Nucleic Acids Research, 1996, Vol. 24, No. 14 **2835–2840**

In vitro expansion of GGC:GCC repeats: identification of the preferred strand of expansion

Jiuping Ji[1], Nigel J. Clegg[3], Karen R. Peterson[4], Aimee L. Jackson[1], Charles D. Laird[2,4] and Lawrence A. Loeb[1,*]

[1]The Joseph Gottstein Memorial Cancer Research Laboratory, Department of Pathology and Biochemistry, [2]Departments of Zoology and Genetics, and [3]Department of Biochemistry, University of Washington, Seattle, WA 98195, USA and [4]Program in Molecular Medicine, C3-168, Fred Hutchinson Cancer Research Center, 1124 Columbia Street, Seattle, WA 98104, USA

Received January 18, 1996; Revised and Accepted June 6, 1996

ABSTRACT

The human fragile-X syndrome, a major cause of inherited mental retardation, is associated with expansion of the trinucleotide repeat GGC:GCC. Repetitive sequences in DNA are subject to slippage during catalysis by DNA polymerases. We characterized the extent of slippage of synthetic GGC:GCC repeats by various DNA polymerases: *Taq* DNA polymerase, Klenow fragment of DNA polymerase I, DNA Sequenase®, DNA polymerase-α and polymerase-β, as well as HIV reverse transcriptase. All of these enzymes were found to expand GGC:GCC repeats, with the most extensive expansion exhibited by *Taq* DNA polymerase. Starting with a template and primer, each 15 nucleotides (nt) in length, the product of one round of synthesis by *Taq* polymerase is as long as 250 nt. Sequence analysis of cloned DNA fragments expanded by *Taq* polymerase indicates that expansion involves multiple triplet additions and that it is asymmetric. The asymmetric distribution of terminal nucleotides in the expanded product is consistent with active expansion of the GCC strand and passive additions onto the GGC strand. The preferential elongation and expansion of the GCC strand was confirmed in studies utilizing longer repeats within a single-stranded M-13 template.

INTRODUCTION

Trinucleotide repeat expansions are associated with many human inherited diseases (1). Expansion of repetitive sequences may involve unequal crossing-over between the repeats, gene conversion, or misalignment of DNA strands during DNA replication or repair (2,3). In humans, expansion of a tandemly reiterated GGC:GCC sequence (also described as CGG/CCG, see ref. 4 for discussion of nomenclature), is associated with the expression of different fragile chromosomal sites: FRAXA (5,6), FRAXE (7), FRAXF (8,9), FRA11B (10) and FRA16A (11); the first two of these fragile sites are associated with mental retardation. The best-studied fragile site, FRAXA, is a cytogenetic marker for fragile-X syndrome, the most common cause of inherited mental retardation. Most fragile-X patients have an expansion and hypermethylation of a GGC repeat in the 5′-untranslated region of exon 1 of the fragile-X mental retardation gene, *FMR1* (5,12). In normal individuals the number of GGC repeats varies from 6 to 56; in unaffected carriers it varies from 43 to ~200; and in individuals with fragile-X syndrome it is >200 (13). In normal individuals, CGG repeats contain occasional single base interruptions and it appears that an initial step leading to instability for trinucleotide repeats is loss of single base interruptions in the repeat sequence (14–17).

The association of the GGC repeat with genetic disease prompted us to reinvestigate the ability of a GGC:GCC duplex to expand during *in vitro* replication. Schlotterer and Tautz (18) studied the role of DNA slippage synthesis during trinucleotide repeat expansion *in vitro* using a variety of complementary oligomers and bacterial DNA polymerases. Most trinucleotide repeats did undergo expansion, including a CAG:CTG repeat that has been implicated in several human inherited diseases (19,20). In contrast, the GGC:GCC duplex underwent little or no DNA slippage synthesis under their assay conditions (18). Here we demonstrate that this repeat can be expanded *in vitro* by eukaryotic and prokaryotic DNA polymerases as well as by reverse transcriptase, and we infer from cloning the product of the reaction that there is a striking asymmetry in the propensity of the two complementary strands to expand.

