh3* strain with pTA4, the plasmid lacking a variant repeat) was 2.4×10^{-4} per cell division. This rate is 24-fold elevated over that observed in the wild-type strain (SMH6). In a previous study, we found that *msh3* elevated tract instability 25- to 63-fold (Strand et al., 1995). Of the altered plasmids, four had insertions of one repeat and 17 had deletions of one repeat. The rate of tract alterations in SMH20 (*msh3* with pTA1, the plasmid with the AT variant repeat) was 1.2×10^{-4} per cell division. Thus, the wild-type *MSH3* gene, like the *MSH2* and *PMS1* genes, is required for the stabilizing effects of variant repeats. Sequenced pTA1 plasmids rescued from 5-FOA^R derivatives of SMH20 also had polar alterations in tract length (Table 4). The polarity observed in SMH20 (no alterations on the 5′ side and 16 on the 3′ side) was significantly ($p = 0.04$, Fisher's exact test) greater than that observed for the *msh2* and *pms1* strains (22 alterations on the 5″ side and 83 on the 3′ side).

Discussion

In this report, we show that in yeast, as in humans (Chong et al., 1995; Eichler et al., 1994; Hirst et al., 1994; Kunst and Warren, 1994; Snow et al., 1994) and E. coli (Bichara et al., 1995), variant repeats stabilize simple repetitive tracts. There are at least three possible explanations for this stabilization. First, it is possible that long repetitive tracts have unusual structural characteristics that predispose them to alterations; disruption of the continuity of the tracts by variant repeats causes loss of this structural characteristic (Bichara et al., 1995; Gacy et al., 1995; Kang et al., 1995). Second, following a dissociation between primer and template strands, the variant repeat may help in perfect realignment of primer and template, preventing the formation of displaced repeats. The third explanation is related to the expected configuration of mispaired bases in the helix following a primer–template dissociation. If the primer and template strands dissociate during replication of a poly(GT) tract with a variant repeat, reassociation of these strands will often result in three mismatches, one involving a displacement of a GT repeat and two involving the variant bases (Figure 3). The probability of recognition of at least one of these three mismatches may be considerably higher than the probability of recognizing a single displaced GT repeat by itself. Since excision tracts (measured in meiotic cells) in yeast are usually greater than several hundred base pairs (Detloff and Petes, 1992), an excision tract initiated from any of the mismatches would

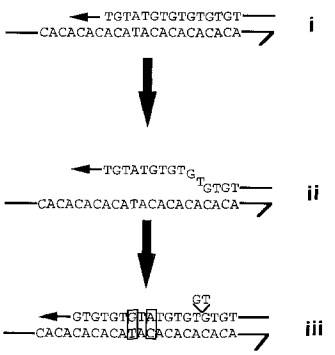

Figure 3. DNA Polymerase Slippage in a Tract with a Variant Repeat

As a consequence of the variant repeat, a slippage event resulting in an unpaired poly(GT) repeat will also result in two other closely linked mismatches (shown by boxes).

probably excise all three. The net result of these events would be a decreased rate of instability in tracts with a variant repeat.

Since we find that the stabilizing effects of the variant repeat are lost in a mismatch repair-deficient background, we favor the last explanation. Two relevant points should be added. First, an alternative interpretation of the effects of these mutations in the elimination of the stabilizing effects of the variant repeats is that the mismatch repair proteins may dissociate heteroduplexes involving three mismatches, allowing the primer and template strands another chance at realigning without mismatches. A role of the mismatch repair system in aborting heteroduplexes with multiple mismatches has been suggested in bacteria (Rayssiguier et al., 1989; Worth et al., 1994), yeast (Alani et al., 1994; Selva et al., 1995), and mammals (de Wind et al., 1995). Second, we found that a single variant AT repeat stabilizes flanking poly(GT) tracts of 32 and 34 bp (Table 3), even though we showed previously that poly(GT) tracts 33 bp in length flanked by single copy sequences are quite unstable (Henderson and Petes, 1992; Strand et al., 1993). This result is explicable in the context of the models discussed above if the dissociation between primer and template strands is extensive enough that the reannealing usually involves poly(GT) sequences on both sides of the variant repeat.

