Overcoming a barrier for DNA polymerization in triplex-forming sequences

Vladimir N. Potaman* and John J. Bissler¹

Institute of Biosciences and Technology, The Texas A&M University System Health Science Center, 2121 West Holcombe Boulevard, Houston, TX 77030, USA and ¹Children’s Hospital Research Foundation, 3333 Burnette Avenue, Cincinnati, OH 45229, USA

Received May 30, 1999; Accepted June 7, 1999

ABSTRACT

Folded structures in the DNA template, such as hairpins and multi-stranded structures, often serve as pause and arrest sites for DNA polymerases. DNA polymerization is particularly difficult on mirror-repeated homopurine-homopyrimidine templates where triple-stranded (triplex) structures may form between the nascent and folded template strands. In order to use a linear PCR amplification approach for the structural analysis of DNA in mirror-repeated sequences we modified a conventional protocol. The barrier for DNA synthesis can be eliminated using an oligonucleotide that hybridizes with the template to prevent its folding and is subsequently displaced by the progressing polymerase. The described approach is potentially useful for sequencing and analysis of chemical adducts and point mutations in a variety of sequences prone to the formation of folded structures, such as long hairpins and quadruplexes.

INTRODUCTION

Primer extension-based methods are often used for mutational analysis as well as for analysis of DNA structure, susceptibility to chemicals and UV light, and protein–DNA interactions (1–5). The conformational state of the analyzed template is important for polymerization efficiency. Folded structures in the template, such as hairpins and three- and four-stranded structures, impede the progression of DNA polymerase (6–11). The DNA polymerase-blocking effects are more pronounced at lower temperatures; however, in (G+C)-rich sequences they often persist even at the elevated temperatures employed in PCR procedures (11,12). DNA polymerization is particularly difficult in mirror-repeated homopurine-homopyrimidine (Pu·Py) sequences (9,10). With these templates polymerization arrest is best understood in terms of triple-stranded structure (triplex) formation (13,14). Triplex may either pre-form locally in an otherwise double-stranded DNA or form during polymerization between the nascent and folded template strands (7,9). DNA polymerization-based primer extension analysis at elevated temperatures worked well in our experiments with mirror-repeated sequences of moderate (G+C) content and relatively high (G+C) content but rather imperfect mirror symmetry (15–17). However, during structural studies of a largely Pu·Py sequence in intron 21 from the polycystic kidney disease-associated gene PKD1 (18) the primer extension approach proved difficult for a very (G+C)-rich mirror repeat (74% G+C; Fig. 1A). The presumed triplex barrier for DNA synthesis can be overcome using an oligonucleotide that hybridizes with the template to prevent its folding and is subsequently displaced by the progressing polymerase. The relief of DNA polymerization arrest in two other (G+C)-rich triplex-forming sequences from the PKD1 gene demonstrates the generality of this approach.

MATERIALS AND METHODS

Materials

The plasmid DNA contained several Pu·Py sequences cloned in the EcoRI sites of pUC8 (pCRAS27) and pUC19 (pCW1070, pCW2242 and pCW2966). Oligonucleotide primers and unmodified ‘helpers’ for each DNA template were custom synthesized by Genosys Biotechnologies (The Woodlands, TX) whereas 3’-NH₂-modified ‘helpers’ were synthesized and chromatographically purified by Integrated DNA Technologies (Coralville, IA). Primers were 5’-end-labeled with [γ-³²P]ATP (New England Nuclear, Boston, MA) using T4 polynucleotide kinase. DNA polymerases used were Taq ( Gibco BRL, Gaithersburg, MD), the Stoffel fragment of AmpliTaq (Perkin-Elmer, San Francisco, CA), KlenTaq (Ab Peptides, St Louis, MO) and Pfu (Stratagene, La Jolla, CA). An LA-16 enzyme mix was made by mixing 15 vol of KlenTaq (25 U/μl) and 1 vol of Pfu (5 U/μl) polymerases. T4 polynucleotide kinase and restriction endonucleases were from New England Biolabs (Beverly, MA). Betaine and dimethylsulfoxide (DMSO) were purchased from Sigma (St Louis, MO) and Aldrich (Milwaukee, WI), respectively. Nucleoside triphosphates were from Gibco BRL, whereas dideoxynucleoside triphosphates (ddNTP) were from Boehringer Mannheim (Indianapolis, IN).

