Mechanism of in Vitro Expansion of Long DNA Repeats: Effect of Temperature, Repeat Length, Repeat Sequence, and DNA Polymerases †

ARTICLE *in* BIOCHEMISTRY · FEBRUARY 2002

Impact Factor: 3.02 · DOI: 10.1021/bi0110950 · Source: PubMed

Mechanism of in Vitro Expansion of Long DNA Repeats: Effect of Temperature, Repeat Length, Repeat Sequence, and DNA Polymerases†

Wirote Tuntiwechapikul and Miguel Salazar*

Division of Medicinal Chemistry, College of Pharmacy, and Institute for Cellular and Molecular Biology, The University of Texas at Austin, Austin, Texas 78712

Received May 29, 2001

ABSTRACT: Studies of sequence repeat expansions from duplexes consisting of DNA repeat sequences greater than three bases are currently lacking. These studies are needed in order to gain a better understanding of DNA expansions in general and as a first step in understanding expansions of longer sequence repeats that have been implicated in human diseases. We have undertaken an in vitro study of tetranucleotide, hexanucleotide, and octanucleotide repeat expansions from short DNA duplexes using Taq DNA polymerase. Expansions of hexanucleotide repeats were also studied with the Klenow fragment of DNA polymerase I and with T4 DNA polymerase. Studies with Taq DNA polymerase show that expansions occur more readily as the length of the repeat sequence decreases but are generally more efficient at reaction temperatures closer to the melting point of the starting duplex. A mechanism for the observed expansions with Taq DNA polymerase is proposed that does not invoke strand slippage or DNA structure. Studies at 37 °C with Klenow pol I and T4 DNA polymerase indicate that the template-switching and/or strand-displacement activities of the polymerases used can play a major role in the apparent in vitro expansions of short repetitive DNA duplexes.

Repetitive DNA sequences are ubiquitous in nature. The best-known examples of eukaryotic repetitive DNA sequences are the trinucleotide repeats CCG, GAA, and CTG, whose expansion is believed to be the cause of various human disorders (*1*). Examples of other well-known repetitive DNA sequences consisting of longer sequence repeats include the pentanucleotide repeat $(GGAAT)_n$ found in human centromeric DNA (*2*, *3*), the repeat sequences found at the chromosome ends (the telomeres) in various organisms (*4*), the 14 base pair repeat $[ACAGGGGT(G/C)(T/C)GGGG]_n$ located in the insulin-linked polymorphic region (*5*), and the 12 base pair repeat $(GTACGGGACCGA)_n$ found in centromeric DNA of *Drosophila melanogaster* (*6*, *7*).

Repetitive DNA sequences are known to elongate in vivo through expansion of the repeat number. Trinucleotide repeats and the mechanism(s) by which they expand or elongate have been the subject of intense investigation following the discovery that expansion of the CCG triplet is the source of fragile X syndrome (*8*). Since then, mechanistic and structural in vitro studies have yielded some clues regarding the possible cause(s) of trinucleotide repeat expansions in vivo (*9*). However, despite extensive research, the mechanism(s) of repetitive DNA expansions in vivo remain largely unknown and the only well-understood mechanism for expansion/elongation of repetitive DNA sequences in vivo is that for telomeric DNA, which is elongated by telomerase, a ribonucleoprotein enzyme complex with reverse transcriptase activity (*10*).

One of the most widely accepted mechanisms for triplet repeat expansions in vivo invokes strand slippage, whereby the 3′-end of one strand dissociates and then reassociates upstream and is subsequently extended by DNA polymerase. A number of in vitro studies that address the possibility of this mechanism as the source for expansion of triplet repeats have been carried out with short duplexes and various polymerases (*11*−*15*). Although these studies have yielded some clues regarding triplet repeat expansion mechanisms, in most cases they failed to take into account the thermal stability of the starting duplex as well as the strand-displacement and/or template-switching activities of the polymerases used. As will be shown below, these are important parameters that should be taken into account in studies of this kind.

Obviously, more in vitro (or in vivo) studies will be necessary to unambiguously determine the mechanisms responsible for trinucleotide repeat expansions. However, studies on expansions of longer sequence repeats may yield a better understanding of repetitive DNA expansions in general, particularly triplet repeat expansions, and might also provide some clues regarding the in vivo expansions of longer sequence repeats, particularly those that have been implicated in human diseases (*16*). Here we show that repeats of up to eight bases can be efficiently expanded regardless of the repeat sequence. However, expansions are more efficient at temperatures closer to the melting point of the starting duplex. Expansion experiments with Klenow fragment and duplexes having longer starting lengths (>100 base pairs) indicate that expansion may occur more readily at 37 °C as the length of the starting duplex increases. However, experiments with Klenow fragment and T4 DNA polymerase and duplexes containing single-strand overhangs indicate that

† This work was supported by grants from the NIH (Grant CA77000), The Robert A. Welch Foundation (Grant F-1409), and the William Orr Dingwall Foundation.

* Corresponding author: Tel 512-471-4551; fax 512-232-2606; e-mail m-salazar@mail.utexas.edu.

