(dT—dC)n and (dG—dA)n tracts arrest single stranded DNA replication in vitro

Aviva Lapidot, Nava Baran and Haim Manor

Department of Biology, Technion-Israel Institute of Technology, Haifa 32000, Israel

Received December 8, 1988; Accepted January 6, 1989

ABSTRACT

Previous in vivo studies have indicated that (dT—dC)n.(dG—dA)n tracts (referred to here as (TC)n.(GA)n), which are widely dispersed in vertebrate genomes, may serve as pause or arrest signals for DNA replication and amplification. To determine whether these repeat elements act as stop signals for DNA replication in vitro, single stranded DNAs including (TC)n or (GA)n tracts of various lengths, were prepared by cloning such tracts into phage M13 vectors, and were replicated with the Klenow fragment of the E. coli DNA polymerase 1, or with the calf thymus DNA polymerase α, by extension of an M13 primer. Gel electrophoresis of the reaction products revealed that the replication was specifically arrested around the middle of both (TC)n and (GA)n tracts of \( n > 16 \). However, whereas in the (TC)n tracts the arrests were less prominent at pH=8.0 than at pH=6.5-7.5, and were completely eliminated at pH=8.5, the arrests in the (GA)n tracts were stronger at the higher pH values. These results, and previous data, suggest that the arrests were caused by formation of unusual DNA structures, possibly triple helices between partially replicated (TC)n or (GA)n tracts, and unreplicated portions of these sequences.

INTRODUCTION

The repeating dinucleotide sequence elements (dT—dC)n.(dG—dA)n (henceforth referred to as (TC)n.(GA)n) are widely dispersed in eukaryotic genomes (1-6). Detailed studies of the abundance and the genomic distribution of these elements have been carried out so far in a number of primates, rodents and the chicken (7). Quantitative hybridization assays have shown that in the genomes of man and monkey, and in the chicken genome, the abundances of (TC)n.(GA)n tracts are 0.05\% to 0.07\% of all DNA sequences, while in the mouse, rat and Chinese hamster genomes, their abundances are 0.30\% to 0.40\%. Plaque hybridization assays of genomic libraries, and a genbank survey, have indicated that these elements are spread throughout the genomes of primates and rodents, and that they map primarily within introns and intergenic regions of these orders (7).

Other studies have indicated that (TC)n.(GA)n tracts are sensitive to single strand specific endonucleases and to chemical reagents that do not react with B-DNA (4,6,8-11). These observations have led to the inference that the (TC)n.(GA)n sequences may assume a non-B, non-Z, helical structure (8-11). Of various models proposed for this structure, the H-DNA (12) which contains triple stranded and single stranded regions, has recently gained strong experimental support (13-17).

We have recently suggested that (TA)n.(GA)n tracts might act as pause or arrest signals for DNA replication and amplification (18). This suggestion was based on a discovery made in the course of studies of polyoma virus (Py)
induction in a line of Py transformed rat cells, designated LPT (19). In LPT cells, multiple rounds of replication of the chromosomally associated Py DNA and flanking cellular sequences can be induced in a single cell cycle by exposure to carcinogens (20). The replication was found to be arrested within a cell DNA segment containing a (TC)$_{27}$-(GA)$_{27}$ tract, as well as inverted and direct repeats of six base pairs (18,20). Subsequently, a 210 bp segment including these sequence elements, has been cloned into Simian Virus 40, and found to cause a pause in fork progression during replication of the viral DNA in monkey cells (21).

To further examine the ability of (TC)$_n$-(GA)$_n$ tracts to arrest DNA replication, we have subcloned segments including the genomic (TC)$_{27}$-(GA)$_{27}$ tract, or synthetic (TC)$_n$-(GA)$_n$ sequences of various lengths, into M13 phage vectors (22), and carried out in vitro replication assays of single stranded DNA molecules prepared from these phage recombinants. These reactions were catalysed by the Klenow fragment of the E. coli DNA polymerase I, or by the calf thymus DNA polymerase α, and were initiated at M13 primers. Here, we present the results of these assays, which, in accordance with the in vivo data, showed that replication of the single stranded DNA templates was specifically arrested at the genomic (TC)$_{27}$ and (GA)$_{27}$ tracts, and at synthetic (TC)$_n$ and (GA)$_n$ tracts of $n \geq 16$. Furthermore, this approach allowed a detailed analysis of the patterns of arrests that occur at these repeat elements.