EXPERIMENTAL PROCEDURES

Chemicals and enzymes

Recombinant *Taq* DNA polymerase was purchased from Boehringer Mannheim Corp. The Klenow fragment of *Escherichia coli* DNA polymerase I [Pol I(kf)], human immunodeficiency virus reverse transcriptase (HIV-RT) and calf thymus DNA polymerase-α (Pol-α) were purified as previously described (21,22). Recombinant mouse DNA polymerase-β (Pol-β) was a gift from Dr S. Wilson (University of Texas at Galveston). DNA Sequenase®, a modified T7 DNA polymerase and T4 polynucleotide kinase were purchased

* To whom correspondence should be addressed

from United States Biochemicals; T4 ligase was purchased from Gibco-BRL. One unit of Pol-α or HIV-RT catalyzes the incorporation of 1 nmol of [α-^{32}P]dTTP into an acid-insoluble product in 30 min at 37°C on activated calf thymus DNA (22). One unit of the other DNA polymerases is the amount of enzyme required to incorporate 1 nmol of [^{3}H]TTP into an acid-insoluble product in 10 min at 37°C using poly dA as a template. Deoxy- and dideoxy-ribonucleoside triphosphates were purchased from Sigma. DuPont-NEN supplied [γ-^{32}P]ATP used for end-labeling and [α-^{32}P]dCTP for DNA synthetic reactions. In the case of oligonucleotides with repetitive sequences, the subscript after the parenthesis indicates the number of repeated units.

DNA synthesis

Oligomer (20 pmol) was labeled at the 5′-end with [γ-^{32}P]ATP and mixed with 20 pmol of a complementary unlabeled oligomer [or single-stranded M13 DNA containing FMR-1 DNA (see below)] in 10 µl of 40 mM Tris–HCl pH 8.0, 50 mM NaCl, 20 mM $MgCl_2$. The primers were annealed by heating the mixture to 95°C for 10 min, followed by incubation at 70°C for 10 min, and by cooling to room temperature for 30 min. These conditions are not sufficient for the formation of quadruplex structures (23). The standard reactions in a final volume of 15 µl contained: 1 pmol of each of the annealed complementary oligonucleotides, 100 µM each of dCTP and dGTP, 1 U each of the indicated DNA polymerase and buffer. Incubation was at 37°C for 90 min. The buffer for *Taq* polymerase, Pol I(kf), DNA Sequenase and HIV-RT contained 50 mM Tris–HCl pH 8.0, 10 mM NaCl and 4 mM $MgCl_2$. The buffer for DNA polymerase-α contained 20 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (pH 7.8), 1.0 mM dithiothreitol and 3 mM $MgCl_2$, and for DNA polymerase-β the buffer contained 20 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (pH 8.4), 50 mM NaCl and 10 mM $MgCl_2$. The elongated oligonucleotides were desalted by centrifugation through a Sephadex G-50 mini-column, electrophoresed on 8% polyacrylamide/7.0 M urea gels, and quantified by scanning the gels with a PhosphorImager (Molecular Dynamics).

Cloning of the extended and expanded repeats

The expansion reaction catalyzed by *Taq* polymerase was incubated for 60 min at 37°C, followed by an additional 30 min incubation with 1 mM dATP. The reaction mixture was extracted with phenol/chloroform, and DNA was purified by centrifugation through a Sephadex G-50 mini-column prior to cloning. The DNA fragments were ligated into a linearized TA vector (containing a 3′-overhanging dT residue), pCRTM II (Invitrogen) by incubation overnight at 16°C in 10 µl of 50 mM Tris–HCl at pH 7.6, 10 mM $MgCl_2$, 1 mM ATP, 1 mM DTT, 5% polyethylene glycol-8000 and 1 U T4 ligase; the product was transfected into *E.coli* CSH50 [Δ(*pro-lac*) *thi ara strA/F′*(*proAB IacIa-zΔM15*)] using a ‘Gene-pulser’ electroporator (BioRad). Recombinant DNA from individual transformed clones was isolated and sequenced with a universal primer as described (24,25).

Converting FMR-1 fragment into single-stranded DNA template

A human fragile-X trinucleotide DNA fragment from a normal individual was PCR-amplified from genomic DNA using the method of Fu *et al.* (13). In order to clone repetitive sequences in either orientation, the FMR-1 fragment was amplified with two sets of PCR primers. The FMR-1 fragment was then inserted into an M13 vector to obtain a single-stranded template. Both sets of PCR primers contain restriction sites for *Eco*RI or *Bam*HI at a 5′ termini, as well as the FMR sequence complementary to the region from –20 to –1 upstream or from +19 to +38 downstream of the trinucleotide repeats: FMRJJ1, 5′-CGG GAA TTC CAG GGG GCG TGC GGC AGC G-3′; FMRJJ3, 5′-CGC CGG ATC CAG GGG GCG TGC GGC AGC G-3′; FMRJJ2, 5′-CCG GGA TCC CGC CCC CGA GAG GTG GGC TG-3′; and FMRJJ4, 5′-CCG GAA TTC CGC CCC CGA GAG GTG GGC GTG-3′. The PCR-amplified FMR-1 fragments and M13mp19 vector DNA were, respectively, incubated in a 50 ml reaction mixture at 37°C for 60 min, containing: 50 mM Tris–HCl pH 8.0, 10 mM $MgCl_2$, 100 mM NaCl and 10 U of both *Eco*RI and *Bam*HI. The DNA was extracted with phenol/chloroform and precipitated with ethanol. The oligonucleotide ends were removed by Centricon-30 filtration (Amicon). The FMR-1 fragment was ligated into M13 DNA and single-stranded DNA was prepared (25).