10.1021/bi0110950 CCC: $22.00
Published on Web 12/22/2001

Table 1: DNA Duplexes Used in This Study

name	sequence	melting point (°C)
TAG_2	5′-$(TAGG)_{10}$-3′•5′-$(CCTA)_{10}$-3′	76.1
T_2AG_3	5′-GG$(TTAGGG)_6$TT-3′•5′-AA$(CCCTAA)_6$CC-3′	76.0
$T_2A_2G_4$	5′-$(TTAAGGGG)_5$-3′•5′-$(CCCCTTAA)_5$-3′	76.0
$(T_2AG_3)_5$	5′-$(TTAGGG)_5$-3′•5′-$(CCCTAA)_5$-3′	73.0
$(T_2G_2AG)_5$	5′-$(TTGGAG)_5$-3′•5′-$(CTCCAA)_5$-3′	74.0
$(TGTGAG)_5$	5′-$(TGTGAG)_5$-3′•5′-$(CTCACA)_5$-3′	75.0
$(ACGTCA)_5$	5′-$(ACGTCA)_5$-3′•5′-$(TGACGT)_5$-3′	78.0
T_2G_4	5′-$(TTGGGG)_5$-3′•5′-$(CCCCAA)_5$-3′	nd[a]
CG54	5′-$(TTAGGG)_9$-3′•5′-$(CCCTAA)_9$-3′	nd
F-0	5′-$(TTAGGG)_3$GCAGGATATCTGGATCCA$(TTAGGG)_3$-3′• 5′-$(CCCTAA)_3$TGGATCCAGATATCCTGC$(CCCTAA)_3$-3′	nd
F-1	5′-$(TTAGGG)_2$GCAGGATATCTGGATCCA$(TTAGGG)_4$-3′• 5′-$(CCCTAA)_3$TGGATCCAGATATCCTGC$(CCCTAA)_3$-3′	nd
F-2	5′-(TTAGGG)GCAGGATATCTGGATCCA$(TTAGGG)_5$-3′• 5′-$(CCCTAA)_3$TGGATCCAGATATCCTGC$(CCCTAA)_3$-3′	nd
F-3	5′-(TTAGGG)GCAGGATATCTGGATCCA$(TTAGGG)_5$-3′• 5′-$(CCCTAA)_2$TGGATCCAGATATCCTGC$(CCCTAA)_4$-3′	nd

[a] Not determined.

the apparent expansions of longer duplexes at 37 °C could arise from the strand-displacement and template-switching activities of the enzyme used. We included in this study duplexes consisting of the human and *Tetrahymena* telomeric repeat sequences, T_2AG_3 and T_2G_4, respectively, and show that they can be expanded in vitro. We propose a mechanism for expansion of these and other duplexes studied here that does not invoke strand slippage.

MATERIALS AND METHODS

Preparation of Oligonucleotides. All oligonucleotides were synthesized on a PerSeptive Biosystems Expedite 8909 DNA synthesizer and purified by denaturing polyacrylamide gel electrophoresis. DNA concentrations were determined by UV absorption at 260 nm. The oligonucleotides were diluted into 50 μM stock solutions in 1/10 TE[1] buffer, pH 8.0, prior to use in the polymerase reactions. Labeled duplexes were prepared by first labeling one strand with [γ-^{32}P]ATP by use of T4 polynucleotide kinase. After enzyme deactivation, the labeled strand was purified by elution through a Bio-Rad Bio-Spin column and mixed with 2 molar equiv of nonlabeled complementary strand in 1/10 TE buffer, pH 8.0. The mixture was heated to 95 °C and then slowly cooled to room temperature before purification by 15% nondenaturing PAGE. The DNA duplexes used in this study are listed in Table 1.

Polymerase Reactions. All the reactions were performed in an Eppendorf Mastercycler personal thermocycler. In the reactions carried out at high temperature (60−85 °C), the 50 μL reaction mixtures contained DNA duplex, 200 μM dNTPs, and 2.5 units of Taq DNA polymerase (Promega) in a reaction buffer consisting of 10 mM Tris-HCl (pH 9.0), 50 mM KCl, 1.5 mM $MgCl_2$, and 0.1% Triton X-100. The samples were heated to 95 °C for 5 min to denature any preformed structure and then incubated at the indicated temperature and time. In the reactions carried out at 37 °C with Klenow pol I, the 50 μL reaction mixtures contained DNA duplex, 200 μM dNTPs, and 2 units of Klenow pol I (New England Biolabs) in a reaction buffer consisting of 10 mM Tris-HCl (pH 7.5), 50 mM KCl, 5 mM $MgCl_2$, and 7.5 mM DTT. In the reactions carried out at 37 °C with T4 DNA polymerase, the 50 μL reaction mixtures contained DNA duplex, 100 μM dNTPs, and 2 units of T4 DNA polymerase (New England Biolabs) in a reaction buffer consisting of 10 mM Tris-HCl (pH 7.9), 50 mM KCl, 10 mM $MgCl_2$, and 1 mM DTT. For the experiments with unlabeled DNA, the duplex concentration was 50 μM. The reaction products were separated by 1.5% agarose gel electrophoresis and visualized by ethidium bromide staining. For the experiments with ^{32}P-labeled DNA, 40 kcpm of the indicated purified duplex was mixed with 10 μM of unlabeled duplex of the same sequence. The reaction was stopped by adding 1 μL of 0.5 M EDTA. The products were then purified with phenol/chloroform, precipitated with ethanol, and analyzed either by 4% nondenaturing PAGE or by 6% denaturing PAGE. The images were captured and visualized by PhosphorImager (Molecular Dynamics model 445 S1).