**MATERIALS AND METHODS**

**M13 Constructions**

mpTC27 and mpGA27: A plasmid including the (TC)$_{27}$-(GA)$_{27}$ tract and the 6 bp repeat elements (18), was cut either with the restriction enzymes SmaI and Rsal, or with Sau3A and Rsal. The SmaI/Rsal and the Sau3A/RsaI fragments, respectively, were cloned into the SmaI restriction site, or between the BamHI and the SmaI restriction sites, of the phage vectors M13mpl8 and M13mpl9 (22,23). The recombinant phages were screened by plaque hybridization (24,25), for the presence of the (TC)$_{27}$ tract, or the (GA)$_{27}$ tract, using as hybridization probes synthetic (TC)$_g$ and (GA)$_g$ oligonucleotides that had been labeled with $\gamma^32P$PO$_4$ in a reaction catalyzed by T4 polynucleotide kinase (26). Phage clones that hybridized with these probes were characterized by sequencing (27). Two phages were used here: mpTC27 is an M13mp19 recombinant including a genomic Sau3A/RsaI 190 bp fragment; this phage yields the strand containing the (TO27 tract and the 6 bp repeat. mpGA27 is an M13mp18 recombinant including a genomic SmaI/Rsal 210 bp fragment; this phage yields the strand containing the (GA)$_{27}$ tract and the 6 bp repeat.

mpTC9-A and mpGA9-T: The synthetic unphosphorylated (TC)$_g$ and (GA)$_g$ preparations were mixed at a concentration of 25 µg/ml each, in a buffer containing 70 mM tris-HCl pH=7.6, 10mM MgCl$_2$ and 5mM dithiothreitol. The mixture was annealed 18 hrs at 50°C. The double stranded oligonucleotides were cloned into the SmaI site of the vector M13mp18, as described (22,23). Of the recombinant phages, several contained (TC)$_n$ and (GA)$_n$ tracts with $n \geq 8$. Two phages were used here- mpTC9-T containing a (TC)$_g$-T tract and mpAG9-A containing an (AG)$_g$-A tract.

mpTC17-C-TC2, mpTC23-T, mpTC31, mpGA16: The (TC)$_g$ and (GA)$_g$ oligonucleotide preparations were phosphorylated with polynucleotide kinase and were annealed, ligated and cloned into the phage vectors M13mp18 and M13mp19, as described above. The recombinant phages used here were: mpTC23-T - an M13mp19 derivative containing a (TC)$_{23}$-T tract. mpTC17-C-TC2, mpTC31 and mpGA16 - M13mp18 derivatives containing (TC)$_{17}$-(TC)$_2$, (TC)$_{31}$ and (GA)$_{16}$, respectively. Assays of single stranded DNA replication catalysed by the Klenow fragment of the E. coli DNA polymerase I. Most assays were carried out under conditions generally employed for
sequencing (27). In these assays, single stranded DNA purified from each of the M13 phage recombinants was annealed with the 17-mer sequencing primer for 2 hrs at 60°C, in 10 µl of a buffer containing 15 mM tris-HCl pH=8.0, 7.5 mM MgCl2, 50 µg/ml of phage DNA and 0.50 µg/ml of primer. Replication was performed in 5 µl of a buffer containing 5 mM tris HCl pH=8.0 and 2.5 mM MgCl2, 20 µg/ml of primer-template complex and 1.2 µM of (α-32P)-dATP. In addition, the following concentrations of dNTPs and ddNTPs were used in the various reactions (unless specified otherwise): 30 µM dCTP, 30 µM dGTP, 30 µM dTTP and 40 µM ddATP (A lanes), or 40 µM dGTP, 40 µM dTTP, 2.0 µM dCTP and 40 µM ddCTP (C lanes), or 40 µM dCTP, 40 µM dTTP, 2.0 µM dGTP and 80 µM ddCTP (G lanes), or 40 µM dCTP, 40 µM dGTP, 2.0 µM dCTP and 125 µM ddTTP (T lanes), or 1.0 µM dCTP, 1.0 µM dGTP, 1.0 µM dTTP (−) lanes. The polymerase/template ratios are specified in the legends to the figures in terms of units/µg DNA (1.0 unit of Klenow polymerase is defined as the amount of enzyme that incorporates 10 nmoles of each of the four dNMPs into acid insoluble material in 30 min at 37°C; the enzyme preparations were obtained from New England Biolabs, Amersham and from Dr L. Loeb, University of Washington, Seattle). The reaction mixtures were incubated 20 min at 37°C; then, 150 µM of each of the four dNTPs were added and the mixtures were further incubated for 15 min at 37°C. The reactions were stopped by addition of EDTA, xylene cyanol, bromophenol blue and formamide at final concentrations of 10mM, 0.05%, 0.05% and 50%, respectively. The mixtures were loaded on sequencing gels and electrophoresed, as described below.

Revolution of single stranded DNA catalysed by the calf thymus DNA polymerase α.

Polyuridine kinase was used to label the 17-mer sequencing primer with 32P (26); the primer was then annealed 2 hrs at 60°C with the single stranded DNA templates in 5 µl of a solution containing 25 mM tris HCl pH=7.6, 5 mM MgCl2, 2.5 mM dithiothreitol, 0.05 mM spermidine, 0.05 mM EDTA, 75 µg/ml of DNA template and 1.25 µg/ml of primer. Replication was carried out 90 min at 37°C in 25 µl of a solution containing 50 mM tris HCl pH=7.4, 10 mM MgCl2, 2 mM dithiothreitol, 50 µM each of dATP, dTTP, dGTP and dCTP, 5–10 µg/ml of primer-template and 10 units/ml of pol α (1 unit of pol α is an amount of enzyme that incorporates 1 nmoles of dNMP into acid insoluble form in 1 hr at 37°C (28); pol α was obtained from Dra L. Loeb and F. Perrino). Then, an equal volume of a solution containing 100 mM tris HCl pH=8.0 and 100 mM EDTA and 50 µg/ml of yeast tRNA, was added to each sample. The nucleic acids were precipitated with ethanol, washed and dried, as described (29). Dried pellets were resuspended in 5 µl of a solution containing 0.05% xylene Cyanol, 0.05% bromophenol blue, 10 mM EDTA, 5 mM tris HCl pH=8.0 and 50% formamide. The samples were loaded on sequencing gels and electrophoresed, as described below.

