The EMBO Journal vol.12 no.13 pp.4975–4983, 1993

Suicidal nucleotide sequences for DNA polymerization

George M.Samadashwily[1], Andrey Dayn[1,2] and Sergei M.Mirkin[1]

[1]Department of Genetics, University of Illinois at Chicago, Chicago, IL 60612, USA
[2]Present address: Abbott Laboratories, Molecular Diagnostics D907, Abbott Park, IL 60064, USA

Communicated by A.Rich

Studying the activity of T7 DNA polymerase (Sequenase) on open circular DNAs, we observed virtually complete termination within potential triplex-forming sequences. Mutations destroying the triplex potential of the sequences prevented termination, while compensatory mutations restoring triplex potential restored it. We hypothesize that strand displacement during DNA polymerization of double-helical templates brings three DNA strands (duplex DNA downstream of the polymerase plus a displaced overhang) into close proximity, provoking triplex formation, which in turn prevents further DNA synthesis. Supporting this idea, we found that Sequenase is unable to propagate through short triple-helical stretches within single-stranded DNA templates. Thus, DNA polymerase, by inducing triplex formation at specific sequences in front of the replication fork, causes self-termination. Possible biological implications of such 'conformational suicide' are discussed. Our data also provide a novel way to target DNA polymerases at specific sequences using triplex-forming oligonucleotides.

Key words: DNA polymerase/*H-DNA/replication/termination/triple-helical DNA

Introduction

Many DNA polymerases, including *Escherichia coli* Pol I, thermophilic Vent polymerase, DNA polymerases of phages T5 and ϕ29, and chemically or genetically modified T7 DNA polymerase (Sequenase), are able to displace the non-template DNA strand in the course of DNA synthesis without the assistance of energy-driven helicases and other proteins (reviewed in Kornberg and Baker, 1992). The role of displaced overhangs in genetic recombination has been discussed previously (Lundquist and Olivera, 1982). An interesting structural consequence of strand displacement has not, however, been investigated. The displaced strand may fold back, forming an intramolecular triplex downstream of the replication fork (Figure 1A).

It is well known that intramolecular DNA triplexes are formed by special sequences, perfect or near perfect homopurine–homopyrimidine mirror repeats, which for simplicity we will call H motifs (Mirkin *et al.*, 1987; Dayn *et al.*, 1992). Under mild acidic conditions they consist of two pyrimidine and one purine strand (H-DNA) (Mirkin *et al.*, 1987; Hanvey *et al.*, 1988; Htun and Dahlberg, 1988; Johnston, 1988; Voloshin *et al.*, 1988), while under neutral pH, in the presence of bivalent cations, they include two purine and one pyrimidine strand (*H-DNA) (Kohwi and Kohwi-Shigematsu, 1988; Bernues *et al.*, 1989, 1990). In both cases the orientation of the two chemically similar strands is antiparallel, and free energy from DNA supercoiling is required for triplex extrusion. Conditions of DNA polymerization *in vitro* are close to optimal for the formation of *H-DNA. We suggested, therefore, that if DNA polymerase met a potential triplex-forming DNA sequence, the displacement of the purine-rich strand would lead to triplex formation. The strand orientation in this case will be antiparallel by definition, and DNA polymerase driven overhang formation might efficiently substitute for DNA supercoiling.

Formation of triplexes prior to or during DNA polymerization may prevent further elongation. Manor and his co-authors (Lapidot *et al.*, 1989; Baran *et al.*, 1991) found that DNA polymerase terminates in the middle of d(G-A)$_n$ and d(T-C)$_n$ stretches in single-stranded templates. It was hypothesized that when DNA polymerase synthesizes half of the homopolymer strand, the rest of the stretch folds back forming an H-like structure which serves as a trap for the enzyme. We have recently found that different *H type triplexes present in double-helical templates prior to polymerization cause DNA polymerase to terminate at their borders (Dayn *et al.*, 1992). One would expect, therefore, that if polymerase-driven strand displacement promotes triplex formation, this would block further polymerization. This hypothesis gives three strong predictions: (i) H motifs should be suicidal for DNA polymerases capable of strand displacement; (ii) the displacement of the purine-rich, rather than the pyrimidine-rich strands would cause premature termination, since only purine/purine/pyrimidine triplexes are stable under polymerization conditions; and (iii) termination should occur at the pseudosymmetry site of H motifs where the purine-rich strand folds back to adopt the *H conformation.

To check this hypothesis, we studied the activity of modified T7 DNA polymerase on nicked double-helical templates containing sequences with different *H-forming potential. Due to the lack of DNA supercoiling those templates did not contain any non-B structures including triplexes prior to polymerization. In complete agreement with our predictions, we found that when the purine-rich strand was displaced during polymerization, Sequenase was practically completely blocked at the middle of these sequences as it should be in the case of triplex formation. Conversely, when the pyrimidine-rich strand was displaced, no termination occurred. We also saw a perfect correlation between triplex-forming ability of a given sequence and its termination efficiency.

To obtain direct evidence that the triplexes presumably formed by strand displacement cause termination, we studied DNA polymerization on otherwise single-stranded templates