Thermal Denaturation Studies. Oligonucleotide duplexes were prepared from purified oligonucleotide single strands in a buffer containing 50 mM Tris-HCl (pH 8.0) and 100 mM NaCl. Each sample was heated to 95 °C for 5 min and then slowly cooled to room temperature. The UV thermal denaturation studies were conducted on a Cary 300 spectrophotometer equipped with a Peltier 12-cell holder. The melting profiles were obtained by monitoring the absorbance at 260 nm while the sample temperature was increased from 55 to 85 °C at a rate of 1 °C/min. Melting points were determined from the first derivative of the absorbance vs temperature curve and are accurate to within 1 °C.

Preparation of DNAs of Various Lengths for the Length-Dependent Expansion Experiments. The DNA duplexes were first expanded with Taq DNA polymerase at the optimal temperature (75 °C for T_2AG_3, 80 °C for T_2G_4) for 2 h. The products were then separated by 1.5% agarose gel electrophoresis and visualized by ethidium bromide staining. Gel fragments containing DNAs of various lengths were cut in increments of approximately 100 base pairs, with the 100 base pair DNA ladder as a guide. The DNAs were then

[1] Abbreviations: TE, tris(hydroxymethyl)aminomethane (Tris)−ethylenediaminetetraacetic acid (EDTA) buffer; PAGE, polyacrylamide gel electrophoresis; dNTPs, deoxynucleoside triphosphates; DTT, dithiothreitol; kcpm, kilocounts per minute.

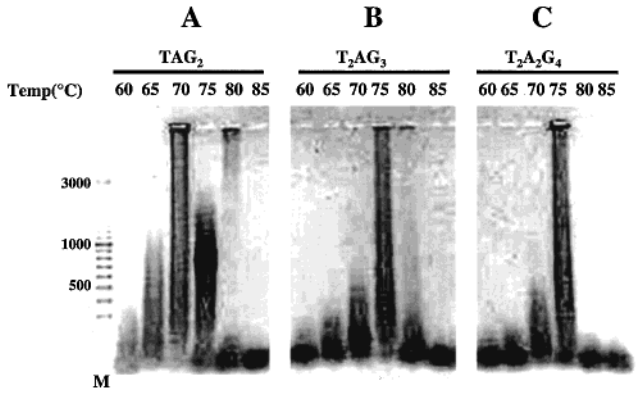

FIGURE 1: Temperature-dependent expansion with Taq DNA polymerase of 40 base pair duplexes consisting of (A) tetranucleotide (TAG_2), (B) hexanucleotide (T_2AG_3), and (C) octanucleotide ($T_2A_2G_4$) repeats. The reactions were incubated at the indicated temperatures for 2 h, and the products were analyzed by 1.5% agarose gel electrophoresis and ethidium bromide staining. Lane M represents a 100 base pair DNA ladder.

electroeluted from the gel fragments, dialyzed extensively in purified water, dried, and redissolved in DNA polymerase I reaction mixture as above. Each DNA sample was then separated into two tubes (with and without DNA polymerase I Klenow fragment) and incubated at 37 °C for 2 h. The reaction products were then analyzed by 1.5% agarose gel electrophoresis and ethidium bromide staining.

RESULTS

Figure 1 shows the agarose gel electrophoresis results for the expansion of the tetranucleotide, hexanucleotide, and octanucleotide repeats from duplexes with a starting length of 40 base pairs. The melting points of these three starting duplexes are all around 76 °C (Table 1), as determined by UV spectroscopy under DNA and salt concentrations similar to those used in the polymerase reactions. Clearly, in all cases the expansions are more efficient and the products are longer at temperatures that are closer to the melting point of the starting duplex. Only the duplex consisting of tetranucleotide repeats (TAG_2) showed significant expansion (longer than that expected from extension of complementary strands annealed through one terminal repeat) over a wider range of temperatures and an optimum expansion temperature (ca. 70 °C) somewhat lower than that for the T_2AG_3 and the $T_2A_2G_4$ duplexes. Denaturing polyacrylamide gel electrophoresis of the temperature-dependent polymerase reaction products for these duplexes using ^{32}P-labeled DNA (Figure 2) showed that in all cases the G-rich DNA strand is incrementally increased in length by discrete units corresponding to the length of the repeat. At temperatures closer to the melting point of the starting oligonucleotide, the reaction products increase in length beyond the resolution limit of the gel. Similar results were obtained for the C-rich strands (see Supporting Information). Only the TAG_2 duplex gave rise to longer products when incubated with Klenow pol I at 37 °C. However, as will be seen below, this apparent expansion is likely a result of the combined strand-displacement and template-switching activities of Klenow pol I.

Studies with triplet repeat sequences indicate that strand slippage may be mediated by DNA structures, since many of these triplet repeat sequences are known to form hairpins, G-quadruplexes, and triplexes in vitro (*1*). To rule out the possibility that G-quadruplex structures might be involved in the observed expansions of G-rich sequences, we examined the effect of the repeat sequence under the same reaction conditions as those in Figures 1 and 2. Figure 3 shows the expansion with Taq DNA polymerase of duplexes consisting of the six-base repeats T_2AG_3, T_2G_2AG, TGTGAG, and ACGTCA, respectively. All of these duplexes are 30 base pairs in starting length and have the corresponding melting points shown in Table 1 under similar DNA concentrations and reaction buffer conditions. Clearly, the $(T_2AG_3)_5$, $(T_2G_2AG)_5$, and $(TGTGAG)_5$ duplexes are all expanded readily at around 75 °C. This is not surprising since these three starting duplexes have melting temperatures of 73, 74, and 75 °C, respectively (Table 1). The starting duplex consisting of the relatively random hexanucleotide repeat ACGTCA, also expands readily but at the higher temperature of 80 °C, which is in agreement with the higher melting point of 78 °C for this duplex (see Table 1). Denaturing polyacrylamide gel electrophoresis of the expansion products for these duplexes under similar conditions but with ^{32}P-labeled DNA showed that the DNA is also incrementally increased in length by discrete units corresponding to the length of the repeat (see Supporting Information). Thus, it seems unlikely that DNA structures are involved in the observed expansions. Rather, the observed expansions appear to be largely dependent on the temperature of the reaction, with