Since the stabilizing effect of the variant repeat is lost in strains with *msh2*, *msh3*, or *pms1* mutations, we suggest that all three gene products are involved in the correction of mismatches (Figure 3) or the dissociation of heteroduplexes with multiple mismatches. As described above, the effects of the *msh3* mutation on the frequencies of spontaneous mutation, repeat instability, and postmeiotic segregation are smaller than those observed for *msh2* or *pms1*. These results argue that there is more than one DNA mismatch repair complex in yeast and that these complexes have different specificities (Strand et al., 1995). The significantly increased polarity observed in the *msh3* mutant (Table 4) may represent another manifestation of this specificity.

As discussed in the Introduction, for both fragile X syndrome and SCA1, unstable mutant alleles lack the variant repeats observed in the stable wild-type alleles. It is unclear whether the variant repeats were lost during the event that led to the expansion or whether the variant repeats were lost by one mechanism (for example, DNA polymerase slippage) with a second mechanism (for example, recombination) involved in the expansion. None of the altered plasmids examined in our study had lost the variant repeat. Since our studies involve dinucleotide repeats that are usually altered by small additions or deletions, we cannot rule out the possibility that the stabilizing effects of variant repeats in preventing expansions of arrays of trinucleotide repeats involve different mechanisms.

The alterations in tract length that occur when plasmids with the variant repeat are in a mismatch repair–deficient genetic background are polar. Both additions and deletions are found 3- to 4-fold more frequently in the 3′ half of the tandem array than in the 5′ half. Sequence analysis of alleles of the *FMR1* gene indicates that small changes in the numbers of repeats also occur preferentially at one end of the array (Eichler et al., 1994; Hirst et al., 1994; Kunst and Warren, 1994; Snow et al., 1994); it was suggested that such effects reflect a polar aspect of DNA replication (Richards and Sutherland, 1994).

In our experiments, because of the location of the replication origin (*ARS* in Figure 2), the repetitive tract is presumably replicated by a replication fork that moves 5′ to 3′ with respect to the coding sequence of the fusion gene. We suggest that the slippage occurs preferentially on the lagging strand during DNA replication, as observed for some classes of deletion events in E. coli (Trinh and Sinden, 1991) and yeast (Gordenin et al., 1992), and that the Okazaki fragment responsible for replicating the poly(GT) tract is initiated outside of the tract, perhaps at a defined site (Figure 4). We further hypothesize that the Okazaki primer fragment tends to dissociate from the template before the poly(GT) tract is completely replicated, shortly after replication of the variant base. This nonrandom dissociation could reflect a tendency of poly(GT) tracts above a certain length to dissociate during replication, a distance-dependent dissociation triggered at the junction of the poly(GT) tract with single copy sequences, or a stimulation of dissociation by the variant base. The net result of these events is that mismatches would be preferentially formed to one side of the variant base. Although this model explains the data, we cannot rule out the influence of other

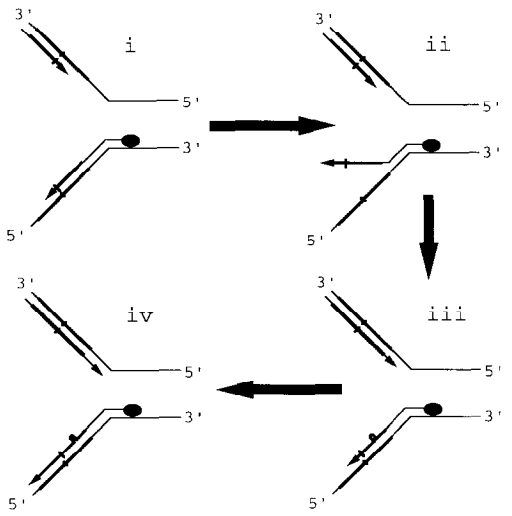

Figure 4. Model for Polarity of Slippage

Leading and lagging strand syntheses are shown on the upper and lower arms of the replication fork, respectively. The site at which the Okazaki fragment is initiated is shown by the closed oval. The poly(GT) tract is indicated by thick lines, and the variant base within the repeat is indicated by a short intersecting line within the poly(GT) tract.

(i) Synthesis on both leading and lagging strands occurs slightly beyond the variant base in the repetitive tract.

(ii) The primer and template strands on the lagging strand partially dissociate.

(iii) The primer and template strands reassociate in a misaligned configuration resulting in a displaced GT pair on the primer strand and mismatches involving the variant repeat.