RESULTS

Slippage during DNA synthesis has been postulated to cause both frameshift mutations (2), and the expansion of homoduplexes and other short repeats (2,26). To measure slippage of the fragile X repeats, we annealed a radiolabeled $(GGC)_5$ oligomer to an unlabeled $(GCC)_5$ oligomer and incubated the duplex with different polymerases in the presence of dCTP and dGTP. The reaction products were resolved on denaturing polyacrylamide gels (Fig. 1). Each of the tested DNA polymerases extended a fraction of the GGC oligomers to a length greater than 30 nucleotides (nt), the combined length of the primer and template strands. In the absence of slippage the product would correspond to the length of the primer plus that of the template minus the number of complementary nucleotides required for the initiation of DNA synthesis by the polymerase (27). *Taq* DNA polymerase synthesized the longest fragments, with chain lengths approaching 250 nt. In the experiment reported here, the percentage of GGC primers extended to sizes >30 nt varied from 6% for Pol-α to 79% for *Taq* DNA polymerase, suggesting that the capacity for DNA expansion is a property of the individual DNA polymerases.

In order to gain insight into the mechanism for chain expansion, we cloned the product synthesized by *Taq* DNA polymerase. We took advantage of the ability of *Taq* to add single deoxyadenosine residues in a template-independent fashion onto the 3′-end of duplex molecules (28). The dA-terminated extended oligonucleotides were subsequently inserted into a vector pCRTMII containing a 3′dT-overhang. Following ligation, the recombined DNA was introduced into *E.coli*, individual clones were isolated, purified, and the repetitive segment was sequenced using an upstream primer. As shown in Figure 2, the expanded sequence was recovered in two orientations. Of the isolated clones, 11 had their trinucleotide repeats in the 5′-GCC-3′ orientation (I) and 12 had the repeats in the 5′-GGC-3′ orientation (II). Hence, no bias was observed in the orientation of the expanded repetitive sequences in the vector. The number of residues within each cloned repeat varied from 17 to 64, indicating that the polymerase extended the original 15 nt primer by as much as 49 residues. The reported inability of *E.coli* to maintain repetitive sequences longer than 150 nt could account for the lack of even longer

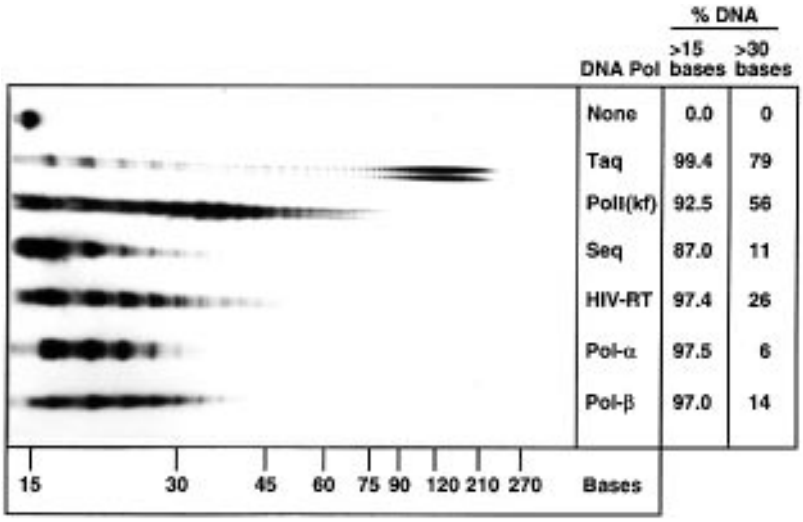

Figure 1. Expansion of trinucleotide repeats of GGC:GCC with various DNA polymerases. A ^{32}P-labeled $(GGC)_5$ oligomer was annealed with a $(GCC)_5$ oligomer and incubated with 1 U each of various DNA polymerases as described in 'Experimental Procedures'. The reaction products were resolved on an 8% denaturing polyacrylamide gel and quantified using a PhosphorImager. The extent of expansion in each reaction is estimated from the percentage of the radioactivity present in molecules extended longer than 30 bases.

inserts (29). The distribution of lengths of the cloned and sequenced products is heterogeneous (Fig. 2), as also reflected in the gel pattern of the extended chains (Fig. 1). Of the 23 clones, seven were longer than twice the length of the template and therefore must have resulted from slippage of the primer relative to the template during synthesis. The size of the longest repeats suggests that multiple slippage events must have occurred during synthesis. The uninterrupted trinucleotide sequence in most of the clones indicates that expansion involves slippage of 3 nt at a time. The only exceptions were the terminal residues (*vide infra*) and, in two of 23 clones, an internal adenine residue interrupted the trinucleotide array. In these two cases, a GCC repeat was added after the incorporated adenine residue, suggesting that *Taq* polymerase can extend beyond a mismatched 3′ terminal nucleotide.