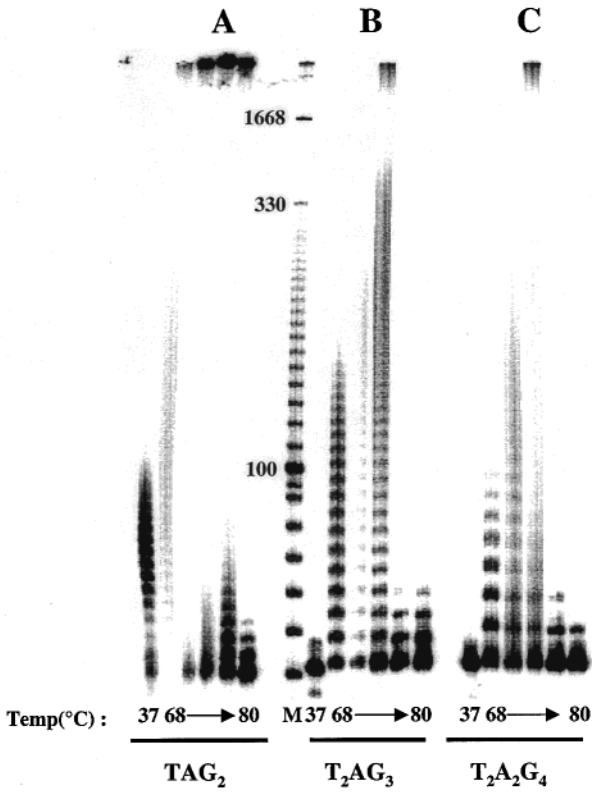

FIGURE 2: Temperature-dependent expansions of duplexes with a starting length of 40 base pairs and consisting of (A) tetranucleotide (TAG_2), (B) hexanucleotide (T_2AG_3), and (C) octanucleotide ($T_2A_2G_4$) repeats. The products were analyzed by 6% denaturing PAGE and only the results for the G-rich DNA strands are shown. The reactions at 68, 71, 74, 77, and 80 °C (68 → 80 °C in the figure) were carried out with Taq DNA polymerase. The reactions at 37 °C were carried out with Klenow pol I. The reaction time was 1 h at the indicated temperature in all cases. Lane M is a 30−330 AFLP DNA ladder. The material in some of the wells is DNA that has been expanded well beyond the resolution limit of the gel.

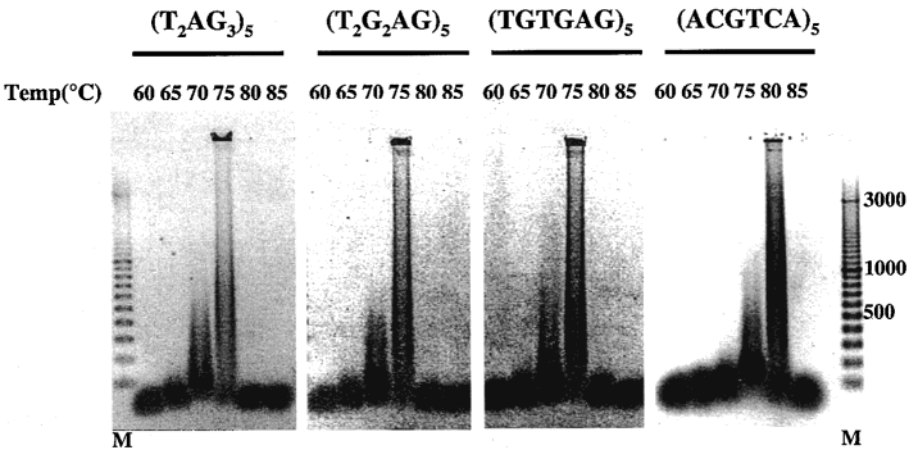

FIGURE 3: Temperature-dependent expansions with Taq DNA polymerase of duplexes with starting lengths of 30 base pairs and consisting of different hexanucleotide repeat sequences. The reactions conditions are the same as in Figure 1. Lane M represents a 100 base pair DNA ladder.

the expansions being more prominent near the melting point of the starting oligonucleotide.