(iv) Synthesis continues on both leading and lagging strands.

factors, for example, different modes of replicating poly(GT) and poly(CA) sequences, on the observed polarity.

An alternative explanation of the observed polarity is that the alterations occur by recombination (gene conversion) events initiated in one of the flanking sequences (Jeffreys et al., 1994). Since the frequency of simple repeat alterations in yeast is unaffected by the *rad52* mutation (Henderson and Petes, 1992) and is not elevated in meiotic cells (Strand et al., 1993), the polarity of alterations observed in our experiments is more likely to reflect some polar aspect of DNA polymerase slippage than recombination.

In summary, our results indicate that, in yeast, variant repeats stabilize simple repetitive sequences through a mechanism that requires a functional mismatch repair system. Alterations that occur in tracts with a variant repeat occur preferentially to one side of the interruption, probably reflecting a polar aspect of DNA polymerase slippage.

Experimental Procedures

Media and Growth Conditions

Media for the growth of yeast strains were standard (Sherman et al., 1983). Media containing 5-FOA were prepared as described by Boeke et al. (1983). Yeast strains were grown at 30°C. The E. coli strain XL1 blue was grown at 37°C in LB broth containing 100 μg/ml ampicillin when appropriate (Sambrook et al., 1989). 5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) plates for E. coli strains were prepared as described by Sambrook et al. (1989).

Plasmid Constructions

The related plasmids pTA1 to pTA4 were derived from the plasmid pSH31 (Henderson and Petes, 1992), which contains a β-galactosidase gene that has a 29 bp poly(GT) tract within in its coding sequence. Digestion of this plasmid with the enzymes BamHI and SalI released the poly(GT) tract as a small fragment. The large vector fragment was gel purified and ligated to various double-stranded oligonucleotides with BamHI- and SalI-compatible ends containing different simple sequence tracts (Table 1). These double-stranded oligonucleotides were prepared by annealing complementary single-stranded oligonucleotides (500 ng each of two complementary oligonucleotides suspended in 25 μl of 1 mM Tris–HCl [pH 7.5], 100 mM NaCl) according to the following cycles of heating and cooling: heated to 100°C for 4 min, cooled to 50°C, and incubated 1 hr; heated to 75°C and incubated for 4 min; cooled to 50°C and incubated for 30 min; and cooled slowly to room temperature. The annealed products were recovered, added to BamHI–SalI-digested pSH31, and treated with T4 DNA ligase at 12°C. Ligation products were transformed into E. coli XL1 blue and plated onto LB media containing ampicillin and X-Gal. Since the insertion of the pTA1 to pTA4 oligonucleotides should lead to an in-frame insertion in the β-galactosidase gene, only blue transformants were further analyzed. The insertions in the plasmids were sequenced, and most had the expected insertion. Plasmids with the correct oligonucleotide insertions were then treated with PvuII, deleting most of the *lacZ* gene but leaving the beginning of the coding sequence (including the oligonucleotide) intact. Into this site, we inserted a 1.5 kb HindIII fragment (cohesive ends filled in using the Klenow fragment of DNA polymerase) derived from pNKY48 (Alani and Kleckner, 1987). This fragment contains a wild-type *URA3* gene fused to a small region of the *HIS4* gene. A diagram of the structure of the resulting plasmids is shown in Figure 2.

Yeast Strains

All yeast strains were isogenic to AMY125 (except for alterations introduced by transformation) and are described in Table 2. MS71 is a *LEU2* derivative of AMY125 described by Strand et al. (1993). SMH1 was a derivative of MS128 (Strand et al., 1995) that had lost the plasmid pSH91. In MS128, the *MSH2* gene is disrupted by an insertion of Tn10LUK (Reenan and Kolodner, 1992b) that has a mutant *ura3* gene in the inserted element. In addition, the strain is Leu$^+$ because of an insertion of the plasmid CV9 (Strand et al., 1995). Strain GCY140 is a *msh3* derivative of MS71 (Strand et al., 1995). Strains AMY101 and AMY125 were provided by A. Morrison. Standard yeast transformation protocols were used in strain constructions (Becker and Guarente, 1991).