DNA sequencing

For dideoxy sequencing termination mixes contained 50 μM each dNTP and one of the following ddNTP: A (1.5 mM ddATP), C (0.5 mM ddCTP), G (10 μM ddGTP) or T (0.8 mM ddTTP). In most cases sequencing analysis was accomplished on PvuII–PvuII fragments of the plasmids to prevent dNTP

*To whom correspondence should be addressed. Tel: +1 713 677 7675; Fax: +1 713 677 7689; Email: vpotaman@ibt.tamu.edu
and ddNTP pools in reactions from depleting via their incorporation into irrelevant neighboring sequences. For sequencing with Taq polymerase each reaction mix containing about 10$^9$ copies of template DNA, 0.6 U Taq polymerase, 45 nM 32P-labeled primer, 10 µM dNTPs and 5-fold diluted termination mix in PCR buffer was subjected to 20 PCR cycles (95°C, 30 s; 55°C, 30 s; 70°C, 1 min). For sequencing with LA-16 each reaction mix containing 1 U KlenTaq and 0.015 U Pfu polymerases, 25 nM 32P-labeled primer, 2 M betaine, 9% DMSO, 10 µM dNTPs and ddNTPs in KlenTaq buffer was subjected to 20 PCR cycles (95°C, 30 s; 55°C, 30 s; 68°C, 1 min). Wherever appropriate, variations in the amount of LA-16 and additions of oligonucleotide helpers and denaturants to sequencing reactions are indicated in the figure legends. Products of sequencing reactions were separated in a 7% denaturing polyacrylamide gel.

**RESULTS**

**Using oligonucleotides to overcome a structural barrier for DNA polymerization**

During structural studies of the longest mirror-repeated Pu·Py sequence from intron 21 of the PKD1 gene (Fig. 1A) we observed a significantly different primer extension on complementary templates. Thermophilic Taq DNA polymerase and its Stoffel fragment efficiently replicate the mirror-repeated pyrimidine template strand of plasmid pCW2966 (data not shown). However, they exhibit multiple strong pause sites in the 3'-half of the mirror-repeated purine strand (Fig. 2A). This is likely due to triplex formation between the newly synthesized strand and the folded template that cannot be easily untangled by DNA polymerase (Fig. 1B). This barrier seems to be particularly strong, which is not surprising as a rather long triplex can be formed consisting of 21 base triads. Changing the primer annealing and DNA synthesis temperatures, as well as the ionic strength of the polymerization reaction, in order to disrupt the triplex did not result in an easier passage of the triplex-forming sequence by Taq DNA polymerase. To overcome the structural barrier, several modifications to the PCR protocol were made. These included the addition of denaturants to destabilize secondary structures in the template and using a mixture of DNA polymerases (LA-16) as well as additional ‘helper’ oligonucleotides. Denaturants, such as DMSO and betaine, destabilize secondary structures in the template, thereby reducing their effects on polymerization efficiency (12,19). A recently developed DNA polymerase system (LA-16) is efficient in amplification of long (up to 35 kb) and very (G+C)-rich (up to 80%) templates (12,20).

**Figure 1.** Possible mechanisms of polymerization arrest and its elimination. (A) The longest mirror-repeated purine sequence from the PKD1 gene and oligonucleotides that hybridize across the repeat center. (B) The polymerization arrest may occur after a nascent pyrimidine strand reaches the center of the mirror repeat. The Watson–Crick duplex between the nascent pyrimidine strand and half of the purine template serves as a target for binding via Hoogsteen hydrogen bonds to the other half of the purine template still unused for DNA polymerization. Polymerase is unable to untangle the triplex and/or continue synthesis around a hairpin tip. (C) Oligonucleotide hybridization across the center of the mirror repeat prevents the purine strand from hairpin formation. A progressing polymerase displaces the oligonucleotide and synthesizes a new strand on an unfolded template.