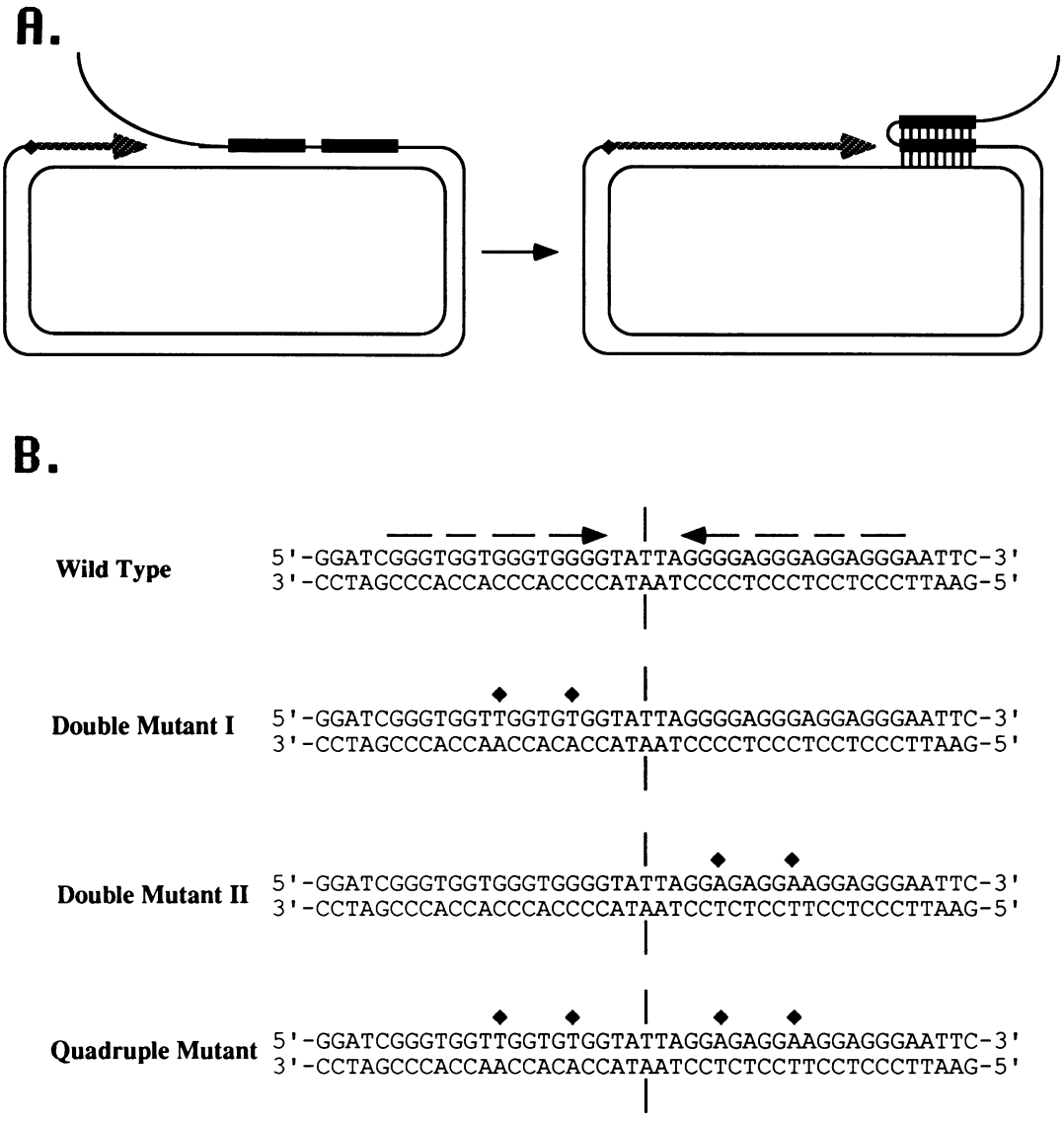

Fig. 1. Conformational suicide within H motifs during DNA polymerization. (**A**) Strand displacement during DNA polymerization may cause triplex formation. The diamond shows the original nick in the double-helical template providing a 3'-OH end for DNA polymerase; the black boxes represent two halves of a hypothetical homopurine–homopyrimidine mirror repeat involved in the formation of an intramolecular triplex; the striated arrow shows the newly synthesized DNA chain. (**B**) Sequences with varying triplex-forming potential. Vertical lines show a pseudosymmetry site. Broken arrows illustrate mirror symmetry within G-clusters. Diamonds indicate point substitutions.

containing triple-helical stretches. We found that DNA polymerase could not overcome a triplex barrier. Remarkably, triplexes of different configurations with or without purine-rich hairpins show similar effects on polymerization. We conclude, therefore, that the triplex itself rather than the complex *H configuration is necessary and sufficient for polymerase blockage. These results are the first indication that DNA polymerase can be stopped at homopurine–homopyrimidine sequences using triplex-forming oligonucleotides. The applications of these findings for anti-gene strategies are discussed.

Results

*DNA sequences with different *H-forming potential*

*H-DNA is an intramolecular triplex formed by H motifs. It is composed of a purine-rich hairpin and half of the pyrimidine-rich strand of an H motif, while the other half of the pyrimidine-rich strand remains single-stranded (Kohwi and Kohwi-Shigematsu, 1988). GGC and AAT triads are common components of the *H-DNA (Bernues *et al.*, 1990). We have recently found that TAT triads may be incorporated in *H-DNA, as well (Dayn *et al.*, 1992). In all cases GGC triads are the main units stabilizing *H-DNA. Sequences that adopt this structure must contain mirror repeated guanine stretches. The existence of a single-stranded pyrimidine-rich stretch in the structure makes possible its detection in supercoiled DNA using chemicals specific towards single-stranded pyrimidines, in particular chloroacetaldehyde (CAA) (Kohwi and Kohwi-Shigematsu, 1988).

For the purpose of our study we synthesized and cloned several sequences with varying *H-forming potential into the pBluescript plasmid (Figure 1B). In the 'wild type' sequence, guanine clusters in the purine-rich strand are mirror repeated, while the intervening thymines in one half of the sequence are reflected by adenines in the other half. As a result, when the 5'-half of the purine-rich strand folds back, GGC and TAT base triads are formed. We found

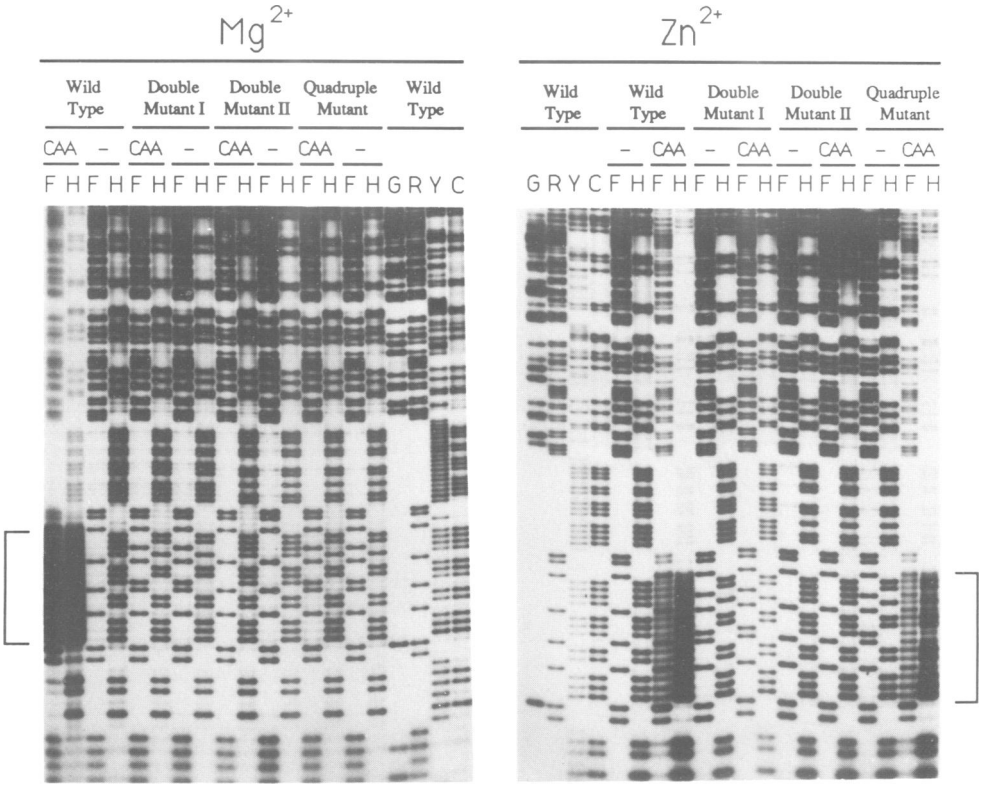

Fig. 2. Chemical modification of different H motifs in supercoiled DNA. (−), control DNA samples; CAA, modification by chloroacetaldehyde; H and F, treatment of modified samples by hydrazine or formic acid, respectively, prior to piperidine. G, R, Y and C are standard Maxam−Gilbert sequencing ladders for guanines, purines, pyrimidines and cytosines, respectively. Brackets indicate DNA bases, all located in the 3′ halves of the pyrimidine-rich strands, where hypersensitivity was observed.

previously that this sequence adopts a corresponding isoform of *H-DNA in supercoiled plasmids (Dayn *et al.*, 1992). Double mutant I has two G-to-T substitutions in the 5′-half of the purine-rich strand, while double mutant II has two symmetrical G-to-A substitutions in its 3′-half. In both cases, the mirror symmetry of the G clusters is destroyed, preventing triplex formation. Finally, the quadruple mutant combines the point substitutions of both double mutants, restoring the original pseudosymmetry and consequent triplex-forming ability.