Gel electrophoresis and autoradiography.

Sequencing gels containing 6% polyacrylamide and 8M urea were prepared, using sharktooth combs (BRL). The gels were 25 cm long and 0.40 mm thick. Samples were heated 3 min at 90°C before being loaded at 3µl/slot. Electrophoresis was carried out at 1300 volts for 45 min to 3 hr, then the gels were dried and exposed to an X-ray film, as described (29).

RESULTS

Arrest of Replication of Single Stranded DNA Molecules at the (TC)27 and the (GA)27 Tracts.

Fragments including the genomic (TC)27-(GA)27 segment and the six bp repeat elements (18; see the Introduction), have been subcloned into the phage vectors M13mp18 and M13mp19 (22). Single stranded DNAs containing either the (TC)27 tract, or the (GA)27 tract, were prepared from the phage recombinants and were replicated with the Klenow fragment of the E. coli DNA polymerase I by extension of a 17-mer M13 primer (27).

Fig. 1A (lanes T,G,C,A) shows the gel patterns obtained in replication
Fig. 1. Arrest of single stranded DNA replication at the genomic (TC)$_{27}$ and (GA)$_{27}$ tracts. (A) The Klenow polymerase was used, at an enzyme/template ratio of 10 units/µg DNA, to replicate single stranded DNA prepared from the phage recombinant mpTC27. The reactions were carried out in the presence of ddTTP (lane T), or ddGTP (lane G), or ddCTP (lane C), or ddATP (lane A), or in the absence of any ddNTP (lane (-)). The samples were electrophoresed in sequencing gels and the gels were autoradiographed. The arrow shows the direction of electrophoresis. The number of nucleotides between the end of the primer and the beginning of the (TC)$_{27}$ tract in the template was 95. (B) The sequence of a DNA segment including the (TC)$_{27}$(GA)$_{27}$ tract is shown. The numbers represent nucleotide residues, No. 1 in each strand being the first residue of the (TC)$_{27}$ or the (GA)$_{27}$, that the polymerase encounters during replication of this strand. The sequence is interrupted, as indicated. (C) The Klenow polymerase was used, at an enzyme/template ratio of 7 units/µg DNA, to replicate single stranded DNA prepared from the phage recombinant mpGA27. The notations are the same as those used in section A. The number of nucleotides between the end of the primer and the beginning of the (GA)$_{27}$ tract in this template was 123.
Fig. 2. Effects of pH and ionic strength on the patterns of arrest at the (TC)27 tract. (A) Effects of pH. The Klenow polymerase was used at an enzyme/template ratio of 10 units/μg DNA to replicate single stranded DNA prepared from the phage recombinant mpTC27. The reactions and the analysis were carried out as described in Materials and Methods, except for the pH which was varied as follows: lanes A and 1- pH=7.0, lane 2- pH=7.5, lane 3- pH=8.0, lane 4- pH=8.5 (The pH values were determined at 20°C). The assay shown in lane A was carried out in the presence of ddATP. The assays shown in lanes 1-4 were carried out in the absence of any ddNTP. The two arrows designate the boundaries of the (GA)27 tract in the newly synthesized DNA, and the arrowhead designates the center of the arrest sites found within the (GA)27. The direction of electrophoresis was from top to bottom. (B) Effects of ionic strength. Replication assays were carried out as described in A., except that the pH was 8.0 in all assays, and NaCl was added to the reaction mixtures at the following concentrations: panel a- no NaCl; panel b- 25 mM NaCl; panel c- 50 mM NaCl; panel d- 75 mM NaCl. The notation A and (-) above the lines designate reactions performed in the presence of ddATP and the absence of ddNTPs, respectively. (C) Replication assays were performed as described in B., except that the pH was 7.5.
assays of the strand including the (TC)$_{27}$ tract, carried out under conditions generally employed for sequencing, in the presence of dideoxy T, G, C and A, respectively. This gel has been inverted. Fig. 1C (lanes A, C, G, T) presents the corresponding assays of the strand including the (GA)$_{27}$ tract. Fig. 1B shows a part of the sequence found in these gels, including the (TC)$_{27}$-(GA)$_{27}$ tract. The top and the bottom lines drawn from the sequence in Fig. 1B to the gel in Fig. 1A indicate the boundaries of the (TC)$_{27}$ tract in the template strand, and the (GA)$_{27}$ sequence in the newly synthesized strand; the nucleotides are given the numbers 1-54, in the direction of DNA synthesis. Other lines designate bands that were clearly not generated by termination of the growing chains with dideoxynucleotides, but instead, represent arrest sites for the polymerase. The lane designated (−) in panel A shows the gel profile obtained in a replication assay carried out in the absence of any ddNTP. It can be seen that the pattern of arrests generated in this assay was similar to the patterns observed in the assays performed with the dideoxynucleotides. The strongest arrests occurred at the T residues No. 28 and 30, that is one and three residues, respectively, beyond the middle of the (TC)$_{27}$ tract, and arrests of lesser strengths occurred at several additional T residues on both sides of these sites. The replication was also less efficiently arrested at C residues found in this region, as revealed by the rather weak bands observed between the stronger bands.