**Figure 2.** Test of polymerization efficiency on the mirror-repeated purine template in plasmid pCW2966 using sequencing reactions with the M13 Reverse primer hybridizing 80 nt apart from the Pu-Py sequence. Plasmid DNA was digested with PvuII so that the full-length primer extension proceeds 92 nt past the mirror repeat. (A) Sequencing with Taq polymerase. (B) Sequencing with LA-16 (15:1 v/v mixture of KlenTaq1 and Pfu polymerases). (C) Sequencing with LA-16 and helper oligonucleotide I. Reaction conditions were as in (B) except that the LA-16 concentration was doubled and 0.5 µM helper oligonucleotide I was included in reactions. (D) Sequencing with LA-16 and helper oligonucleotide III. Reaction conditions as in (C) were modified to change the PCR cycling (95°C, 30 s; 55°C, 30 s; 63°C, 3 min) and include 2.5 µM helper oligonucleotide III instead of oligonucleotide I.
polymerases. The former is a proofreading-deficient Taq polymerase fragment analogous to the Klenow fragment of Escherichia coli DNA polymerase, whereas Pfu polymerase is proofreading proficient. Using LA-16 and the addition of betaine to 2 M and DMSO to 9% results in efficient elimination of the pause sites in the 3'-half of the mirror-repeated purine strand (Fig. 2B). However, sequencing bands in the 5'-half of the purine mirror repeat are only 20–30% as intense as those in the 3'-half. There is also a strong polymerase blockage at the middle of the mirror repeat. Assuming that the polymerization barrier arises due to triplex formation, we used several pyrimidine oligonucleotides (Fig. 1A) that hybridize with the middle and 5'-half of the purine sequence. These 'helper' oligonucleotides should prevent the DNA template from folding into a part triplex conformation (Fig. 1C).

Two types of helper oligonucleotides have been tested for their efficiency in assisting DNA polymerization on the purine template. When sequencing reactions contained a 20-fold excess of the helper 20mer I relative to the radiolabeled primer, the strong blockage site was largely eliminated and LA-16 efficiently polymerized DNA in the homopurine region and beyond (Fig. 2C). Oligomer I with a free 3'-OH group was extended under the reaction conditions (data not shown) resulting in more stable hybridization at elevated PCR temperatures and efficient inhibition of template folding. The thermophilic polymerase seemed to efficiently unwind the double-stranded complex between the template and extended oligomer I since we detected normal products initiated by the labeled primer. The 3'-NH₂-modified 20mer II, which is unable to serve as a primer, did not alleviate polymerase pausing (data not shown). It is likely that at elevated PCR temperatures this oligomer does not bind to the template strongly enough to prevent its folding. At 100-fold excess over the radiolabeled primer, the longer 3'-NH₂-modified 30mer III reduced the amount of polymerase pausing by ~3-fold. An additional 3-fold reduction in pausing was achieved by stabilizing the template–oligomer III complex by reducing the extension temperature from 68 to 63°C (Fig. 2D). Although under these conditions polymerase pausing was still three to four times stronger than with an extendable 20mer, we believe that it can be further minimized by increasing oligomer length and optimizing its position with respect to the repeat center.