Plasmid Rescue and DNA Sequence Analysis

Plasmids were rescued from yeast using the protocol of Ward (1990) and transformed into E. coli by electroporation (Sambrook et al., 1989). Double-stranded sequencing templates were prepared according to the method of Kraft et al. (1988). Sequencing was done using a commercially available −40 primer (5′-GTTTTCCCAGTCACGAC) and the United States Biochemical Sequenase kit.

Frameshift Rate Determination

Yeast strains with the plasmids pTA1 to pTA4 are Ura$^+$ and, therefore, sensitive to 5-FOA. Alterations of the poly(GT) tract that result in a frameshift in the plasmid-borne *URA3* fusion gene can be detected by plating the cells on 5-FOA-containing plates (Henderson and Petes, 1992). To determine the rate at which 5-FOA^R derivatives arise, we streaked strains containing plasmids for single colonies on omission medium (SD complete lacking tryptophan) to force retention of the plasmid. After 3 days of growth at 30°C, individual colonies were picked and resuspended in 100 μl of sterile water. Suitable dilutions of each sample were plated on either solid SD medium lacking tryptophan (to determine cell number) or medium containing 5-FOA (to determine the number of Ura$^-$ cells). The 5-FOA-containing medium also lacked leucine and threonine to derepress the *LEU2* promoter of the *URA3* fusion gene. The 5-FOA^R colonies were counted after 3 days of growth at 30°C. For each rate measurement, the frequencies of 5-FOA^R colonies were measured for 20 colonies, and two independent rate measurements were done for each strain. Frequencies were converted to rates using the method of the median (Lea and Coulson, 1949).

When plasmids were rescued from strains SMH6, SMH7 to SMH14, and SMH20 and SMH21, all of the plasmids but three (one from SMH11 and two from SMH20) had alterations in the poly(GT) tract. Consequently, the rates of appearance of 5-FOA^R colonies are about the same as the rates of tract alterations for these strains. Of 18 plasmids derived from SMH4, only one had an altered tract. To obtain a rate of tract alterations for SMH4, we multiplied the rate of appearance of 5-FOA^R colonies by the proportion of plasmids with altered tracts (1/18). Of 18 plasmids derived from 5-FOA^R colonies of SMH3 and 15 plasmids derived from SMH5, none had altered repeats. To obtain nonzero rates of tract alterations for these strains, we multiplied the rates of appearance of 5-FOA^R colonies by 1/18 for SMH3 and 1/15 for SMH5. These rate estimates represent maximum estimates since no altered tracts were detected. The rate data are summarized in Table 3.

Acknowledgments

The research was supported by a grant from the National Institutes of Health (GM50892). We thank S. Warren, H. Zoghbi, and H. Orr for useful suggestions and R. Farber for comments on the manuscript.

Received August 24, 1995; revised October 13, 1995.

References

Alani, E., and Kleckner, N. (1987). A new type of fusion analysis applicable to many organisms: protein fusions to the *URA3* gene of yeast. Genetics *117*, 5–12.

Alani, E., Reenan, R.A.G., and Kolodner, R.D. (1994). Interaction between mismatch repair and genetic recombination in *Saccharomyces cerevisiae*. Genetics *137*, 19–39.

Becker, D.M., and Guarente, L. (1991). High-efficiency transformation of yeast by electroporation. Meth. Enzymol. *194*, 182–187.

Bichara, M., Schumacher, S., and Fuchs, R.P.P. (1995). Genetic instability within monotonous runs of CpG dinucleotides in *Escherichia coli*. Genetics *140*, 897–907.

Boeke, J.D., Lacroute, F., and Fink, G.R. (1984). A positive selection for mutants lacking orotidine-5'-phosphate decarboxylase activity in yeast: 5-fluorootic acid resistance. Mol. Gen. Genet. *197*, 345.

Chong, S.S., McCall, A.E., Cota, J., Subramony, S.H., Orr, H.T., Hughes, M.R., and Zoghbi, H.Y. (1995). Gametic and somatic tissue-specific heterogeneity of the expanded *SCA1* CAG repeat in spinocerebellar ataxia type1. Nature Genet. *10*, 344–350.