As shown in Fig. 1C, replication of the strand including the (GA)$_{27}$ tract was also arrested at several sites around the middle of this tract; the first arrest occurred at the same distance (26 nucleotides) from the site of entry of the polymerase into the (GA)$_{27}$ tract, as that found in the replication of the other strand. However, in this reaction the polymerase was arrested more efficiently at the G residues and less efficiently at the A residues.

The assays presented in Fig. 1 were carried out in a low ionic strength buffer at pH=8.0. The next series of experiments were designed to find out whether variations in the pH and the ionic strength affect the patterns of arrest found within the (TC)$_{27}$ and the (GA)$_{27}$ tracts. Fig. 2A shows the gel profiles obtained in replication assays of the strand including the (TC)$_{27}$ tract, carried out (in a low ionic strength buffer) in the presence of ddATP at pH=7.0 (lane A), or in the absence of any ddNTP, at pH values of 7.0, 7.5, 8.0 and 8.5, respectively (lanes 1, 2, 3 and 4). A clear dependence on the pH is evident—the arrests were stronger and more numerous at the pH values of 7.0 and 7.5 than at pH=8.0. No arrests were detectable at pH=8.5. Assays performed at pH=6.5 gave arrest patterns that were similar to those obtained at pH=7.0 (not shown). In contrast, similar assays of the strand including the (GA)$_{27}$ tract (fig. 3A) revealed that the arrests at the GA$_{27}$ tract were stronger at pH=8.5 than at pH=8.0, and their strengths decreased further at pH=7.0.

Figure 2B shows replication assays of the strand including the (TC)$_{27}$ tract carried out at pH=8.0 and at various concentrations of NaCl. It can be seen that the arrest reactions were less efficient at the higher ionic strengths and were almost totally inhibited at 75mM NaCl. Yet, similar ionic strength dependence assays performed at pH=7.5, showed that at this pH value strong arrests occurred even at 75 mM NaCl (Fig. 2C). Fig. 3B shows replication assays of the strand including the (GA)$_{27}$ tract, carried out at pH=8.0 and at various concentrations of NaCl. Clearly, the dependence of the arrests on ionic strength resembles the dependence observed in the assays of the other strand at pH=8.0 (fig. 2B). Unlike the strand including the (TC)$_{n}$ tract, similar assays of this strand performed at pH=7.5, gave the same dependence on ionic strength as that found at pH=8.0 (not shown). It should be noted that the dependence of the arrests on pH and ionic strength was not due to inactivation of the enzyme; for the sequencing reactions in the A and the T lanes, and the synthesis of high molecular weight DNA (that remained at the top of the gel), were not significantly affected by the variations in the pH and the salt concentration. Instead, it appears that these two parameters affected the unusual helices.
Fig. 3. Effects of pH and ionic strength on the patterns of arrest at the (GA)$_{27}$ tract. (A) Effects of pH. The Klenow polymerase was used at an enzyme/template ratio of 8 units/µg DNA to replicate single stranded DNA prepared from the phage recombinant mpGA27. The reactions and the analysis were carried out as described in Materials and Methods, except for that pH was varied as follows: lanes T and 1- pH=7.0, lane 2- pH=8.0, lane 3- pH=8.5. The assay shown in lane T was carried out in the presence of ddTTP. The assays shown in lanes 1-3 were carried out in the absence of any ddNTP. The two arrows designate the boundaries of the (TC)$_{27}$ tract in the newly synthesized DNA, and the arrowhead designates the center of the arrest sites found within the (TC)$_{27}$. (B) Effects of ionic strength. Replication assays were carried out as described in A., except that the pH was 8.0 in all assays, and NaCl was added to the reaction mixtures at the following concentrations: panel a- no NaCl; panel b- 25 mM NaCl; panel c- 50 mM NaCl; panel d- 75 mM NaCl. The lanes designated T present assays carried out in the presence of ddTTP and the lanes designated (-) present assays carried out in the absence of any ddNTP.
generated at the (TC)$_{27}$ and the (GA)$_{27}$ tracts, or the interactions between these helices and the polymerase (see Discussion). Further studies of the arrest reactions were carried out in a low ionic strength buffer at pH 8.0, unless specified otherwise.