The position of the terminal residues suggests a mechanism for expansion. Of the 23 clones that were sequenced, 16 contained non-triplet terminal additions only on the 3′-terminus of the GCC strand, whereas none contained non-triplet terminal additions on the 5′-end of the GCC repeat. In control reactions, non-triplet additions were lacking in all three sequenced clones derived from oligomers that were not extended by the polymerase (data not shown). The bias in the position of the terminal nucleotides is not due to unidirectional insertion of the expanded repeats or to selection during cloning, since nearly an equal number of clones were recovered in either orientation. Of the 11 clones in orientation I, eight contained non-triplet residues at the 3′-terminus and no non-triplet residues at the 5′-terminus of the GCC strand (Fig. 2). In contrast, eight of the 12 clones recovered in orientation II contained non-triplet residues at the 5′ terminus of the GGC strand, but there were no non-triplet residues at the 3′-terminus. A consensus sequence for all 23 clones is, therefore:

5′-$(GCC)_{5–21}$N-3′
3′-$(CGG)_{5–21}$N-5′

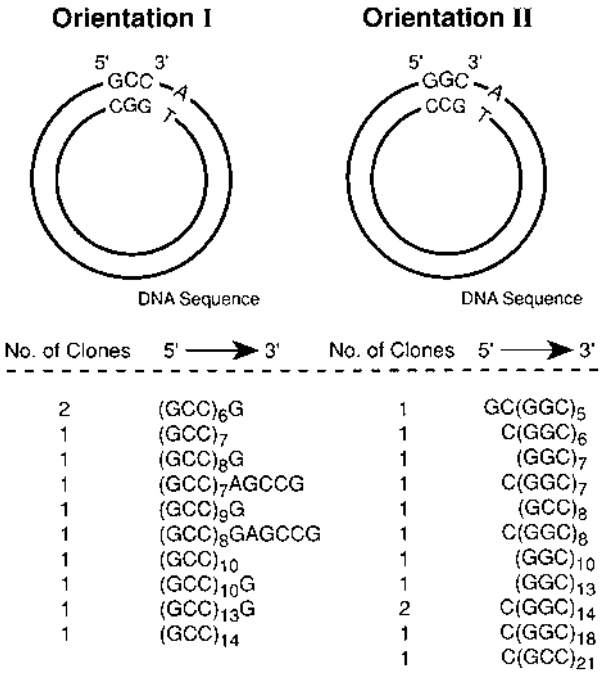

Figure 2. The nucleotide sequence of trinucleotide repeats expanded by *Taq* DNA polymerase. The expansion of the GGC:GCC repeats by *Taq* polymerase and the cloning of the expanded products is given in 'Experimental Procedures'. Based on the orientation of the repetitive segments, 23 clones were sequenced and divided into two groups. The DNA sequence of the transcribed strand is written in the 5′ to 3′ direction and the number of repeats is given as a subscript.

All known DNA polymerases add nucleotides exclusively onto the 3′-terminus. Addition of nucleotides onto the GCC strand provides a direct explanation for the variable sequences found on