A theoretical thermodynamic study of triplet repeat expansions was recently carried out by Harvey (*17*), who concluded that at 37 °C entropic considerations should favor expansion of repeats in duplexes longer than 300 base pairs, regardless of the sequence. This is in agreement with the observations in vivo that expansions of triplet repeats are length-dependent. Thus, below a threshold of 105−120 base pairs, triplet expansions are less likely to take place, but the propensity for expansion increases beyond this length (*9*). The starting duplexes studied above are rather short, and except for the TAG_2 duplex, which shows some expansion at 37 °C, they do not expand at this lower temperature. It was thus of interest to determine whether expansion of longer sequence repeats could be achieved in vitro at 37 °C with starting duplexes longer than 100 base pairs, and we chose for these studies duplexes containing the hexanucleotide repeats T_2AG_3 and T_2G_4. Faced with the technical difficulties of synthesizing telomeric repeat duplexes greater than 100 base pairs in length, we opted to synthesize DNA duplexes of various lengths via the Taq polymerase expansion reaction as in Figures 1 and 3. We next carefully cut out segments in the agarose gel corresponding to lengths of approximately <100 base pairs, 100−200 base pairs, 200−300 base pairs, and 300−400 base pairs, using the DNA marker as a guide. The DNAs were then eluted from the agarose gel, dried and redissolved in polymerase buffer, and then incubated with Klenow fragment and dNTPs at 37 °C, and the products were analyzed by agarose gel electrophoresis. Figure 4 shows that, in the case of the T_2AG_3 repeat, there is apparent expansion of the DNA for starting duplex lengths greater than 200 base pairs, whereas for the T_2G_4 sequence there is apparent expansion of the DNA for starting duplex lengths greater than 100 base pairs. Although the T_2G_4 sequence appears to expand at shorter duplex lengths than the T_2AG_3 sequence, this apparent difference is most likely due to the difficulties in accurately cutting the corresponding DNA bands from the agarose gel.

The apparent observed expansions in Figure 4 appear to confirm Harvey's predictions that the propensity for repeat expansion increases with increasing duplex length, perhaps through strand slippage. However, it is entirely possible that

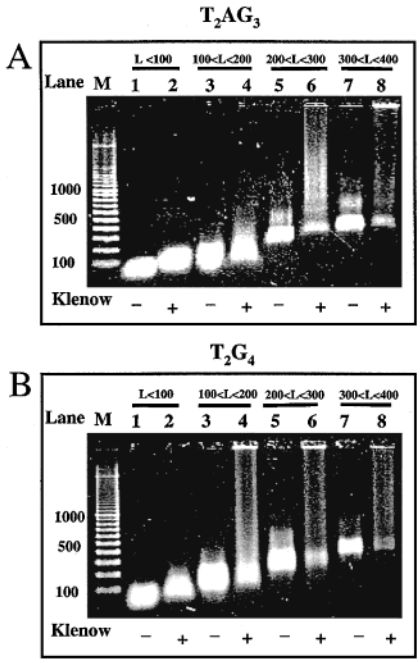

FIGURE 4: Length-dependent expansions of (A) the hexanucleotide repeat (T_2AG_3) and (B) the octanucleotide repeat (T_2G_4) by DNA polymerase I Klenow fragment at 37 °C. The approximate ranges of lengths of the starting duplexes are shown above the lane numbers. Lane M is a 100 base pair DNA ladder. The samples in the odd-numbered lanes were incubated in the same reaction mixture but without DNA polymerase I Klenow fragment.

the longer products in lanes 6 and 8 in Figure 4A and lanes 4, 6, and 8 in Figure 4B may not arise from strand slippage. Because the DNA for the reactions at 37 °C was isolated from fragments of different lengths derived from the Taq polymerase reaction, it is very likely that the DNA duplexes might contain overhangs and bulges. These irregularities, particularly overhangs, combined with the known strand-displacement and template-switching activities of Klenow fragment might give rise to expansion through simple extension of noncovalently bound DNA fragments. To examine whether a pool of DNA duplexes with single-stranded overhangs could give rise to similar expansion profiles, we carried out the same reactions using duplexes (Table 1) with one to three repeat overhangs (duplexes F-1 to F-3). These duplexes contain a middle random sequence

to ensure duplex formation and to rule out the possibility of strand slippage. The CG54 duplex, which lacks a random sequence, and the F-0 duplex, which lacks overhangs, were both used as controls. To determine the effects of strand displacement and/or template switching on the observed expansions, the duplexes were incubated in the presence of dNTPs and either Klenow fragment, which is known to have strand-displacement and template-switching activities, or T4 DNA polymerase, which has been reported to have strand-displacement activity (*18*).

Figure 5A shows that when the starting duplex contains three repeat overhangs (F-3 lanes), nondenaturing agarose gel electrophoresis analysis of the ^{32}P-labeled G-rich DNA strand gives rise to bands that appear to correspond to expansion products in the presence of either Klenow (F-3, K lane) or T4 (F-3, T lane) DNA polymerases. The control lane (F-3, C lane) also shows what appear to be expansion products, but these are in fact long hybrid DNA duplexes that result from annealing through the three repeat overhangs. This is clearly evident from the regular banding pattern observed in this lane. In the case of the CG54 duplex (consisting of nine repeats of TTAGGG; see Table 1) and F-0, neither the control lane nor the lanes corresponding to added Klenow and T4 DNA polymerases show any slow-migrating bands. However, the F-1 through F-3 groups of lanes show that slow-migrating bands begin to appear as the number of repeat overhangs increases, particularly in the presence of Klenow fragment. The control lanes also show slow-migrating bands that begin to appear when there are at least two repeat overhangs in the duplex (F-2 and F-3, C lanes). Similar results were obtained for the complementary DNA strand containing C-rich flanking regions (not shown).