So far, we have used for the replication assays the experimental protocol devised for sequencing reactions, in which replication was first performed with radioactively labeled dATP (and unlabeled dTTP, dCTP and dGTP), and was then extended in the presence of a large excess of unlabeled dATP. The bands observed in this type of assay must represent termination sites for replication, rather than temporary pause sites for the polymerase. To find out whether there are also temporary pause sites within the (TC)$_{27}$ and the (GA)$_{27}$ tracts, we have performed similar replication assays (with no ddNTPs), in which the reactions were stopped either at the end of a short labeling period (a pulse), or after additional chase periods. Fig. 4 shows the gel patterns obtained in one experiment of this type, in which the strand including the (GA)$_{27}$ tract was used as a template. It can be seen (lane 1) that after a pulse of 3 min. the number of bands found within the (TC)$_{27}$ tract in the newly synthesized strand exceeded the number observed in Fig. 1, and that many bands were also observed at sequences located outside the (TC)$_{27}$ tract. However, the number of bands decreased after a 15 min chase period (lane 2), and further decreased after a 45 min. chase (lane 3). The bands that remained were all located within the (TC)$_{27}$ tract and corresponded to those seen in Fig. 1. These bands must represent strong arrest sites, while the bands that were chased represent temporary pauses in the replication. Similar results were obtained when the other strand was used as a template (not shown). For the assays presented below, we used the experimental protocol described in Materials and Methods—a labeling period followed by a chase.

The next series of experiments were carried out to determine whether variations in the concentrations of the dNTP precursors and the polymerase might affect the positions and the relative strengths of the arrest sites. Fig. 5A presents a series of assays designed to examine the effects of variations in the concentrations of the dNTP precursors. The template used for these experiments was the strand including the (TC)$_{27}$ tract. Lane A shows the gel pattern obtained in a reaction that was carried out in the presence of ddATP. Lanes 1-6 show the results of assays in which replication was performed in the absence of any ddNTP. The concentrations of dGTP, dTTP and dCTP were varied over a range of 100-fold during the first step of these reactions, while the concentration of the (α$^{32}$P)-dATP was kept constant. This labeling period was followed by a chase. It can be seen that the locations of the arrest sites were not significantly affected by the changes in the concentrations of the unlabeled dNTPs. At the lower concentrations the stronger bands became less prominent and other bands disappeared, presumably because most polymerase molecules failed to reach the (TC)$_{27}$ tract during the initial phase of the reaction. Similar results were obtained in similar reactions performed with the template strand including the (GA)$_{27}$ tract. It should be noted that, for economical reasons, the concentration of the radioactively labeled dATP could not be increased during the first step of these reactions. However, using radioactively labeled primers, we have carried out replication assays, in which the concentrations of all four dNTPs were as high as those employed in the present assays during the chase period, and obtained similar patterns of arrest (results not shown).

Fig. 5B shows the dependence of the arrest patterns on the polymerase/template ratios. In these assays, the strand including the (GA)$_{27}$ tract was used as a template. The reactions were carried out either in the presence of ddTTP (T lanes), or in the absence of any ddNTP (-) lanes. At the highest enzyme/template ratio (panel a), the patterns of arrest resembled the patterns observed in Fig. 1. The number of residues at which arrests occurred, increased in a symmetrical fashion as the enzyme/template ratio was decreased (panels b-
Fig. 4. Pulse-chase assays. The Klenow polymerase was used at an enzyme/template ratio of 5 units/μg DNA to replicate single stranded DNA prepared from the phage mpGA27. Three identical reactions were initiated by incubating the enzyme with a template-primer complex at a concentration of 20 μg/ml, in the presence of 5 mM tris HCl pH=8.0, 2.5 mM MgCl₂, 1.2 μM dTTP, 1.2 μM dGTP, 1.2 μM dCTP and 1.0 μM (α⁻³²P)-dATP (3000 Ci/mM). The reactions were allowed to proceed for 3 min at 33°C, after which time one assay was terminated (lane 1). 150 μM of each of the four dNTPs were then added to the other two tubes and the reactions were further incubated at 33°C for 15 min (lane 2) and for 45 min (lane 3). The arrows designate the boundaries of the (TC)₂₇ tract in the newly synthesized strand and the arrowhead designates the center of the arrest sites.

d). At the lowest ratio used in these assays (panel d), arrests occurred throughout most of the (GA)₂₇ tract. The variation in the patterns of arrest observed in this experiment, resembles the variation found in the pulse-chase assays shown in Fig. 4. This similarity indicates that many, or all the polymerase molecules do not simply stall at the pause sites found in Fig. 4; instead, dissociation and reassociation of enzyme molecules accounts for these pauses.

Another interesting feature of the data shown in Fig. 5B is the periodicity of strong and weak arrest sites (panel d). Such periodicity was often observed in assays conducted at relatively low enzyme/template ratios, particularly when such assays were carried out in the presence of ddNTPs. Two more clear examples of periodical arrests are shown in Fig. 5C. Lane 1 presents a periodical arrest pattern obtained in a replication assay of the strand including the (GA)₂₇ tract. This assay was carried out in the presence of ddCTP. It can be seen that the period of the arrest signals found in this lane is 12. (note that the number of bands between arrowheads represents the number of TC repeats). Lane 2 presents the gel pattern obtained in a replication assay of the strand including the (TC)₂₇ tract. This assay was also carried out in the presence of ddCTP. The period of arrest signals in this gel is 10.

Arrest of single stranded DNA replication at cloned synthetic (GA)n and (TC)n tracts.