To study the efficiency of *H formation in all four cases we used chemical footprinting of supercoiled DNAs. We had previously shown that triplex formation in the wild type sequence causes the 3′-half of the pyrimidine-rich strand to become hyperreactive to CAA (Dayn *et al.*, 1992). Therefore, we used CAA to test triplex formation in all four cases. Plasmids were CAA-modified in the presence of bivalent cations (Mg^{2+} or Zn^{2+}) followed by restriction digestion and end-labeling. Samples were then treated with either formic acid (purine reaction) or hydrazine in high salt (cytosine reaction) followed by piperidine cleavage and sequencing gel electrophoresis. CAA modification of cytosines leads to the enhancement of corresponding bands on the cytosine ladder and appearance of new bands on a purine ladder, while the adenine modification enhances corresponding bands in a purine ladder and leads to the appearance of new bands in the cytosine ladder (Kohwi and Kohwi-Shigematsu, 1988). Modification results are presented in Figure 2. As one can see, only for the wild type sequence is the 3′-part of the pyrimidine-rich strand CAA-reactive in the presence of magnesium ions. Thus, under these conditions only this sequence adopts an *H conformation. In the presence of zinc ions, however, both the wild type and the quadruple mutant are chemically hyperreactive. The double mutants show no CAA-reactivity under any conditions studied.

Thus we believe that both wild type and quadruple mutant sequences may adopt the *H-conformation in the presence of bivalent cations, while the double mutants are incapable of doing this. Clearly, the wild type sequence has a better *H-forming ability since it forms a triplex under all conditions tested and demonstrates a more prominent modification pattern than the quadruple mutant in the presence of zinc ions. The difference between the two sequences in cation requirements is most probably due to the difference in their AT content. The quadruple mutant was derived from the wild type by four GC-to-AT substitutions and is, therefore, significantly more AT-rich. It was previously found that while *H-DNA formed by $d(G)_n \cdot d(C)_n$ sequences is stabilized by Mg^{2+}, the same structure formed by $d(G\text{-}A)_n \cdot d(T\text{-}C)_n$ stretches is stable only in the presence of Zn^{2+}. Similar effects are observed for intermolecular triplexes as well (reviewed in Malkov *et al.*, 1992). Thus, changes in GC content from 100% to 50% cause a switch in cation requirements. Our data indicate that even moderate alterations in GC content (from 75% to 63%) lead to the same cation switch for a particular sequence to form *H-DNA.

Blockage of DNA polymerization by H motifs

To analyze the influence of H motifs on DNA polymerization, we studied the activity of modified T7 DNA polymerase (Sequenase) on circular, double-helical DNA templates containing a single-stranded nick (serving as a

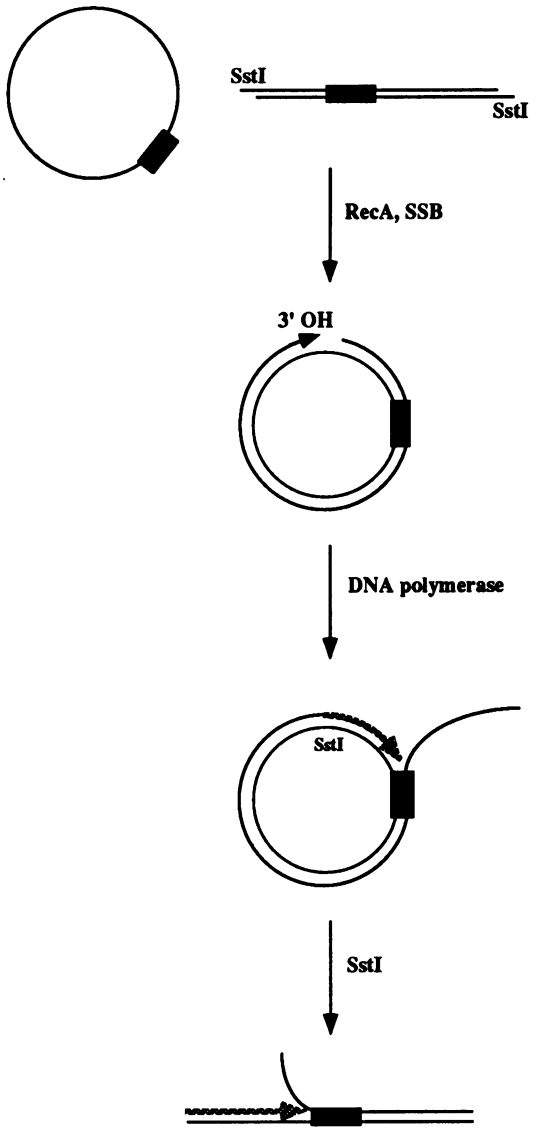

Fig. 3. Construction and sequencing strategy of open circular DNA templates. The black box represents an H motif; the striated arrow shows newly synthesized labeled DNA (see text for details).