Testing the generality of the method

Triplex formation at physiological pH and salt is easier with an increasing percentage of guanines in the purine tract (13,14). Such (G+C)-rich sequences are necessarily made of blocks of guanines separated by more rare adenines. To verify the usefulness of helper oligonucleotides in overcoming structural barriers in other triplex-forming sequences, test cases included two other Pu·Py templates from the PKD1 gene, plasmids pCW1070 and pCW2242, and the Pu·Py sequence from oncogene c-Ki-ras, plasmid pCRAS27 (Fig. 3). The sequence patterns include blocks of two to six guanines. Helper oligonucleotides IV, V and VI were designed to overlap the centers of mirror repeats and prevent template folding in pCRAS27, pCW1070 and pCW2242, respectively. A shorter 27 bp mirror-repeated sequence in pCRAS27 (Fig. 3) may form a hairpin with 11–12 Hoogsteen-type base pairs and a 3–5 nt loop as part of a triplex structure. This structure presents a significant barrier for Taq polymerase (Fig. 4A). Using the LA-16 polymerase mix almost fully eliminates this barrier; only minor
DNA pausing was detected in the mirror-repeated sequence (Fig. 4B). Addition of the helper oligonucleotide IV completely eliminated non-specific chain termination on a purine template of plasmid pCRAS27 (Fig. 4C). With the pCW1070 template (Fig. 3), 14 Hoogsteen-type base pairs may form a hairpin with a 4 nt loop. The longer triplex (compared with that in pCRAS27) may form a stronger polymerization barrier. Several pause sites were prominent when Taq DNA polymerase was used (Fig. 5A). This pausing was greatly reduced by using the LA-16 mix (Fig. 5B). No barrier for DNA polymerization was noticeable when helper oligonucleotide V was included in the reaction (Fig. 5C). DNA polymerization on a purine template of plasmid pCW2242 (Fig. 3), where 19–20 Hoogsteen-type base pairs may form a hairpin with a 3–5 nt loop, was very similar to that for plasmid pCW2966 (Fig. 6). Polymerization pausing was significantly reduced when LA-16 was used and minimized by addition of helper oligonucleotide VI. Thus, the relief of the structural barrier for DNA polymerization in several tested sequence patterns demonstrates the generality of the helper oligonucleotide approach.

**DISCUSSION**

In the four mirror-repeated sequences tested in this work the strength of a polymerization barrier progressively increases with the length of possible triplex structures ($11 < 14 < 19 < 21$ triads). Plasmid pCW2966 has one of the longest known perfect mirror repeats (14,18). It therefore forms the strongest barrier to DNA polymerization and residual pausing is detected even with the helper oligonucleotide. Although it is impossible to provide direct evidence for the presence of triplex in PCR reactions, experimental facts in our study are well explained by triplex formation in mirror-repeated Pu-Py sequences. Polymerization arrest occurs after the Watson–Crick duplex between a nascent pyrimidine strand and half of the purine template is formed to the center of the mirror repeat. The other half of the purine template, still unused for DNA polymerization, binds to this duplex via Hoogsteen hydrogen bonds, thus forming a triplex. Magnesium in the reaction buffer is a factor that promotes triplex formation. To proceed past the middle of the mirror repeat, DNA polymerase must unwind the triplex, which is not feasible because the triple-stranded complex is stabilized by the hydrogen bonds behind the polymerase. Polymerase is unable to continue synthesis on a folded template and most likely dissociates, thereby terminating the nascent strand.

Several DNA polymerases tested in this work vary in the replication efficiency of the mirror-repeated Pu-Py sequences. The polymerase mixture, LA-16, is able to read through the short mirror-repeated sequence (plasmid pCRAS27). However, even the increased potency of LA-16 is insufficient to overcome pausing in the longer sequences from the PKD1 gene. For the long mirror-repeated sequences, prevention of purine template folding by hybridization of helper oligonucleotides results in overcoming the DNA polymerization barriers. In a reaction
with two oligonucleotides the helper oligonucleotide forms a Watson–Crick duplex with the template, thereby preventing its folding into a hairpin stabilized by Hoogsteen hydrogen bonds. DNA polymerase that extends the 5'-most oligomer (primer) then displaces the helper. Although oligonucleotide helpers that may be extended by DNA polymerase bind stronger to the template, the hybridization of non-extendable 3'-modified oligonucleotides might be stabilized by increasing their length or by decreasing the polymerization temperature. This may open up the possibility of using the 3'-NH$_2$-modified oligonucleotides it is possible to linearly amplify strong template folding may solve this problem.