Clearly, the patterns of slow-migrating bands for the Klenow and T4 DNA polymerase lanes for the F-1 through F-3 duplexes in Figure 5A are very different. This prompted us to analyze the reaction products through denaturing polyacrylamide gel electrophoresis. Figure 5B shows the results for the ^{32}P-labeled DNA strand containing G-rich flanking regions. Thus, only duplexes with overhangs (F-1 through F-3) show products longer than the starting strand length of 54 bases. These longer products must be a result of the template-switching and/or strand-displacement activities of the polymerases, since strand slippage is unlikely to occur in these duplexes. In the reactions run with T4 DNA polymerase (T lanes in the F-1 through F-3 groups of lanes), slow-migrating bands can be observed that we attribute to products arising from annealing of two duplexes through their repeat overhangs, followed by 3′-end extension through strand displacement by T4 DNA polymerase. The overall product length does not exceed 100 base pairs in length and the products gradually decrease in length by one repeat in the series of duplexes F-1 through F-3, as would be expected if the products arose from strand displacement. Thus, we find that T4 DNA polymerase has significant strand-displacement activity under the conditions studied, in agreement with previous observations (*18*). In the reactions run with Klenow fragment (K lanes in the F-1 through F-3 groups of lanes), the length of the reaction products is longer than would be expected as a result of only the strand-displacement activity of this DNA polymerase. However, since the products are not significantly longer than about 150 bases and since the CG54 duplex is not expanded at all, it is

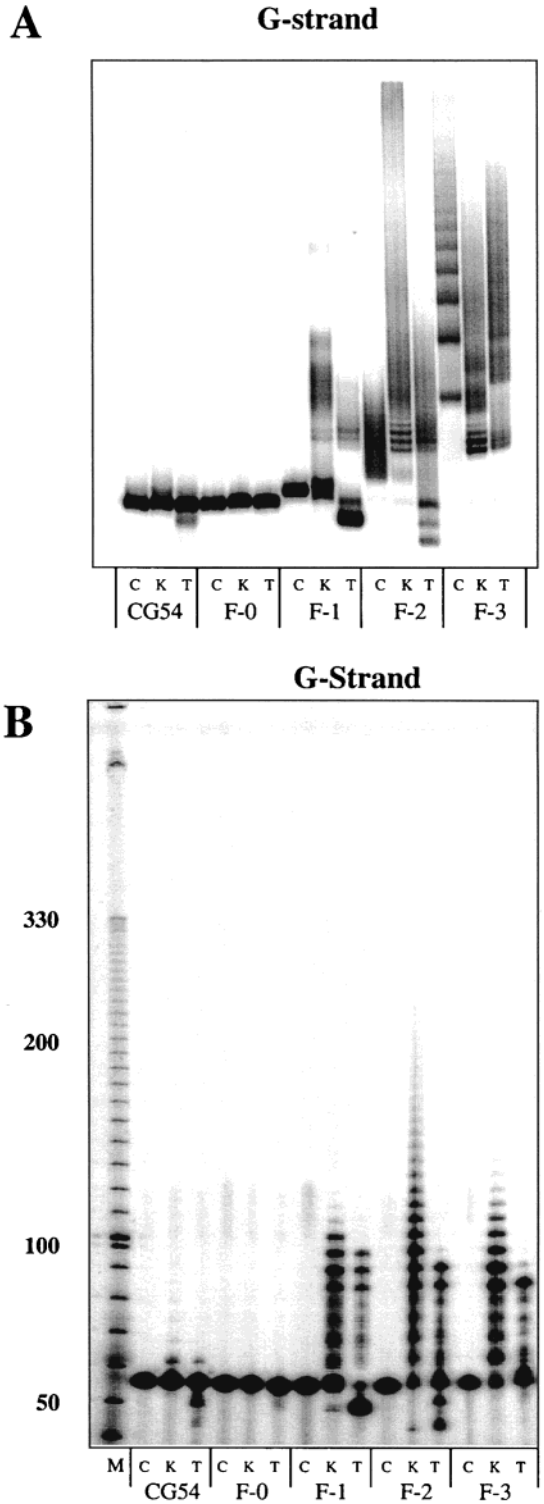

FIGURE 5: (A) Nondenaturing polyacrylamide gel electrophoresis of the reaction products resulting from the incubation of duplexes F-0 through F-3 (Table 1) in the presence of either Klenow fragment (K lanes) or T4 DNA polymerase (T lanes). C indicates control lanes without polymerase added. F-1 through F-3 are experiments with duplexes containing one, two, and three repeat overhangs, respectively. F-0 indicates experiments run with a duplex devoid of overhangs. CG54 indicates experiments run with a 54 base pair duplex consisting of nine TTAGGG repeats. (B) Denaturing polyacrylamide gel electrophoresis of the reaction products run under similar conditions as in Figure 4B. In both panels, only the results for the strands containing G-rich flanking regions are shown. The fast-migrating bands in the T lanes are likely due to shorter duplexes resulting from the 3′ → 5′ exonuclease activity of T4 DNA polymerase.

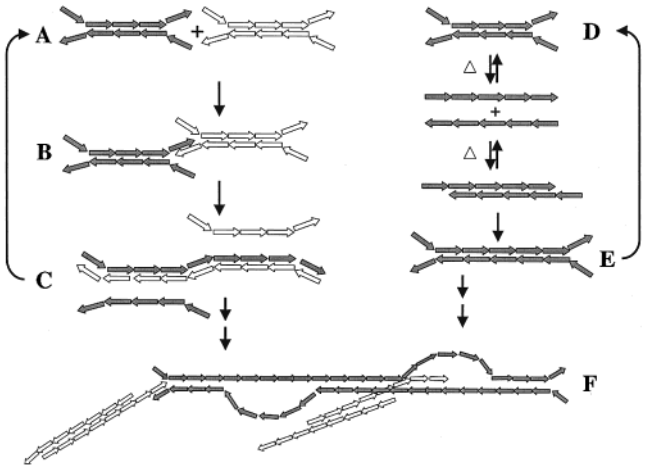

FIGURE 6: Proposed mechanism for the observed in vitro expansions of long DNA repeats. See text for explanation.

unlikely that these longer products arise from strand slippage. Rather, it is highly likely that the observed products are the result of the combined strand-displacement and template-switching activities of Klenow fragment.