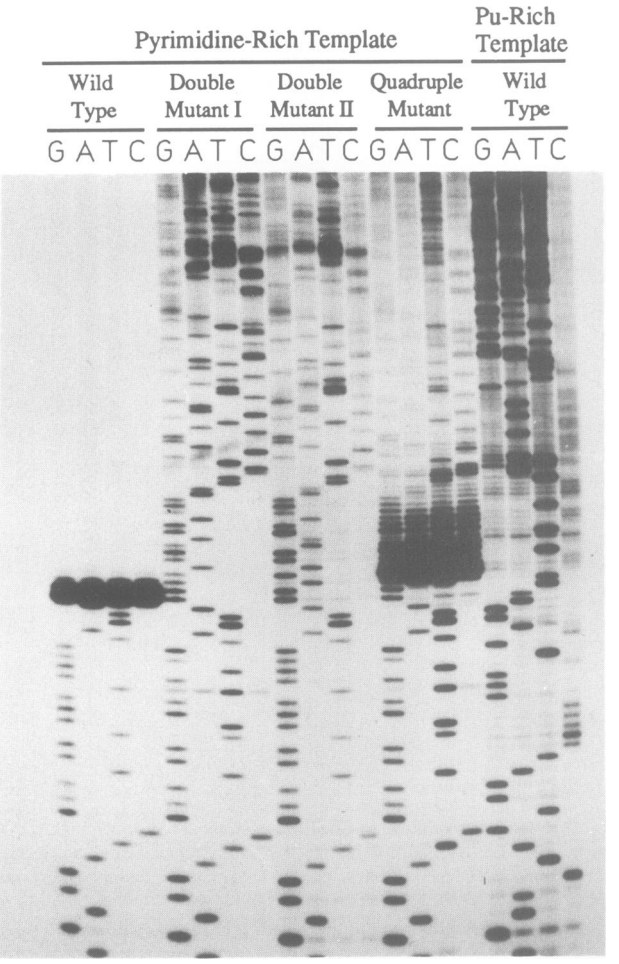

Fig. 4. The influence of triplex-forming DNA sequences on DNA polymerization in open circular templates. G, A, T and C are ddGTP, ddATP, ddTTP and ddCTP sequencing reactions respectively. Termination sites are represented by strong bands in all four sequencing ladders.

primer). The strategy for obtaining such templates is outlined in Figure 3. We first recovered single-stranded circular phage DNA containing either the purine-rich or the pyrimidine-rich strand of our inserts from *E.coli* cells carrying pBluescript plasmids as described in Materials and methods. Then RecA-mediated strand transfer to these single-stranded molecules from corresponding linear double-stranded DNAs was carried out (Soltis and Lehman, 1983). The resultant open circle contained a unique 3'-OH end available for DNA polymerase. Depending on the orientation of the triplex-forming sequence relative to the nick, either the purine- or the pyrimidine-rich strand served as a template. Note that in open circular templates, the formation of *H-DNA prior to polymerization is forbidden, since triplex extrusion depends highly upon DNA supercoiling (Dayn *et al.*, 1992).

To analyze the fine pattern of DNA polymerization we carried out DNA sequencing reactions on open circular templates as described in Materials and methods. To provide a unique reference end for newly synthesized fragments, DNA samples were digested with *Sst*I after polymerization (the *Sst*I site could only be restored in the course of DNA synthesis). The sequencing ladders are presented in Figure 4. As can be seen, the most striking results were obtained for the wild type sequence. When the pyrimidine-rich strand served as a template, we observed virtually complete termination of polymerization at the middle of this sequence. Conversely, when the purine-rich strand of the same sequence served as a template, we saw no termination. This remarkable difference between the two strands may be explained through triplex formation. Indeed, for the pyrimidine-rich template, the purine-rich strand is displaced and may form a triplex under polymerization conditions, while a displaced pyrimidine-rich strand cannot form a triplex, since pyrimidine/purine/pyrimidine triplexes are stable only under acidic pH. In the above experiments, we used a DNA sequencing protocol to locate termination sites at a base level. Thus one may wonder if termination is specific for dideoxy-NTP substrates for the polymerase. Control experiments showed that this is not the case: the same extent of termination was observed when we used 'normal' dNTP substrates in the experiment (data not shown).

Studies of the mutated sequences provided additional support for this interpretation. For both double mutants, we observed no termination of DNA polymerase even when the

pyrimidine-rich strand served as a template. With the quadruple mutant we detected prominent termination starting from the middle of the pyrimidine-rich template (Figure 4). Note, however, that the termination is not complete in the quadruple mutant, where stop signals tend to slide downstream from the center of this sequence, showing that DNA polymerase may partially overcome termination. Overall, we observed a perfect correlation between the ability of a sequence to form *H conformation and its termination ability: the wild type sequence causes a complete termination, the quadruple mutant has a lower triplex potential and weaker termination and the double mutants are unable to form triplexes and show no signs of termination.

Still, the quadruple mutant requirement for zinc, rather than magnesium ions, to adopt *H conformation (see previous section) remained a concern since we polymerized DNA in the presence of magnesium ions. One possible explanation is that during DNA polymerization, the displacement of the purine-rich strand provides a sufficiently high local concentration of the third strand at the double-helical stretch to favor triplex formation even in the presence of magnesium ions. To check this hypothesis, we studied intermolecular triplexes reflecting triple-helical portions of *H-like structures formed by strand displacement. This allowed us to estimate the efficiency of triplex formation when the third strand was present at a high molar excess. For this purpose we synthesized two oligodeoxynucleotides corresponding to the 5′-part of the purine-rich strand of the wild type (dGGGTGGTGGGTGGGG) or double mutant I (dGGGTGGTTGGTGTGG). Each of these oligonucleotides was then incubated in Sequenase buffer with linear double-stranded DNAs containing either the wild type or the quadruple mutant sequences. This gave us four oligonucleotide−template combinations corresponding to all of the triplexes formed by all our sequences.

To detect triplex formation we used a photo-footprinting approach originally described by Lyamichev *et al.* (1990). It was previously found that formation of inter- and intramolecular triplexes prevents <6−4> dipyrimidine accumulation in DNA after UV irradiation (Lyamichev *et al.*, 1990; Tang *et al.*, 1991). These photoproducts are easily detected at a sequence level since they can be cleaved by piperidine. We therefore studied the photoreactivity of the pyrimidine-rich strand in all oligonucleotide−template combinations described above. DNA mixtures were UV irradiated and treated with piperidine followed by sequencing gel electrophoresis (in all cases the pyrimidine-rich strands of double-stranded DNAs were end-labeled). The results in Figure 5 clearly demonstrate photoprotection in the pyrimidine strand in two cases: (i) when both oligonucleotide and template were of the wild type, and (ii) using mutant oligonucleotide mixed with mutant template. We did not detect triplex formation in the other two cases. The former cases correspond to triplexes formed by the wild type and quadruple-mutant sequences, while the latter cases reflect potential triplexes in both double mutants. These data show that a high molar excess (100-fold) of the third strand allows triplex formation for both wild type and quadruple mutant sequences under polymerization conditions.

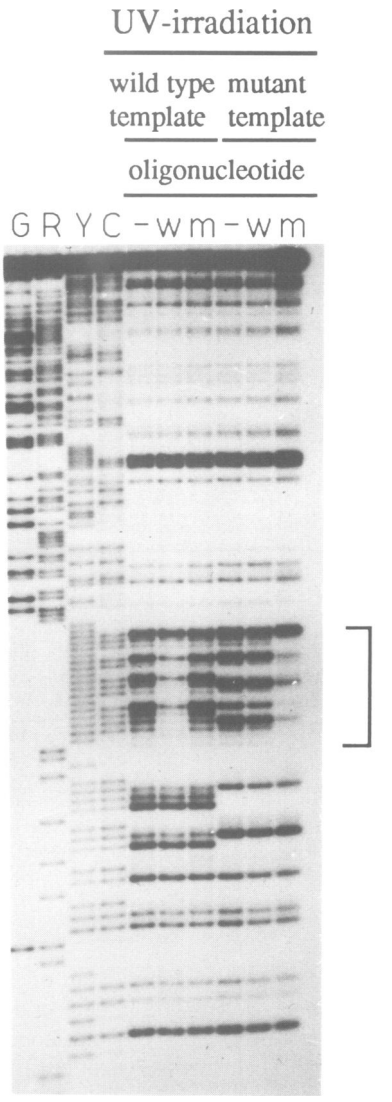

Fig. 5. Analysis of intermolecular triplex formation by photo-footprinting. Wild type and mutant templates are linear double-stranded DNAs containing wild type or quadruple mutant sequences. −, 6,4-photoproduct pattern in the absence of oligonucleotide; w, the same pattern in the presence of a 100-fold molar excess of an oligonucleotide corresponding to the 5′-part of the wild type sequence; m, 100-fold molar excess of an oligonucleotide corresponding to the 5′-part of the double mutant I. G, R, Y and C, standard Maxam−Gilbert sequencing ladder for the wild type sequence. The bracket indicates the area of photoprotection.