The templates used in the experiments described so far included the six bp inverted repeats and the six bp direct repeats found in the rat genome next to the (TC)₂₇-(GA)₂₇ tract (18). It appeared possible that these other elements might play a role in the replication arrests that were observed within the (TC)₂₇ and the (GA)₂₇ tracts. In order to examine this question, and to find out the minimal lengths of (GA)n and (TC)n tracts that are needed to arrest DNA replication, we have cloned synthetic (TC)n,(GA)n sequences of various lengths.
Fig. 5. Patterns of arrest obtained at various concentrations of dNTPs and polymerase. (A) Dependence on concentrations of dNTPs. The Klenow polymerase was used at an enzyme/template ratio of 7 units/µg DNA to replicate single stranded DNA prepared from the recombinant phage mpTC27. The reaction shown in lane A was carried out in the presence of ddATP. The other reactions were carried out in the absence of any ddNTP, as described in Materials and Methods, except that during the first stage of each reaction the concentrations of dGTP, dCTP and dTTP were all varied, as follows: lane 1- 60 µM; lane 2- 30 µM; lane 3- 6 µM; lane 4- 3 µM; lane 5- 1 µM; lane 6- 0.60 µM. (B) Dependence on polymerase concentration. The Klenow polymerase was used to replicate single stranded DNA prepared from the phage recombinant mpGA27, as described in Materials and Methods. The enzyme/template ratio was varied as follows: panel a- 10 units/µg DNA; panel b- 5 units/µg DNA; panel c- 3 units/µg DNA; panel d- 1 unit/µg DNA. The lanes T and (-) designate assays that were carried out in the presence of ddTTP, or in the absence of any ddNTP, respectively. (C) Periodical patterns of arrest. Lane 1- The Klenow polymerase was used at an enzyme/template ratio of 5 units/µg DNA to replicate single stranded DNA purified from the phage recombinant mpGA27. This reaction was carried out in the presence of ddCTP. The arrowheads denote the periodic arrest signals, and the arrows designate the boundaries of the (GA)27 sequence. Lane 2- The Klenow polymerase was used at an enzyme/template ratio of 1 unit/µg DNA to replicate single stranded DNA purified from the phage recombinant mpTC27. The reaction was carried out in the presence of ddCTP.

Fig. 6A (panel a) shows the gel patterns obtained in assays of a template including an (AG)9-A tract, carried out at an enzyme/template ratio of 1.0
Fig. 6. Arrest of DNA replication at cloned synthetic (GA)n and (TC)n tracts of various lengths. (A) The Klenow polymerase was used to replicate single stranded DNA molecules containing (GA)n tracts. Lanes G, T and (-) designate assays that were carried out in the presence of ddGTP, or ddTTP, or in the absence of any ddNTP, respectively. The arrows point at the boundaries of the (TC)n tracts in the newly synthesized strands. The arrowhead designates arrest sites. Panel a—replication of DNA containing (AG)g-A. The enzyme-template ratio was 1 unit/µg DNA. Panel b—replication of DNA containing (GA)_16. The enzyme-template ratio was 2 units/µg DNA. Panel c—the same template as in panel b. The enzyme-template ratio was 1 unit/µg DNA. (B) The Klenow polymerase was used to replicate single stranded DNA molecules containing (TC)n tracts. The notations are the same as in 6A. Panel a—replication of DNA containing (TC)g-T. The enzyme-template ratio was 1 unit/µg DNA. Panel b—replication of DNA containing (TC)^{17}_C-(TC)^2. The enzyme-template ratio was 2 units/µg DNA. Panel c—the same template as in b. The enzyme-template ratio was 1 unit/µg DNA. Panel d—replication of DNA containing (TC)^{23}_T. The enzyme-template ratio was 2 units/µg DNA. Panel e—replication of DNA containing (TC)^{31}. The enzyme-template ratio was 10 units/µg DNA. Panel f—the same template as in e. The enzyme-template ratio was 1 unit/µg DNA.

unit/µg DNA, in the presence of ddGTP (G lane), or ddTTP (T lane), or in the absence of any ddNTP ((-) lane). It can be seen that the polymerase was not arrested at the (AG)g-A tract. Other assays, carried out at higher enzyme/template ratios gave similar results. At lower ratios of enzyme/template, the polymerase was arrested non-specifically at many sites inside and outside the (AG)g-A tract (not shown). Panels b and c in Fig. 6A show similar assays of a template including a (GA)_16 tract, carried out at enzyme/template ratios of 2.0 units/µg and 1.0 unit/µg, respectively. It can be seen that the polymerase was arrested in the (GA)_16 tract at the lower enzyme/template ratio (panel c), but
Fig. 7. Arrest of the calf thymus DNA polymerase α at (GA)n and (TC)n tracts. The calf thymus DNA polymerase α was used to replicate single stranded DNA molecules containing (GA)n and (TC)n tracts. These reactions were carried out with a radioactively labeled primer, as described in Material and Methods. The DNA templates included the following tracts: (GA)_{16} (panel a); (TC)_{17}−C−(TC)_{2} (panel b); (TC)_{23}−T (panel c); (TC)_{31} (panel d). The lanes designated T, A, G show the patterns obtained in reactions performed with the Klenow polymerase, in the presence of ddTTP, ddATP and ddGTP respectively. The lanes designated α show the patterns obtained in the assays performed with pol α. The products of the reactions catalyzed by the bacterial and the mammalian enzymes were coelectrophoresed in the same gels.

was not arrested at the higher ratio (panel b). Thus, unlike the (GA)_{27} tract, the (GA)_{16} tract arrested DNA replication only within a small range of enzyme/template ratios. In both cases, however, the arrests occurred predominantly at G residues around the middle of the tracts. It should be noted that despite many attempts to clone longer synthetic (GA)n tracts, we have not obtained phage clones including tracts of n>16.