DISCUSSION

One reasonable mechanism for the observed repeat expansions is shown in Figure 6. Thus, at lower temperatures, the duplexes are stable, and little or no expansion can take place. However, fraying at the ends of the duplexes can occur and eventually becomes more pronounced as the reaction temperature is increased (Figure 6A). Due to end fraying, there may be transient intermolecular base pairing between duplexes in solution (Figure 6B). The strand-displacement activity of the polymerase may then be responsible for extending the complementary strands of the DNA duplexes in these intermolecular complexes (Figure 6C). On the other hand, at reaction temperatures near the melting point of the duplex, repeated melting and staggered reannealing, followed by polymerase extension, could lead to more efficient extension of the two DNA strands (Figure 6D,E). Eventually, as the DNA increases in length, bulges may form in the duplex at elevated temperatures. Single-stranded DNA in bulged regions may then serve as templates for the extension of other duplexes present in solution (Figure 6F). Of course, as expected, at temperatures much higher than the melting point of the starting duplex, there is only single-stranded DNA and extension of the DNA is abolished.

The relevance of the mechanism in Figure 6 for expansion of repetitive DNA sequences in vivo is unclear at this point. However, the mechanism explains many aspects of the observed expansions in vitro. It explains the observed increase in DNA length by discrete repeat units (Figure 2 and Supporting Information), since staggered reannealing could easily lead to one repeat overhangs that are then replicated by the DNA polymerase. Furthermore, base pairing of discrete sequence repeats through end fraying and subsequent expansion through the mechanism indicated in Figure 6A−C could also occur. The latter mechanism also explains why some expansion of duplexes containing shorter repeats is observed at temperatures much lower than the melting point of the starting duplex. Expansion through this mechanism would be favored with decreasing sequence repeat length, since the number of base pairs needed to be unwound to allow for transient intermolecular base pairing through one repeat unit is reduced. One example is the TAG_2 duplex, which is expanded somewhat at 37 °C, whereas the $T_2A_2G_4$ duplex is not (Figure 2), despite the fact that both duplexes have similar melting points. Also, neither the CG54 nor the F-0 duplexes in Figure 5 are expanded in any way at 37 °C, since unwinding of six base pairs from either or both ends of the duplex (needed for expansion as shown in Figure 6A−C) is unlikely to occur at this temperature. Thus, a reasonable explanation is that intermolecular base pairing through one repeat unit is more likely to occur for the TAG_2 duplex, giving rise to some expansion of the duplex via strand displacement as shown in Figure 6A−C.

It is apparent from a comparison of the expansion products in Figure 2 that expansion becomes more efficient at higher temperatures, since in this case all of the intermolecular interactions in Figure 6 that lead to expansion can take place. The overall mechanism in Figure 6 may also explain the lack of sequence effects on the expansions with Taq DNA polymerase shown in Figure 3. To further test the model in Figure 6, we carried out expansions of hexanucleotide repeats with a variety of duplexes containing nonrepeat sequences that would prevent the type of molecular events shown in Figure 6 from taking place. As expected, none of these duplexes was expanded when compared to a control duplex (see Supporting Information), lending further support for this in vitro expansion model. Thus, from the expansion model in Figure 6, the role of DNA structure in the expansion of the short repetitive DNA duplexes studied here need not be invoked, and it is unlikely that strand slippage is responsible for the observed expansions.

Length-dependent in vitro expansion studies of DNA sequence repeats are hampered by the technical difficulties in constructing duplexes greater than 100 base pairs in length without irregularities, including single-strand overhangs and bulges. Although it might be possible to minimize the presence of such irregularities in these longer duplexes through nuclease pretreatment, it would be difficult to completely eliminate them. Thus, their presence makes it difficult to test the theory proposed by Harvey (*17*) [supported by in vivo observations] that the propensity for expansion through strand slippage should increase with duplex length. From the expansion experiments at 37 °C in Figure 4 (even-numbered lanes and DNA greater than 100 base pairs), one cannot completely rule out that the products result from true expansions that occur through strand slippage. However, since the DNA was obtained following expansion with Taq DNA polymerase, it is likely that the starting DNA duplexes contained the above-mentioned irregularities. Thus, the combined strand-displacement and template-switching activities of Klenow pol I may give rise to the observed products due to intermolecular interactions involving complexes such as those in Figure 6 panels B and F. Indeed, some smearing of the DNA in the odd-numbered lanes in Figure 4 indicates that intermolecular base pairing may be taking place. The experiments in Figure 5 with duplexes containing overhangs indicate that intermolecular base pairing, combined with the strand-displacement and template-switching activities of the polymerases, may be partly responsible for the observed expansions. Thus, when the DNA duplexes contain at least one repeat overhang (F-1

through F-3 duplexes), nondenaturing gel electrophoresis shows bands that appear to result from expansion by the DNA polymerases (Figure 5A, lanes K and T). However, when the same reaction products are examined under denaturing conditions (Figure 5B), the results indicate that the reaction products are most likely a result of the strand-displacement activity of T4 DNA polymerase (T lanes), and the combined strand-displacement and template-switching activities of Klenow fragment (K lanes). Therefore, the apparent expansions in Figure 4 could very well arise from interduplex base pairing and subsequent filling in by the polymerase. However, as mentioned above, the role of some strand slippage in longer duplexes cannot be completely eliminated at this point. Furthermore, the possibility exists that slippage in long duplexes may only work with shorter repeats and not longer repeats. More comprehensive studies will be needed to determine whether this is the case.