Blockage of DNA polymerization by triplexes within single-stranded templates

Though the above data are consistent with the hypothesis that triplex formation in the course of polymerization causes premature termination, direct proof that these structures block DNA polymerase was lacking. Another important question is whether a triplex itself or a complex *H-like configuration formed in the course of strand displacement is responsible for termination. In other words, is a hairpin loop configuration of the purine-rich strand essential for polymerase blockage?

To address these questions we studied DNA polymerization on single-stranded templates containing artificially designed triple-helical stretches reflecting the above intramolecular triplexes. Our design is illustrated in Figure 6. We obtained circular single-stranded template containing the 5′-half of the pyrimidine-rich strand of the wild type sequence. (This part of the pyrimidine-rich strand is involved in triplex formation based on our modification data.) To this template we then annealed different triplex-

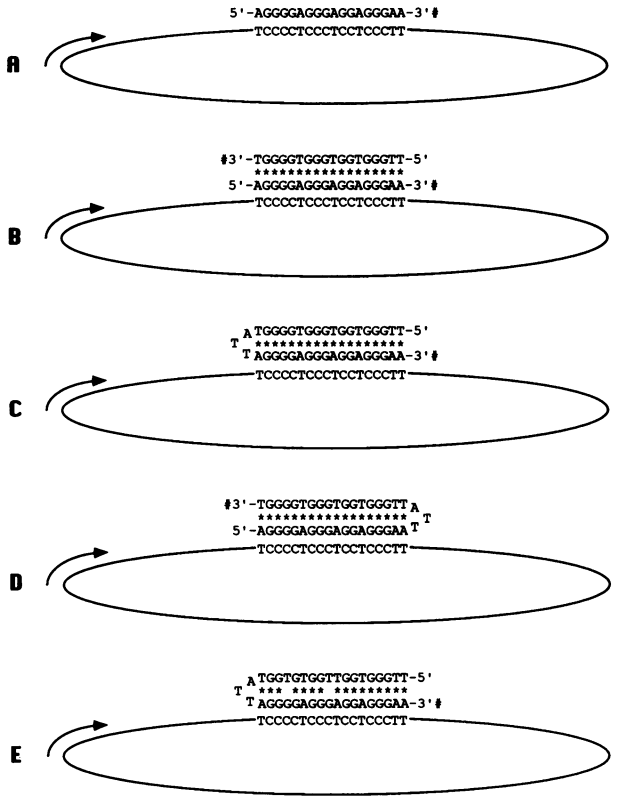

Fig. 6. Short triplexes within single-stranded templates. The 3′ amine-ON group is shown by a # mark. The arrow corresponds to the reverse primer.

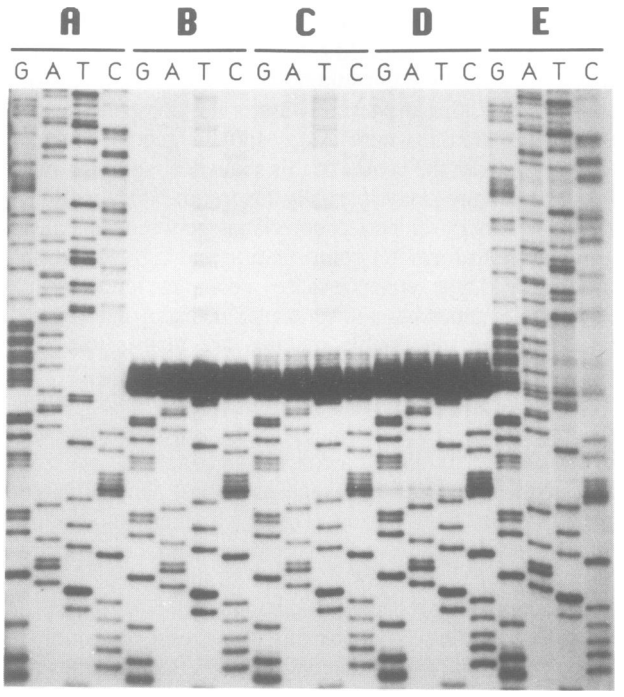

Fig. 7. Blockage of DNA polymerase by triplexes within single-stranded templates. Sections A–E corresponds to templates presented in Figure 6. G, A, T and C are ddGTP, ddATP, ddTTP and ddCTP sequencing reactions respectively. Termination sites are represented by strong bands in all four sequencing ladders.

forming oligonucleotides. In the case of A, only the complementary duplex-forming oligonucleotide was added. In B we annealed two purine-rich oligonucleotides that allow triplexes with antiparallel orientations of purine-rich strands to form. This intermolecular triplex does not contain any hairpin-loop structures. C represents the case when the above two purine-rich oligonucleotides were covalently bound via an ATT loop. This configuration is identical to an intramolecular triplex formed by the wild type sequence in supercoiled DNA and during strand displacement. D is another hairpin-looped triplex with the loop distal from DNA polymerase. Finally, E is similar to C but contains two G-to-T point substitutions disrupting triplex ability. This configuration reflects an intramolecular triplex that could be formed by the double mutant I and which, as described above, is unfavorable. The formation of triplexes in cases B–D was confirmed by footprinting assays (data not shown).

To study the pattern of DNA polymerization we annealed a reverse primer to all these templates and carried out sequencing reactions from it as described in Materials and methods. To prevent the triplex-forming oligonucleotides from serving as primers for DNA polymerase, their 3′-ends were blocked by 3′ amine-ON group (Nelson *et al.*, 1989). This was done using an Applied Biosystems oligonucleotide synthesizer and 3′ amine-ON CPG (Cruachem). The sequencing results are presented in Figure 7. First, one can see that Sequenase easily reads through the duplex area (Figure 7A), due to its strand displacement activity. Second, in cases B–D we observed almost complete termination of polymerization exactly at the junction between single- and triple-stranded DNA. Finally, destabilization of a triplex by two point substitutions allows DNA polymerase to elongate normally.