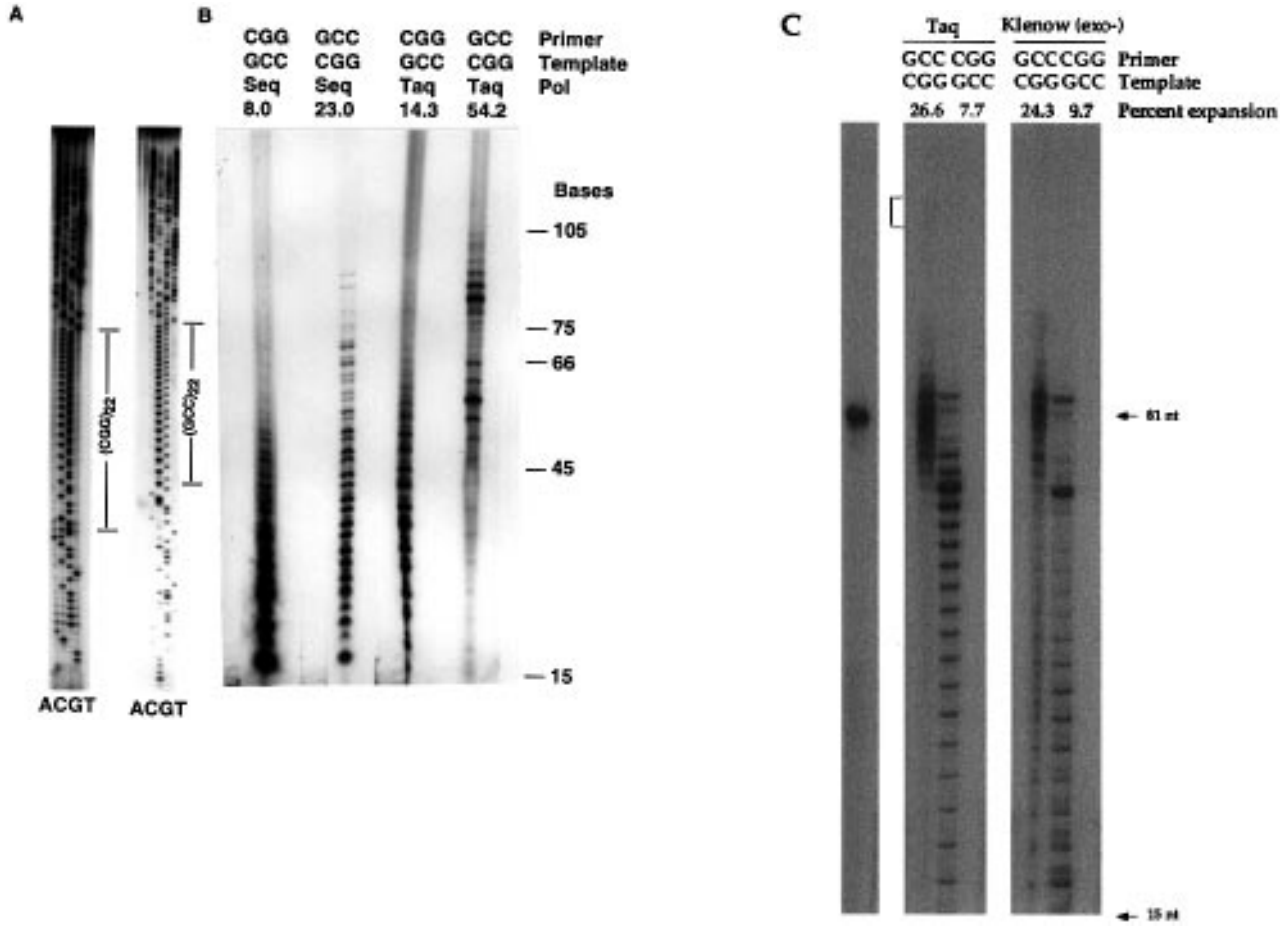

Figure 3. DNA slippage synthesis on circular single-stranded DNA template containing repetitive trinucleotide sequence. (**A**) DNA sequence of single-stranded M13 recombinant DNA containing GGC:GCC repeats. The single-stranded M13 phage containing fragile-X trinucleotide repeats of either $(GCC)_n$ or $(GGC)_n$ were identified and sequenced using the universal primer (24). Each dideoxy A, C, G, T reaction mixture was loaded onto an 8% denaturing polyacrylamide gel. The sequences shown in (A) are that of a trinucleotide repeat of 22 U in opposite orientations with arrows indicating the orientation of trinucleotide repeats. (**B**) DNA slippage synthesis on FMR-1 DNA. The single-stranded M13 DNA containing trinucleotide repeats of either $(CGG)_{22}$ or $(GCC)_{22}$ was annealed with the complementary oligomer and incubated with the indicated DNA polymerases, in the presence of dGTP and [α-^{32}P]dCTP at 37°C for 90 min. The extended DNA molecules were analyzed by electrophoresis on an 8% denaturing polyacrylamide gel and quantified by using a Phosphorimager. Percent expansion in each lane reflects the percentage of counts incorporated into extension products >66 nt. (**C**) DNA slippage synthesis on FMR-1 DNA. The single-stranded M13 DNA containing trinucleotide repeats of either $(CGG)_{22}$ or $(GCC)_{22}$ was annealed with the complementary 5′-end labeled 15mer primer and incubated with the indicated DNA polymerases, in the presence of dGTP and dCTP. The extension products were analyzed by electrophoresis on an 8% denaturing polyacrylamide gel and quantified by using a Phosphorimager, using a 5′-end labeled 81 nt oligomer as the marker for full extension (66 nt repeats plus 15 nt primer). The percent expansion in each lane reflects the percentage of counts incorporated into expansion products larger than 81 nt. The bracket indicates the presence of expansion products of 160 nt in the 5′-GCC-3′ strand, as determined by comparison to a labeled 10 bp DNA size ladder.

the 3′-terminus of the GCC strand in orientation I. In contrast, the lack of terminal additions at the 3′-terminus of the GGC strand in orientation II suggests that this strand is less subjected to slippage during catalysis. Since DNA synthesis occurs exclusively in a 5′ to 3′ direction, the 5′ sequence variability observed in orientation II clones could not be the result of nucleotide additions by *Taq* polymerase, but instead must have occurred by elongation of the vector during cloning of the expanded product.