Chung, M., Ranum, L.P.W., Duvich, L.A., Servadio, A., Zoghbi, H.Y. and Orr, H.T. (1993). Evidence for a mechanism predisposing to intergenerational CAG repeat instability in spinocerebellar ataxia type I. Nature Genet. *5*, 254–258.

de Wind, N., Dekker, M., Berns, A., Radman, M., and Te Riele, H. (1995). Inactivation of the mouse *Msh2* gene results in mismatch repair deficiency, methylation tolerance, hyperrecombination, and predisposition to cancer. Cell *82*, 321–330.

Detloff, P., and Petes, T.D. (1992). Measurements of excision repair tracts formed during meiotic recombination in *Saccharomyces cerevisiae*. Mol. Cell. Biol. *12*, 1805–1814.

Eichler, E.E., Holden, J.J.A., Popvich, B.W., Reiss, A.L., Snow, K., Thibodeau, S.N., Richards, C.S., Ward, P.A., and Nelson, D.L. (1994). Length of uninterrupted CGG repeats determines instability in the *FMR1* gene. Nature Genet. *8*, 88–94.

Farber, R.A., Petes, T.P., Dominska, M., Hudgens, S.S., and Liskay, R.M. (1994). Instability of simple sequence repeats in a mammalian cell line. Hum. Mol. Genet. *3*, 253–256.

Freund, A., Bichara, M., and Fuchs, R.P. (1989). Z-DNA-forming sequences are spontaneous deletion hot spots. Proc. Natl. Acad. Sci. USA *86*, 7465–7470.

Gacy, A.M., Goellner, G., Juranic, N., Macura, S., and McMurray, C.T. (1995). Trinucleotide repeats that expand in human disease expand hairpin structures in vitro. Cell *81*, 533–540.

Gordenin, D.A., Malkova, A.L., Peterzen, A., Kulikov, V.N., Pavlov, Y.I., Perkins, E., and Resnick, M.A. (1992). Transposon Tn5 excision in yeast: influence of DNA polymerases α, δ, and ε and repair genes. Proc. Natl. Acad. Sci. USA *89*, 3785–3789.

Hamada, H., Petrino, M.G., and Kakungaga, T. (1982). A novel repeated element with Z-DNA potential is widely found in evolutionary diverse eukaryotic genomes. Proc. Natl. Acad. Sci. USA *79*, 6465–6469.

Henderson, S.T., and Petes, T.D. (1992). Instability of simple sequence DNA in *Saccharomyces cerevisiae*. Mol. Cell. Biol. *12*, 2749–2757.

Hirst, M.C., Grewal, P.K., and Davies, K.E. (1994). Precursor arrays for triplet repeat expansion at the fragile X locus. Hum. Mol. Genet. *3*, 1553–1560.

Jeffreys, A.J., Tamaki, K., MacLeod, A., Monckton, D.G., Neil, D.L., and Armour, J.A.L. (1994). Complex gene conversion events in germline mutation at human minisatellites. Nature Genet. *6*, 136–145.

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

Karran, P., and Bignami, M. (1994). DNA damage tolerance, mismatch repair and genome instability. Bioessays *16*, 833–839.

Kraft, R.J., Tardiff, K.S., Krauter, K.S., and Leinward, L.A. (1988). Using mini-prep plasmid DNA for sequencing double stranded templates with Sequenase. Biotechniques *6*, 544–545.

Kramer, B., Kramer, W., Williamson, M.S., and Fogel, S. (1989). Cloning and nucleotide sequence of DNA mismatch repair gene *PMS1* from *Saccharomyces cerevisiae*: homology of *PMS1* to prokaryotic *mutL* and *hexB*. J. Bacteriol. *171*, 5339–5346.

Kremer, E.J., Pritchard, M., Lynch, M., Yu, S., Holman, K., Baker, E., Warren, S.T., Schlessinger, D., Sutherland, G.R., and Richards, R.I. (1991). Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence $p(CCG)_n$. Science *252*, 1711–1714.

Kunst, C.B., and Warren, S.T. (1994). Cryptic and polar variation of the fragile X repeat could result in predisposing normal alleles. Cell *77*, 853–861.

Lea, D.E., and Coulson, C.A. (1949). The distribution of numbers of mutants in bacterial populations. J. Genet. *49*, 264–285.

Levinson, G., and Gutman, G.A. (1987). High frequency of short frameshifts in poly-CA/GT repeats borne by bacteriophage M13 in *Escherichia coli* K12. Nucl. Acids Res. *15*, 5323–5328.