These data provide the first direct evidence that DNA polymerase cannot overcome a triplex barrier. Termination appears to be qualitatively the same for a putative intermolecular triplex and the hairpin-looped triplexes. We conclude, therefore, that a triple helix itself rather than a complex *H type configuration is responsible for polymerization blockage.

Discussion

Our results show that H-forming repeats in double-stranded templates serve as efficient terminators for modified T7 DNA polymerase. Several lines of evidence indicate that this effect is due to triplex formation rather than sequence specificity. First, we see a remarkable correlation between the triplex potential of a sequence and its terminator strength. Mutations destroying triplex formation release polymerase blockage, while compensatory mutations restoring triplex formation restore termination (Figure 4). Second, we observed characteristic strand asymmetry: termination occurred only when the purine-rich strand of an H motif was displaced during polymerization (Figure 4). Since only purine/purine/pyrimidine triplexes are stable under polymerization conditions, the displacement of the purine-rich strand may provoke their formation, preventing further elongation. Finally, when intermolecular triplexes having the same sequence and configuration as ones presumably formed by strand displacement occurred within single-stranded templates, we saw a complete termination of DNA polymerase at the junction between single- and triple-stranded DNA (Figure 7). We believe, therefore, that strand displacement in the course of polymerization promotes *H-

DNA formation, causing self-termination (Figure 1A). We call self-termination of DNA polymerases within triplex-forming sequences 'conformational suicide'.

The observed termination may be due to either thermodynamic or kinetic reasons, i.e. DNA polymerase may be unable to polymerize through the triplex part of the template, or it may just slow down at triplex boundaries. Quantitative studies of this question are yet to be provided. To get a qualitative answer, we compared termination efficiency while increasing the polymerization time from the standard 3–5 min period up to 30 min. This 10-fold increase in polymerization time did not change the termination pattern for the wild type sequence. Hence we believe that DNA polymerization through the triplex portions of a template is thermodynamically forbidden.

H motifs in double-stranded DNAs terminate not only for Sequenase but also for other polymerases with similar mechanisms of action. For example, we observed the same pattern of termination for thermophilic Vent DNA polymerase and the large fragment of DNA polymerase I (data not shown). Another group of polymerases, including Taq, allows DNA synthesis on double-stranded templates, chewing the non-template strand with its 5′→3′ exonuclease activity. As was recently demonstrated, this is an endo- rather than exonuclease activity which cleaves the non-template strand at a junction between single- and double-stranded DNA (Lyamichev *et al.*, 1993). After initial strand displacement and endonuclease cleavage, short pieces of non-template DNA strand dissociate, opening a template for polymerization. This nuclease degradation makes polymerase-driven triplex formation less likely. In agreement with this, we found that Taq polymerase does not terminate within H motifs in open circular templates. This is not due to an elevated temperature of the Taq reaction, since we used a wide range of temperatures between 37 and 70°C (data not shown). We conclude, therefore, that H motifs are suicidal for DNA polymerases capable of non-catalytic, mechanical strand displacement.

The role of H motifs in regulation of replication *in vivo* remains to be determined. It was suggested in a few studies that such motifs may account for termination of replication *in vivo*. One example came from the analysis of polyomavirus-transformed rat cells. Polyomavirus DNA integrates in a particular chromosomal site (Mendelsohn *et al.*, 1982); treatment of the transformed cells with mitomycin C leads to the amplification of virus DNA and adjacent cellular sequences (Baran *et al.*, 1987). The boundary of the amplified DNA segment lies within a homopurine–homopyrimidine stretch, $d(G\text{-}A)_{27}/d(T\text{-}C)_{27}$. It was suggested that this DNA motif could be a natural replication terminator. Supporting this hypothesis, this motif, when cloned into SV40 DNA adjacent to the origin of replication, led to a pause in replication fork progression, causing slower viral growth (Rao *et al.*, 1988). Another case came from studies of the dihydrofolate reductase (*dhfr*) locus. This locus is amplified up to 1000 times in methotrexate-resistant Chinese hamster cells due to the activity of the strong bidirectional replication origin (Milbrandt *et al.*, 1981). An unusual cluster of simple repeats, including $d(A\text{-}C)_{18}$, $d(A\text{-}G)_{21}$, $d(G)_9$ and $d(A\text{-}G)_{27}$, is located 2 kb 3′ to the origin of replication (Caddle *et al.*, 1990). This cluster was cloned in the replication vector pSV011 and the efficiency of episome replication was studied by different approaches. It was shown that when cloned on either side of the SV40 origin, it halved episomal replication, while, when placed on both sides of the origin, it blocked replication almost completely. Due to the complex nature of the cloned DNA segment, the observed effect is difficult to link with a particular element within the cluster. However, since it was found that several DNA polymerases were unable to read through $d(A\text{-}G)_n$ tracts, the authors speculated that these triplex-forming repeats may play the key role in the termination of DNA replication (Brinton *et al.*, 1991).

We believe that our model may explain these data through the formation of triplexes in the course of replication. It is important that according to the model, triplex formation is a simple consequence of strand displacement and a high level of DNA supercoiling is not required. This makes it especially attractive for eukaryotic cells where an actual torsion tension in intracellular DNA is questionable (Sinden *et al.*, 1980; Petryniak and Lutter, 1987). Detailed studies of several prokaryotic and eukaryotic replication systems revealed remarkably similar steps including initial unwinding of *ori*, synthesis of the leading strand accompanied by the displacement of a non-template strand, assembly of a complex responsible for lagging strand synthesis on a displaced strand and, finally, coordinated synthesis of both DNA strands (reviewed in Kornberg and Baker, 1992). Polymerase-driven triplex formation could potentially impair leading and lagging strand synthesis. If it indeed occurs, leading strand synthesis should be prematurely terminated, while the template for lagging strand synthesis is folded into an unusual DNA conformation. We realize, of course, that though our system resembles the replication fork in some detail, it contains only DNA polymerase rather than the complex replication machinery. Experiments are in progress to study the influence of helicases, single-strand DNA binding protein (SSBs) and other accessory proteins on triplex-promoted termination.

An increasing number of studies use oligonucleotide-directed triplex formation for targeting genetic processes. The majority of those studies are concerned with transcription. In many cases, triplex-forming oligonucleotides served as artificial repressors preventing the binding of transcriptional regulators. The very first results were obtained with the human c-*myc* promoter. It was found that the binding of a purine-rich oligonucleotide to the imperfect homopurine–homopyrimidine sequence that is located 100 bp upstream of the P1 promoter blocks transcription from this promoter both *in vitro* and *in vivo* (Cooney *et al.*, 1988; Postel *et al.*, 1991). Similar observations were made for the methallothionein gene promoter. In this case a homopyrimidine oligonucleotide formed a triplex with the upstream portion of the promoter preventing the binding of the transcriptional activator Sp1 (Maher *et al.*, 1989). This in turn drastically reduced the promoter's activation in a cell-free transcription system (Maher *et al.*, 1992). Quite recently, a triplex-forming oligonucleotide–intercalator conjugate was shown to act as a transcriptional repressor of the interleukin-2 receptor gene both *in vitro* and *in vivo* (Grigoriev *et al.*, 1992, 1993). This was due to the fact that the formation of a triplex additionally stabilized by cross-linking to target DNA efficiently prevented the binding of the transcriptional activator NF-κB. Different mechanisms of transcriptional inhibition have also been revealed. The pBR322 *bla* gene contains a 13 bp homopurine–homopyrimidine target just immediately downstream from the transcriptional start site. A 13mer