A number of studies on trinucleotide repeat expansions have been carried out at 37 °C and higher temperatures with short DNA duplexes (*11*−*15*). In many instances these short starting duplexes contained overhangs and the studies were carried out with polymerases known to possess strand-displacement or both strand-displacement and template-switching activities. Although the melting points of these short starting duplexes were not measured, they most likely fall close to the optimum expansion temperatures observed for these duplexes. However, even at 37 °C, intermolecular base pairing through at least one triplet repeat unit is more likely to occur in these short duplexes, giving rise to expansion through the process described in Figure 6A−C. Thus, it is likely that the observed in vitro triplet repeat expansions in previous reports that utilized short duplexes also result from the expansion mechanism shown in Figure 6, rather than the conventional strand slippage mechanism. It is worth mentioning, however, that most of these previous studies were carried out with lower DNA concentrations (typically 2−5 μM in duplex) than those used here (10 and 50 μM in duplex). Thus, it could be argued that only the higher DNA concentrations lead to expansion through intermolecular interactions coupled with strand displacement, as shown in Figure 6. However, we have repeated the experiment in Figure 2 with a DNA concentration of 0.2 μM in duplex using [α-^{32}P]dGTP as a tracing element with similar results (see Supporting Information). Thus, even at these lower DNA concentrations, expansion seems to take place through the mechanism shown in Figure 6.

In summary, although we did not examine all available DNA polymerases and all possible DNA repeats, the combined data suggest that in vitro expansions of repetitive DNA from short DNA duplexes may not necessarily arise from a conventional strand-slippage mechanism. Rather, a mechanism such as that shown in Figure 6 is most likely responsible for the observed expansion of short repetitive DNA duplexes. Thus, the melting point of the starting oligonucleotide and the length of the sequence repeat, as well as the strand-displacement and template-switching activities of the DNA polymerases used, can play a major role in the observed expansions. Therefore, these factors should be taken into account in in vitro expansion studies of short repetitive DNA duplexes.

SUPPORTING INFORMATION AVAILABLE

Three figures showing denaturing PAGE of temperature-dependent expansion of the G-rich strands for the 30-mer duplexes, the C-rich strands for the 40-mer duplexes, and the expansion products analogous to those shown in Figure 2 but at lower concentration, and a table and a figure showing the negative effect of random flanking sequences on the expansion reactions. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Timchenko, L. T., and Caskey, C. T. (1996) *FASEB J. 10*, 1589−1597.
2. Grady, D. L., Ratliff, R. L., Robinson, D. L., McCanlies, E. C., Meyne, J., and Moyzis, R. K. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 1695−1699.
3. Catasti, P., Gupta, G., Garcia, A. E., Ratliff, R., Hong, L., Yau, P., Moyzis, R. K., and Bradbury, E. M. (1994) *Biochemistry 33*, 3819−3830.
4. Henderson, E. (1995) in *Telomeres* (Blackburn, E. H., and Greider, C. W., Eds.) pp 11−34, Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY.
5. Bell, G. I., Selby, M., and Rutter, W. J. (1982) *Nature 295*, 31−35.
6. Abad, J. P., Carmena, M., Baars, S., Saunders, R. D. C., Glover, D. M., Ludeña, P., Sentis, C., Tyler-Smith, C., and Villasante, A. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 4663−4667.
7. Carmena, M., Abad, J. P., Villasante, A., and Gonzalez, C. (1993) *J. Cell Sci. 105*, 41−50.
8. Verkerk, A. J. M. J., Pieretti, M., Sutcliffe, J. S., Fu, Y. H., Kuhl, D. P. A., Pizzuti, A., Reiner, O., Richards, S., Victoria, M. F., Zhang, F., Eussen, B. E., Van Ommen, G. J. B., Blonden, L. A. J., Riggins, G. J., Chastain, J. L., Kunst, C. B., Galjaard, H., Caskey, C. T., Nelson, D. L., Oostra, B. A., and Warren, S. T. (1991) *Cell 65*, 905−914.
9. Bowater, R. P., and Wells, R. D. (2000) *Prog. Nucleic Acids Res. Mol. Biol. 66*, 159−202.
10. Greider, C. W., and Blackburn, E. H. (1985) *Cell 43*, 405−413.
11. Behn-Krappa, A., and Doerfler, W. (1994) *Hum. Mutat. 3*, 19−24.
12. Schlötterer, C., and Tautz, D. (1992) *Nucleic Acids Res. 20*, 211−215.
13. Ji, J., Clegg, N. J., Peterson, K. R., Jackson, A. L., Laird, C. D., and Loeb, L. A. (1996) *Nucleic Acids Res. 24*, 2835−2840.
14. Topal, M. D., and Lyons-Darden, T. (1999) *Nucleic Acids Res. 27*, 2235−2240.
15. Topal, M. D., and Lyons-Darden, T. (1999) *J. Biol. Chem. 274*, 25975−25978.
16. Lalioti, M. D., Scott, H. S., Buresi, C., Rossier, C., Bottani, A., Morris, M. A., Malafosse, A., and Antonarakis, S. E. (1997) *Nature 386*, 847−851.
17. Harvey, S. C (1997) *Biochemistry 36*, 3047−3049.
18. da Silva, E. F., and Reha-Krantz, L. J. (2000) *J. Biol. Chem. 275*, 31528−31535.

BI0110950