Panels a, b and c in Fig. 6B show the gel patterns obtained in similar
replication assays of templates including (TC)n tracts, whose lengths were comparable to those of the (GA)n tracts in Fig. 6A. The results were similar: The replication was not arrested at a (TC)9-T tract (panel a). Also, specific arrests were observed within a (TC)17-C-(TC)2 tract when the enzyme/template ratio was 1.0 units/µg DNA (panel c), but not when the ratio was 2.0 units/µg DNA (panel b). In the reaction that was carried out in the absence of a ddNTP (panel c, (-) lane), the arrests occurred preferentially at A residues and were located at the end, rather than the middle of the sequence. However, when ddTTP, or ddGTP were included in the reaction (panel c, lanes G and T), periodical arrests occurred at G residues with a period of 10, in agreement with the pattern found in lane 2 of Fig. 5C. It should be noted that these reactions were carried out at pH~8.0. In similar assays carried out at pH=7.0, arrests were observed around the middle of the (TC)17-C-(TC)2 tract even at enzyme/template ratios of 5.0 units/µg DNA; however, no arrests were detectable at the (TC)9-T tract (not shown).

Panel d in Fig. 6B shows the gel pattern obtained in a replication assay of a template including a (TC)23-T tract, carried out at an enzyme/template ratio of 2.0 units/µg DNA. It can be seen that the replication was specifically arrested around the middle of this tract, and that the arrests occurred predominantly at A residues. In other assays, specific arrests were observed at the same positions, at enzyme/template ratios of up to 10 units/µg DNA (not shown).

Panels e and f present the gel patterns obtained in replication assays of a template including a (TC)31 tract, that were carried out at two different enzyme/template ratios. It can be seen that at the higher ratio of 10 units/µg DNA (panel e), the arrests occurred preferentially at residues found around the middle of the (TC)31 tract. At the lower enzyme/template ratio of 1.0 unit/µg DNA (panel f), the arrests were periodical, with a period of 10. The arrests occurred at both G and A residues, but were slightly stronger at the G residues.

Based on these data, we conclude that the Klenow polymerase is specifically arrested at (GA)n and (TC)n segments of n>16. Unlike the (TC)27 and the (GA)27 tracts, the synthetic cloned (TC)n and (GA)n tracts are not adjacent to other repeat elements. Therefore, the 6 bp repeat elements, which were found next to the (TC)27-(GA)27 tract in the genomic DNA, do not appear to play a significant role in arresting the replication.

(GA)n and (TC)n tracts arrest single stranded DNA replication catalysed by the mammalian DNA polymerase α.

Since DNA polymerase α is presumably the primary replication enzyme in mammalian cells (30), it was interesting to find out whether single stranded DNA replication catalysed by this enzyme is arrested at (GA)n and (TC)n tracts. We used for these assays some of the templates described in the previous section, and an immunoaffinity purified calf thymus DNA polymerase α (28). In these reactions, a radioactively labeled primer was extended at relatively high concentrations of the four (unlabeled) dNTPs, in the absence of ddNTPs. The gel patterns obtained in these assays are shown in Fig. 7 (a lanes), next to lanes displaying sequencing reactions of the same templates carried out with the Klenow polymerase, and designated by the ddNTPs used for the reactions. Panel a shows the results obtained in assays of the template including a (GA)16 tract. It can be seen that pol α was arrested at several residues around the middle of the (GA)16 tract and at residues preceding and following this tract (the Klenow polymerase was not arrested in this assay). It is interesting to note that the mammalian enzyme was arrested more efficiently at A residues than at G residues, unlike the bacterial enzyme which was preferentially arrested at G residues of (GA)n sequences (see Fig. 1). It should also be noted that, like the Klenow polymerase, pol α was not specifically arrested at an (AG)9-A tract (data not shown).

Panel b in Fig. 7 shows the gel patterns obtained in replication assays of
the template including a (TC)$_{17}$-C-(TC)$_{2}$ tract. It can be seen that the mammalian enzyme was not arrested at the (TC)$_{17}$-C-(TC)$_{2}$ tract, but was arrested at sequences that precede and follow this tract. Panel c shows similar assays of the template including a (TC)$_{23}$-T tract. It can be seen that both the mammalian and the bacterial polymerases were arrested around the middle of this tract. However, many pol α molecules also proceeded beyond this tract and were arrested at more distal sequences. Panel d shows replication assays of the template including a (TC)$_{31}$ tract. It can be seen that both pol α and the Klenow polymerase were arrested around the middle of this tract. A comparison of panels c and d reveals that the fraction of enzyme molecules which proceeded beyond the (TC)$_{31}$ tract was smaller than the fraction of molecules which crossed the (TC)$_{23}$-T tract. It appears, therefore, that the mammalian polymerase was more efficiently arrested at the longer tract. Finally, it should be noted that, unlike the Klenow polymerase, in the assays shown in the panels c and d, pol α was preferentially arrested at G residues.