homopyrimidine oligonucleotide forming an intermolecular triplex with this target hindered initiation of transcription by *E.coli* RNA polymerase *in vitro* (Duval-Valentin *et al.*, 1992). Finally, Young *et al.* (1991) followed eukaryotic RNA polymerase II transcription *in vitro* from the adenovirus major late promoter. The transcribed portion of DNA contained a 15 bp homopurine−homopyrimidine tract that formed an intermolecular triplex with the homopyrimidine oligonucleotide. It was found that when the oligonucleotide was added prior to RNA polymerase, a significant portion of the transcripts were truncated. Thus, triplex formation may inhibit all stages of the transcriptional cycle, i.e. formation of the active promoter complex, initiation and elongation.

Surprisingly, analogous studies targeting DNA replication are practically absent. The only published data concern the inhibition of the growth of SV40 by an octathymidilate covalently linked to an acridine derivative (Birg *et al.*, 1990). *In vitro* this complex forms a triplex with the dA_8 stretch located in the minimal origin of replication adjacent to the T-antigen binding site. *In vivo* it inhibits the cytopathic effect of SV40 in CV1 cells presumably due to the inhibition of T antigen binding or unwinding activity. In the course of our studies we found, for the first time, that elongation of DNA polymerization can be blocked at specific sequences by triplex-forming oligonucleotides (Figure 7). It is well established that such oligonucleotides provide highly sequence-specific recognition of target DNA (Strobel and Dervan, 1991; Strobel *et al.*, 1991). We speculate, therefore, that our results point to a novel opportunity to stop DNA replication at defined chromosomal sequences by means of triplex-forming oligonucleotides. Such artificial terminators are interesting for their therapeutic potential as well as for understanding the regulation of DNA replication.

Materials and methods

Enzymes

Sequenase version 2.0 T7 DNA polymerase (E.C. 2.7.7.7), *E.coli* RecA protein and *E.coli* SSB were obtained from US Biochemicals. Restriction enzymes and enzymes for end-labeling of DNA were obtained from Bethesda Research Laboratories.

Oligonucleotides

Oligonucleotides were synthesized on an ABI High Throughput DNA−RNA synthesizer Model 394 as described in the user's manual. They were deprotected by incubation in concentrated ammonium hydroxide for 15 h at 55°C, then concentrated in a SpeedVac and precipitated with 2 vols of a 2 M solution of $LiClO_4$ in acetone. Dried pellets were dissolved in 0.5 ml of TE buffer and additionally purified on NAP-5 columns (Pharmacia).

To obtain oligonucleotides that cannot serve as substrates for DNA polymerases (Figure 6), we introduced a primary aliphatic amine at their 3′ terminal nucleotides. This was done by an automatic conventional solid phase DNA synthesis using a 3′ amine-ON CPG prepacked column (Cruachem) which transfers a primary amine to the 3′ terminus of a synthesized oligonucleotide without changing any chemistry or adding extra steps.

Plasmid construction

Oligonucleotides corresponding to H-forming sequences presented in Figure 1B were cloned in the pUC19 polylinker between the *Bam*HI and *Eco*RI sites. In order to obtain single-stranded DNA, these sequences were further recloned in two orientations in the phagemid pBluescript SK(−) (Stratagene). Either the purine-rich or the pyrimidine-rich strand could be rescued in a single-stranded state after the isolation of phage DNA. Supercoiled plasmid DNA was isolated by standard alkali lysis followed by two equilibrium centrifugations in a cesium chloride−ethidium bromide gradient.

Isolation of single-stranded DNA

Insertion derivatives of the phagemid pBluescript SK(−) were transformed into the *E.coli* XL-1 strain. Fresh transformants were inoculated into 50 ml of 2×YT medium containing 10^8-10^9 p.f.u./ml of VCS helper phage. After 2 h of incubation at 37°C, kanamycin was added up to 70 μg/ml and cultures were subsequently incubated for 16 h with vigorous aeration. Phage particles were separated from cells by centrifugation at 17 000 *g* for 10 min followed by precipitation with 4% PEG in 0.7 M NH_4CH_3COO. Phage particles were resuspended in 300 μl H_2O, and DNA was isolated by phenol−chloroform extraction followed by ethanol precipitation. DNA was rinsed with 70% ethanol and diluted in 60 μl of H_2O.

Chemical footprinting of DNA

Samples containing 10 μg of supercoiled DNA were modified in 25 mM sodium cacodylate, pH 7.1, in the presence of 4 mM $MgCl_2$ or 1 mM $ZnCl_2$ by chloroacetaldehyde as previously described (Dayn *et al.*, 1992). Control DNA samples were incubated under the same conditions without CAA. DNAs were then digested with *Hin*dIII, end-labeled, incubated with either formic acid or hydrazine, treated with piperidine and loaded on to a sequencing gel.

Photo-footprinting of DNA

10 μg of plasmid DNA were linearized with *Hin*dIII and end-labeled using the Klenow fragment of DNA polymerase I and [α-^{32}P]dCTP, followed by spin chromatography on Sephadex G50. Sequenase buffer was added up to 1 × concentration, triplex-forming oligonucleotides were added up to 5 μM, and samples were irradiated using a UV Stratalinker 2400 at 3000 μkW/cm^2 for 2 min. DNA samples were then digested with *BglI*. 250 bp end-labeled fragments were eluted from an agarose gel, treated with piperidine, and loaded on a sequencing gel.

Construction of open circular templates

Single-stranded DNAs containing the purine- or pyrimidine-rich strand of the original and mutated sequences were isolated as described above. 0.8 μg of single-stranded DNA was then mixed with 1.2 μg of the corresponding double-stranded DNA, linearized by *SstI*. A strand transfer reaction was carried out with 4.4 μg of RecA protein in 480 μl of 5% glycerol, 1 mM DTT, 25 mM Tris−HCl (pH 7.2), 10 mM $MgCl_2$ for 10 min at 37°C. *E.coli* SSB protein and ATP were added up to 61 μM and 4 mM, respectively, and incubation continued for another 30 min. The reaction was terminated by adding 0.5% SDS, 20 mM EDTA, followed by proteinase K treatment and phenol extraction. Open circular DNAs were then isolated from an agarose gel by elution on an ion-exchange membrane NA45 (Schleicher & Schuell).