Nancarrow, J.K., Holman, K., Mangelsdorf, M., Hori, T., Denton, M., Sutherland, G.R., and Richards, R.I. (1995). Molecular basis of $p(CCG)_n$ repeat instability at the *FRA16A* fragile site locus. Hum. Mol. Genet. *4*, 367–372.

Nelson, D.L., and Warren, S.T. (1993). Trinucleotide repeat instability: when and where? Nature Genet. *4*, 107–108.

New, L., Liu, K., and Crouse, G.F. (1993). The yeast gene *MSH3* defines a new class of eukaryotic MutS homologues. Mol. Gen. Genet. *239*, 97–108.

Orr, H.T., Chung, M., Banfi, S., Kwiatkowski, T.J., Servadio, A., Beaudet, A.L., McCall, A.E., Duvick, L.A., Ranum, L.P.W., and Zoghbi, H.Y. (1993). Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1. Nature Genet. *4*, 221–225.

Rayssiguier, C., Thaler, D.S., and Radman, M. (1989). The barrier to recombination between *Escherichia coli* and *Salmonella typhimurium* is disrupted in mismatch-repair mutants. Nature *342*, 396–401.

Reenan, R.A.G., and Kolodner, R.D. (1992a). Isolation and characterization of two *Saccharomyces cerevisiae* genes encoding homologues of the bacterial HexA and MutS mismatch repair proteins. Genetics *132*, 963–973.

Reenan, R.A.G., and Kolodner, R.D. (1992b). Characterization of insertion mutations in the *Saccharomyces cerevisiae MSH1* and *MSH2* genes: evidence for separate mitochondrial and nuclear functions. Genetics *132*, 975–985.

Richards, R.I., and Sutherland, G.R. (1994). Simple repeat DNA is not replicated simply. Nature Genet. *6*, 114–116.

Sambrook, J., Fritsch, E., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual, Second Edition (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).

Selva, E.M., New, L., Crouse, G.F., and Lahue, R.S. (1995). Mismatch correction acts as a barrier to homeologous recombination in *Saccharomyces cerevisiae*. Genetics *139*, 1175–1188.

Sherman, F., Fink, G.R., and Hicks, J.B. (1983). Methods in Yeast Genetics (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).

Smith, G.P. (1973). Unequal crossover and the evolution of mutligene families. Cold Spring Harbor Symp. Quant. Biol. *38*, 507–513.

Snow, K., Tester, D.J., Kruckeberg, K.E., Schaid, D.J., and Thibodeau, S.N. (1994). Sequence analysis of the fragile X trinucleotide repeat: implications for the origin of the fragile X mutation. Hum. Mol. Genet. *3*, 1543–1551.

Strand, M.K., Prolla, T.A., Liskay, R.M., and Petes, T.D. (1993). Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. Nature *365*, 274–276.

Strand, M.K., Early, M.C., Crouse, G.F., and Petes, T.D. (1995). Mutations in the *MSH3* gene preferentially lead to deletions within tracts of simple repetitive DNA in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA, in press.

Streisinger, G., Okada, Y., Emrich, J., Newton, J., Tsugita, A., Terzaghi, E., and Inoye, M. (1966). Frameshift mutations and the genetic code. Cold Spring Harbor Symp. Quant. Biol. *31*, 77–84.

Trinh, T.Q., and Sinden, R.R. (1991). Preferential DNA secondary structure mutagenesis in the lagging strand of replication in *E. coli*. Nature *352*, 544–547.

Walmsley, R.M., Szostak, J.W., and Petes, T.D. (1983). Is there left-handed DNA at the ends of yeast chromosomes? Nature *302*, 84–86.

Ward, A.C. (1990). Single step purification of shuttle vectors from yeast for high efficiency back-transformation into *E. coli*. Nucl. Acids Res. *18*, 5345.

Williamson, M.S., Game, J.C., and Fogel, S. (1985). Meiotic gene conversion mutants in *Saccharomyces cerevisiae*. I. Isolation and characterization of *pms1*-1 and *pms1*-2. Genetics *110*, 609–646.

Worth, L., Clark, S., Radman, M., and Modrich, P. (1994). Mismatch repair proteins MutS and MutL inhibit RecA-catalyzed strand transfer between divergent DNAs. Proc. Natl. Acad. Sci. USA *91*, 3238–3241.