To determine whether biased slippage also occurs on repetitive trinucleotide sequences embedded within the flanking sequences found in the FMR-1 gene, we constructed single-stranded M13 vectors containing either $(GGC)_{22}$ or $(GCC)_{22}$ inserts and verified the position of the inserts by DNA sequencing (Fig. 3A). Onto each of these constructs we annealed the corresponding complementary oligonucleotide primer $(GCC)_5$ or $(CGG)_5$. Incubation was carried out with either DNA Sequenase or *Taq* DNA polymerase in the presence of dGTP and [α-^{32}P]dCTP. In principle, the primer could anneal to any of the 22 trinucleotide units in the template, and as a result the newly synthesized molecules could be as long as 81 bases (66 plus the primer) in the absence of slippage. In control experiments, we annealed specific non-repetitive oligonucleotide primers complementary to nucleotide sequences that flank the trinucleotide repeat region and we did not observe any expanded DNA products (data not shown). In contrast, the results presented in Figure 3B indicate that products >81 nt were produced in reaction with either the $(GGC)_5$ or $(GCC)_5$ primers, suggesting slippage DNA synthesis occurs on both strands. Expansion products through slippage synthesis with the oligomer $(GCC)_5$ usually resulted in longer products than those obtained with $(GGC)_5$ as a primer, suggesting that with the

M13 template $(GCC)_5$ is also more frequently subject to slippage DNA synthesis. In a second set of experiments, we utilized ^{32}P-labeled primers. The extent of expansion was greater for the $(GCC)_5$ primer (26.6%) than the $(CGG)_5$ primer (7.7%) (Fig. 3B). Furthermore, extensive pausing at triplet repeats was observed only with CGG primed synthesis. A comparison of expansions observed in Figures 1 and 3 indicates that expansion is greater when both the primer and templates are not flanked by non-repetitive sequences.

DISCUSSION

Our results support suggestions that expansion of repetitive trinucleotide sequences in the *FMR1* gene is caused by slippage of nascent polymerized DNA strands during replication or repair (26,29). The expansion of trinucleotide repeat sequences during DNA synthesis *in vitro* is consistent with earlier studies in which DNA polymerases produced long products using homopolymer or copolymer templates (18,30). In studying the replication of triplet repeats using oligonucleotide templates of 15 residues and complementary primers of nine residues, Schlotterer and Tautz (18) observed that *E.coli* DNA polymerase I yielded products longer than 50 nt. Notable exceptions were GGC:GCC duplexes, which failed to undergo significant expansion. In contrast, by using greater amounts of DNA polymerases we find that trinucleotide GGC:GCC repeats can be expanded *in vitro* by various DNA polymerases.

Our studies suggest that slippage of the primer relative to the template can drive an expansion process, and that with expansion of GGC:GCC repeats this slippage occurs in multiples of three nucleotides. A model based on our results indicates that slippage expansion would involve the looping out of a trinucleotide repeat or its multiple (Fig. 4A). This looping out would result in a repositioning of the primer strand with respect to the template (Fig. 4B). As a result, the primer strand is lengthened by the size of the looped out segment. From these results, we cannot determine the template site for continued synthesis, but the absence of inversions in the sequenced clones argues against dissociation of the polymerase. Interruptions in repeats, such as AGG trinucleotides found in the normal population, are predicted to significantly impede slippage (15–17). Indeed, premutation genotypes that lack the AGG punctuations have been shown to be prone to further expansion (15,17,31–33).

The sequence of the cloned expanded products (Fig. 2) suggests that the mechanism for expansion of the trinucleotide repeats is strand-biased in that slippage occurs more frequently at the 3′-terminus of the GCC strand. Of the 23 cloned *in vitro* expanded repeats that we sequenced, 16 contained terminal additions; all of these are compatable with slippage of the GCC strand but not with slippage of the GGC strand. Our finding that recombinant clones are recovered in nearly equal numbers argues against the possibility that this bias resulted from a cloning artifact. The simplest explanation for the biased extension of the GCC strand is, therefore, that it occured in the course of DNA synthesis. A bias in slippage of repetitive sequences is not without precedent. In studies on CTG repeats, Kang *et al.* (33) observed that both expansion and deletion of repetitive sequences were orientation-dependent during replication in *E.coli*. Furthermore, our designation that the GCC strand is more likely to expand is in agreement with predictions of Chen *et al.* (34), based on physical characterization of oligomers similar to those used here.