Sequencing of open circular DNA

0.1 μg of nicked DNA was dissolved in 10 μl of 1 × Sequenase buffer (US Biochemicals) and labeled in the presence of 1 ml 0.1 M DTT, 1 μl [α-^{32}P]dATP (3000 Ci/mmol, Amersham), 2 μl labeling mix (US Biochemicals) and 13 U of Sequenase version 2 for 3 min at room temperature. Termination was then carried out for 5 min at 37°C according to the manufacturer's protocol (US Biochemicals), followed by 15 min of polymerase inactivation at 65°C. To provide a unique reference end for all synthesized fragments, DNA samples were digested with *SstI* (note that the *SstI* site was restored in the course of DNA synthesis) and loaded on to a sequencing gel.

Sequencing of single-stranded DNA with short triplexes

0.5 μg of single-stranded DNAs containing pyrimidine-rich targets were mixed with 100−200 pmol of duplex- or triplex-forming oligonucleotides (Figure 6) and incubated in Sequenase buffer (40 mM Tris−HCl pH 7.5, 50 mM NaCl, 20 mM $MgCl_2$) for 15 min at 37°C. 15 pmol of the 'reverse' primer was used for DNA sequencing according to the Sequenase version 2.0 sequencing protocol (US Biochemical) with the following modifications. Labeling was done in the presence of 230 nM dNTPs (N = G, C or T) supplemented with 5 μCi of [α-^{32}P]dATP (Amersham) for 2−3 min at room temperature. Then dNTPs were added up to 33.3 μM and ddNTPs up to 3.3 μM, followed by polymerization for 10 min at 37°C.

Acknowledgements

We thank Alexander Rich, Claude Hélène, Donald Crothers, Charles Cantor, James Dahlberg, Maxim Frank-Kamenetskii, Richard Morgan, Jonathan Widom and Dan Gottschling for helpful discussions, Sergey Gryaznov for his advice to use 3′ amine-ON CPG columns, and Angela Tyner and Randal Cox for critical reading of the manuscript. Supported by grants MG-25 from the American Cancer Society and R55GM46405-01A1 from the National Institutes of Health to S.M.M.

References

Baran,N., Lapidot,A. and Manor,H. (1987) *Mol. Cell. Biol.*, **7**, 2636–2640.

Baran,N., Lapidot,A. and Manor,H. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 507–511.

Bernues,J., Beltran,R., Casanovas,J.M. and Azorin,F. (1989) *EMBO J.*, **8**, 2087–2094.

Bernues,J., Beltran,R. and Azorin,F. (1990) *Nucleic Acids Res.*, **18**, 4067–4073.

Birg,F., Praseuth,D., Zerial,A., Thuong,N.T., Asseline,U., Le Doan,T. and Helene,C. (1990) *Nucleic Acids Res.*, **18**, 2901–2908.

Brinton,B.T., Caddle,M.S. and Heintz,N.H. (1991) *J. Biol. Chem.*, **266**, 5153–5161.

Caddle,M.S., Lussier,R.H. and Heintz,N.H. (1990) *J. Mol. Biol.*, **213**, 19–33.

Cooney,M., Czernuszewicz,G., Postel,E., Flint,S.J. and Hogan,M.E. (1988) *Science*, **241**, 456–459.

Dayn,A., Samadashwily,G.M. and Mirkin,S.M. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 11406–11410.

Duval-Valentin,G., Thuong,N.T. and Helene,C. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 504–508.

Grigoriev,M., Praseuth,D., Robin,P., Hemar,A., Saison-Behmoaras,T., Dautry-Varsat,A., Thuong,N.T., Helene,C. and Harel-Bellan,A. (1992) *J. Biol. Chem.*, **267**, 3389–3395.

Grigoriev,M., Praseuth,D., Guieysse,A.L., Robin,P., Thuong,N.T., Helene,C. and Harel-Bellan,A. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 3501–3505.

Hanvey,J.C., Klysik,J. and Wells,R.D. (1988) *J. Biol. Chem.*, **263**, 7386–7396.

Htun,H. and Dahlberg,J.E. (1988) *Science*, **241**, 1791–1796.

Johnston,B.H. (1988) *Science*, **241**, 1800–1804.

Kohwi,Y. and Kohwi-Shigematsu,T. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 3781–3785.

Kornberg,A. and Baker,T.A. (1992) *DNA Replication*. 2nd edn. W.H.Freeman and Company, New York.

Lapidot,A., Baran,N. and Manor,H. (1989) *Nucleic Acids Res.*, **17**, 883–900.

Lundquist,R.C. and Olivera,B.M. (1982) *Cell*, **31**, 53–60.

Lyamichev,V.I., Frank-Kamenetskii,M.D. and Soyfer,V.N. (1990) *Nature*, **344**, 568–570.

Lyamichev,V.I., Brow,M.A.D. and Dahlberg,J.E. (1993) *Science*, **260**, 778–783.

Maher,L.J., Wold,B. and Dervan,P.B. (1989) *Science*, **245**, 725–730.

Maher,L.J., Dervan,P.B. and Wold,B. (1992) *Biochemistry*, **31**, 70–81.

Malkov,V.A, Soyfer,V.N. and Frank-Kamenetskii,M.D. (1992) *Nucleic Acids Res.*, **20**, 4889–4895.

Mendelsohn,E., Baran,N., Neer,A. and Manor,H. (1982) *J. Virol.*, **41**, 192–209.

Milbrandt,J.D., Heintz,N.H., White,W.C., Rothman,S.M. and Hamlin,J.L. (1981) *Proc. Natl Acad. Sci. USA*, **78**, 6043–6047.

Mirkin,S.M., Lyamichev,V.I., Drushlyak,K.N., Dobrynin,V.M., Filippov,S.A. and Frank-Kamenetskii,M.D. (1987) *Nature*, **330**, 495–497.

Nelson,P.S., Frye,R.A. and Liu,E. (1989) *Nucleic Acids Res.*, **17**, 7187–7194.

Petryniak,B. and Lutter,L.C. (1987) *Cell*, **48**, 289–295.

Postel,E.H., Flint,S.J., Kessler,D.J. and Hogan,M.E. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 8227–8231.

Rao,S., Manor,H. and Martin,R.G. (1988) *Nucleic Acids Res.*, **16**, 8077–8094.

Sinden,R.R., Carlson,J.O. and Pettijohn,D.E. (1980) *Cell*, **21**, 773–783.

Soltis,D.A. and Lehman,I.R. (1983) *J. Biol. Chem.*, **258**, 6073–6077.

Strobel,S.A. and Dervan,P.B. (1991) *Nature*, **350**, 172–174.

Strobel,S.A., Doucette-Stamm,L.A., Riba,L., Housman,D.E. and Dervan,P.B. (1991) *Science*, **254**, 1639–1642.

Tang,M., Htun,H., Cheng,Y. and Dahlberg,J.E. (1991) *Biochemistry*, **30**, 7021–7026.

Voloshin,O.N., Mirkin,S.M., Lyamichev,V.I., Belotserkovskii,B.P. and Frank-Kamenetskii,M.D. (1988) *Nature*, **333**, 475–476.

Young,S.L., Krawczyk,S.H., Matteucci,M.D. and Toole,J.J. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 10023–10026.

Received on June 28, 1993; revised on September 2, 1993