Previously, attempts have been made to improve the read-through of secondary structure-forming templates. Using betaine, which reduces the melting temperature of G+C base pairs, resulted in overcoming the small hairpin barriers and pause sites at certain combinations of bases (20). Using LA-16 and a number of melting temperature-reducing agents (including DMSO and betaine) drastically improved amplification of very G+C-rich sequences that probably form hairpin- and quadruplex-like structures (12). It is likely that in these cases DNA polymerase is able to physically separate base paired strands. A very strong structural barrier for DNA polymerization forms through of secondary structure-forming templates. Using LA-16 and a number of melting temperature-reducing agents (including DMSO and betaine) drastically improved amplification of very G+C-rich sequences that probably form hairpin- and quadruplex-like structures (12). It is likely that in these cases DNA polymerase is able to physically separate base paired strands. A very strong structural barrier for DNA polymerization forms through of secondary structure-forming templates. Using LA-16 and a number of melting temperature-reducing agents (including DMSO and betaine) drastically improved amplification of very G+C-rich sequences that probably form hairpin- and quadruplex-like structures (12). It is likely that in these cases DNA polymerase is able to physically separate base paired strands. A very strong structural barrier for DNA polymerization forms through of secondary structure-forming templates. Using LA-16 and a number of melting temperature-reducing agents (including DMSO and betaine) drastically improved amplification of very G+C-rich sequences that probably form hairpin- and quadruplex-like structures (12). It is likely that in these cases DNA polymerase is able to physically separate base paired strands. A very strong structural barrier for DNA polymerization forms through of secondary structure-forming templates. Using LA-16 and a number of melting temperature-reducing agents (including DMSO and betaine) drastically improved amplification of very G+C-rich sequences that probably form hairpin- and quadruplex-like structures (12). It is likely that in these cases DNA polymerase is able to physically separate base paired strands. A very strong structural barrier for DNA polymerization forms through of secondary structure-forming templates. Using LA-16 and a number of melting temperature-reducing agents (including DMSO and betaine) drastically improved amplification of very G+C-rich sequences that probably form hairpin- and quadruplex-like structures (12). It is likely that in these cases DNA polymerase is able to physically separate base paired strands. A very strong structural barrier for DNA polymerization forms through of secondary structure-forming templates. Using LA-16 and a number of melting temperature-reducing agents (including DMSO and betaine) drastically improved amplification of very G+C-rich sequences that probably form hairpin- and quadruplex-like structures (12). It is likely that in these cases DNA polymerase is able to physically separate base paired strands. A very strong structural barrier for DNA polymerization forms through of secondary structure-forming templates. Using LA-16 and a number of melting temperature-reducing agents (including DMSO and betaine) drastically improved amplification of very G+C-rich sequences that probably form hairpin- and quadruplex-like structures (12). It is likely that in these cases DNA polymerase is able to physically separate base paired strands. A very strong structural barrier for DNA polymerization forms through of secondary structure-forming templates. Using LA-16 and a number of melting temperature-reducing agents (including DMSO and betaine) drastically improved amplification of very G+C-rich sequences that probably form hairpin- and quadruplex-like structures (12). It is likely that in these cases DNA polymerase is able to physically separate base paired strands. A very strong structural barrier for DNA polymerization forms through of secondary structure-forming templates. Using LA-16 and a number of melting temperature-reducing agents (including DMSO and betaine) drastically improved amplification of very G+C-rich sequences that probably form hairpin- and quadruplex-like structures (12). It is likely that in these cases DNA polymerase is able to physically separate base paired strands. A very strong structural barrier for DNA polymerization forms through of secondary structure-forming templates. Using LA-16 and a number of melting temperature-reducing agents (including DMSO and betaine) drastically improved amplification of very G+C-rich sequences that probably form hairpin- and quadruplex-like structures (12). It is likely that in these cases DNA polymerase is able to physically separate base paired strands. 

In conclusion, our results demonstrate that by using helper oligonucleotides it is possible to linearly amplify strong triplex-forming sequences. This is important since triplex-forming Pu·Py sequences occur in the human genome more often than statistically expected (21). In particular, mirror-repeated Pu·Py sequences may occur as often as once in 50 000 nt (22). Our modifications to the PCR protocol allowed efficient primer extension analysis of possible secondary structures within intron 21 of the PKD1 gene (manuscript in preparation). The helper oligonucleotide approach is potentially useful for sequencing and analysis of chemical adducts and point mutations in other sequences prone to the formation of folded structures, such as long hairpins and quadruplexes.

ACKNOWLEDGEMENTS

We wish to thank Drs E. A. Oussatcheva and R. R. Sinden for experimental help and discussions. This work was supported by the Texas Advanced Technology Program under Grant no. 999902-067 to V.N.P.

REFERENCES