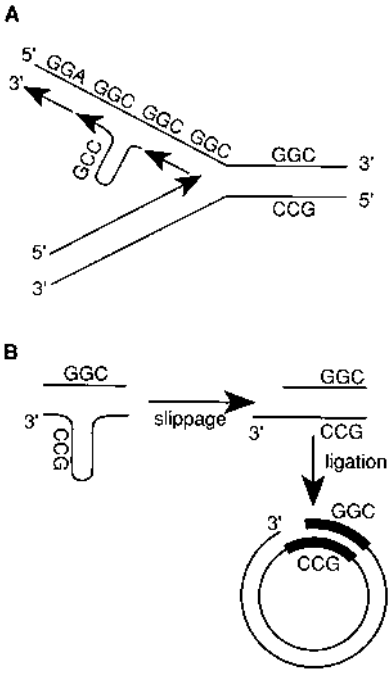

Figure 4. Model for slippage synthesis of trinucleotide GCC repeats. (**A**) Our preferred interpretation of the sequence data in Figures 1 and 3 is that the GCC strand is the predominant strand forming hairpins consisting of one or more repeats, and leading to slippage replication. The 5′ end of the GGC/GCC repeat, oriented relative to the transcription polarity of the *FMR1* gene, contains interruptions in the repeat, here denoted as GGA repeats in the GGC strand. Such interruptions were not used in our experiments, but are included in the diagram to indicate polarity. In principle, hairpins could be formed on either the lagging (upper) or leading (lower) strand. Our inference that the newly synthesized GCC strand is the one more likely to be actively expanded suggests that determining the *in vivo* direction of DNA replication through the *FMR1* gene would help distinguish the roles of leading and lagging strand synthesis in slippage events. (**B**) The slippage of the GCC strand could result in an overhang at the 3′-terminus. The gap on the GGC strand could be filled in by elongation during cloning of the insert.

The formation of an unusual secondary structure by repetitive DNA could stall DNA synthesis and account for the strand bias we observe in copying GCC:GGC repeats by *Taq* DNA polymerase. Non B-form DNA conformations, formed by both GGC and GCC repeats, including hairpins (27,32,34–37) triplex (38) and quadruplex structures (23), could stall the progression of DNA synthesis. Stable hairpins and intrastrand tetraplexes formed on the GGC strand (35,38) could cause pausing by the polymerase and looping out of the GCC strand. Pausing by DNA polymerases at regions of secondary structure has been shown to result in increasing misincorporations (39) and could also result in increased DNA slippage. This scenario would predict that sequences surrounding the fragile-X site would contain a high frequency of mutations, particularly single-base substitutions that result from errors by DNA polymerases.

The expansion of GCC sequences that we observe with purified DNA polymerases could be the underlying mechanism for the pathogenesis of fragile-X syndrome. Since expansion is observed with purifed DNA polymerases, it is not obligatory to postulate a requirement for other proteins that comprise the DNA replication apparatus, or for proteins that bind specifically to the repetitive GGC:GCC trinucleotide repeats. However, the much

larger expansion of hundreds of nucleotides that characterizes the fragile-X syndrome could involve either multiple slippage events (33) or sister chromatid exhanges, and could be facilitated by specific binding proteins. An additional possibility concerns the significant hypermethylation of the CpG dinucleotides within the GGC repeat and flanking DNA of affected individuals (35). This hypermethylation may facilitate aberrant replication of *FMR1* (40,41).

The strand bias that we observe may serve in determining whether expansion occurs on the lagging (Fig. 4) or leading strand during DNA replication. It has been noted that the size of triplet repeats above which instability occurs is similar to the size of the Okazaki fragments on the lagging strand (15,16). Our *in vitro* results would be consistent with involvement of the lagging strand if the direction of DNA replication *in vivo* were from the 5′ end of the *FMR1* gene.

ACKNOWLEDGEMENTS

We thank Alan Herr for his excellent technical assistance in DNA sequencing. This work was supported by an Outstanding Investigator Grant from the NIH (R35-CA-39903), and NIH grant R01 GM19179. N.J.C. was a Guild Fellow of the Fred Hutchinson Cancer Research Center.

REFERENCES

1 Willems, P. J. (1994) *Nature Genet.* **8**, 213–215.
2 Streisinger, G., Okada, Y., Emrich, J., *et al.* (1966) *Cold Spring Harbor Symp. Quant. Biol.* **31**, 77–84.
3 Levinson, G. and Gutman, G. A. (1987) *Nucleic Acids Res.* **15**, 5323–5337.
4 Han, J., Hsu, C., Zhu, Z., Longshore, J. W. and Finley, W. H. (1994) *Nucleic Acids Res.* **22**, 1735–1740.
5 Verkerk, A. J. M. H., Pieretti, M., Sutcliffe, J. S., *et al.* (1991) *Cell* **65**, 904–914.
6 Kremer, E. J., Pritchard, M., Lynch, M., *et al.* (1991) *Science* **252**, 1711–1714.
7 Knight, S. J. L., Flannery, V. A., Hirst, M. C., *et al*. (1993) *Cell* **74**, 127–134.
8 Parrish, J. E., *et al* (1994) *Nature Genet.* **8**, 229–235.
9 Jones, C., Penny, L., Mattina, T., Yu, S., Baker, E., Voullaire, L., Langdon, W. Y., Southerland, G. R., Richards, R. I. and Tunnacliffe, A. (1995) *Nature* **376**, 145–149.
10 Richie, R. J. (1994) *Hum. Mol. Genet.* **3**, 2115–2121.
11 Nancarrow, J. K., Kremer, E., Holman, K., *et al.* (1994) *Science* **264**, 1938–1941.
12 Oberlé, I., Rousseau, F., Heitz, D., *et al.* (1991) *Science* **252**, 1097–1102.
13 Fu, Y. -H., Kuhl, D. P. A., Pizzuti, A., *et al.* (1991) *Cell* **67**, 1047–1058.
14 Chung, M., Ranum, L. P. W., Duvick, L. A., Servadio, A., Zoghbi, H. Y. and Orr, H. T. (1993) *Nature Genet.* **5**, 254–258.
15 Kunst, C. B. and Warren, S. T. (1994) *Cell* **77**, 853–861.
16 Eichler, E. E., Holden, J. J. A., Popovich, B. W., *et al.* (1994) *Nature Genet.* **8**, 88–94.
17 Hirst, M. C., Grewal, P. K. and Davies, K. E. (1994) *Hum. Mol. Genet.* **3**, 1553–1560.
18 Schlotterer, C. and Tautz, D. (1992) *Nucleic Acids Res.* **20**, 211–215.
19 The Huntington's disease collaborative research group (1993) *Cell* **72**, 971–983.
20 Harley, H. G., Brook, J. D., Rundle, S. A., *et al.* (1992) *Nature* **355**, 545–546.
21 Roberts, J. D., Preston, B. D., Johnston, L. A., Soni, A., Loeb, L. A. and Kunkel, T. A. (1994) *Mol. Cell. Biol.* **9**, 469–475.
22 Perrino, F. W., Preston, B. D., Sandell, L. L. and Loeb, L. A. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 8343–8347.
23 Fry, M. and Loeb, L.A. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 4950–4954.
24 Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor, Cold Spring Harbor Laboratory Press, NY.
25 Ji, J. P. and Loeb, L. A. (1992) *Biochemistry* **31**, 954–958.
26 Kunkel, T.A. (1990) *J. Biol. Chem.* **237**, 18251–18258.
27 Hayes, R. N., Hansbury, E., Mitchell, V. E., Ratliff, R. L., Smith, D. A. and Williams, D. L. (1971) *J. Biol. Chem.* **246**, 3631–3638.
28 Patel, P.H. and Preston, B.D. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 549–553.
29 Sinden, R.R. and Well, R.D. (1993) *Genome Anal.* **7**, 107–139.
30 Schachman, H. K., Adler, J., Radding, C. M., Lehman, I. R. and Kornberg, A. (1960) *J. Biol. Chem.* **235**, 3242–3249.
31 Snow, K., Tester, D., Kruckerberg, K., Schaid, D., and Thibodeau, S., (1994) *Hum. Mol. Genet.* **3**, 1543–1551.
32 Gacy, A. M., Goellner, G., Juranic, N. Macura, S. and McMurray, C. T. (1995) *Cell* **81**, 533–540.
33 Kang, S., Jaworski, A., Ohshima, K. and Wells, R.D. (1995) *Nature Genet.* **10**, 213–218.
34 Chen, X., Mariappan, S. V. S., Catasti, P., Ratliff, R., Moyzis, R. K., Laayoun, A., Smith, S.S., Bradbury, E. M., and Gupta, G. (1995) *Proc. Natl Acad. Sci. USA* **92**, 5199–5203.
35 Hansen, R. S., Canfield, T. K., Lamb, M. M., Gartler, S. M. and Laird, C. D. (1993) *Cell* **73**, 1403–1409.
36 Smith, S.S., Laayoun, A., Lingeman, R.G., Baker, D.J. and Riley, J. (1994) *J. Mol. Biol.* **243**, 143–151.
37 Usdin, K. and Woodford, K. J. (1995) *Nucleic Acid Res.* **23**, 4202–4209.
38 Kohwi, Y. and Kohwi-Shigematsu, T. (1991) *Genes Dev.* **5**, 2547–2554.
39 Fry, M. and Loeb, L.A. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 763–767.
40 Hansen, R. S., Gartler, S. M., Scott, C. R., Chen, S.-H. and Laird, C. D. (1992) *Hum. Mol. Genet.* **1**, 571–578.
41 Laird, C., Jaffe, E., Karpen, G., Lamb, M. and Nelson, R. (1987) *Trends Genet.* **3**, 274–281.