**DISCUSSION.**

The single stranded DNA replication assays reported in this article have indicated that (TC)$_{n}$ and (GA)$_{n}$ tracts of $n>16$ are specific stop signals for the Klenow fragment of the E. coli DNA polymerase I. The enzyme was found to stop at multiple sites within these tracts, some of which were temporary pause sites— at these sites the replication continued after a pause. At other sites, replication of many growing chains was permanently arrested. The distribution of the arrest sites was found to depend on the polymerase/DNA ratios. At relatively low ratios of enzyme/template, this distribution was rather broad, and often periodical. At higher ratios, a relatively small number of strong termination sites were observed near the middle of both types of tracts.

We interpret these results as follows: At each step of DNA chain elongation, a polymerase molecule may either dissociate from the DNA, or catalyse addition of one nucleotide to the growing chain, and then move towards the next residue in the template. In fact, dissociation and reassociation of the Klenow polymerase occurs quite often during replication, as indicated by the rather low processivity of the enzyme (31). According to this view, at (TC)$_{n}$ and (GA)$_{n}$ tracts, the dissociation constant of the enzyme is much higher than elsewhere along the DNA sequence and this accounts for the pauses and arrests observed in our assays.

As mentioned in the Introduction, it is now believed that in supercoiled double stranded plasmids, (TC)$_{n}$-(GA)$_{n}$ tracts can be transformed into triple helix+single strand structures (12-17). Each of these structures is generated as follows: Half of the (TC)$_{n}$ tract remains bound to the complementary half of the (GA)$_{n}$ tract by Watson Crick hydrogen bonds. The other half of the (TC)$_{n}$ tract is folded back, occupies the major groove of the Watson Crick double helix and is bound to the (GA)$_{n}$ strand of the helix by Hoogstein hydrogen bonds. The other half of the (GA)$_{n}$ tract remains single stranded. The transition from double stranded DNA into the triple helix+single strand structure occurs most readily at pH<6.0, because formation of the Hoogstein bonds requires hemiprotonation of the cytosines (12-17). However, it also occurs at pH>7.0 in plasmids with a high degree of supercoiling (15). As shown in Fig. 8, the arrests observed in our assays of the templates including (TC)$_{n}$ tracts could be accounted for by formation of similar triple helices consisting of the (TC)$_{n}$ tracts folded around the middle and the newly synthesized half (GA)$_{n}$ tracts. Moreover, unlike double stranded plasmids, no hydrogen bonds should be disrupted to allow formation of these triple helices. Therefore, such helices are expected to form more readily at physiological pH values, and cause arrest of replication. This explanation is compatible with the observation that the strongest arrest sites occur near the middle of the (TC)$_{n}$ tracts, because the most stable triple helices are expected
to be generated at this stage of the replication. The pH dependence of the
arrest reactions also supports this notion— in the range of 6.5<pH<8.0, stronger
arrests occurred at the lower pH values, and no arrests were detectable at
pH=8.5. Furthermore, we have recently carried out replication assays with 7
deaza dGTP, which can form Watson Crick hydrogen bonds with cytosine in DNA, but
cannot form Hoogsteen hydrogen bonds; in these reactions, the polymerase was not
arrested at the (TC)n tracts, as expected if the arrests were due to formation
of triple helices (Baran, Lapidot and Manor, unpublished results).

It is interesting that the arrest sites found in our replication assays of
templates including (GA)n tracts, also mapped near the middle of these tracts.
This result suggests the possibility that formation of (GA)n.(TC)n.(GA)n triple
helices might account for the arrests observed in these reactions. However, even
though formation of (G)n.(C)n.(G)n triple helices, which do not require
protonation of the cytidine residues, has been reported (32), analogous
(GA)n.(TC)n.(GA)n structures have not been observed so far. Clearly, further
experiments would be required to examine this possibility.

The pause sites observed only at relatively low polymerase/template ratios
might be accounted for by formation of shorter triple helices consisting of less
than half of the (TC)n and (GA)n tracts. The periodical distribution of pause
sites (with a period of 10) observed in replication of the (TC)n tracts, might
reflect periodicities of the triple helices (TC)n.(GA)n.(TC)n generated at
various stages of the replication. It is interesting that the period of pause
sites found during replication of the (GA)n tracts was 12 instead of 10. This
difference in the periodicities, and the varying locations of the arrest sites
found in replication of the two strands (see fig. 1), could be due to
differences in the structures of (GA)n.(TC)n.(GA)n and (TC)n.(GA)n.(TC)n triple
helices.

The patterns of arrest found within (GA)n and (TC)n tracts in the assays
performed with the mammalian pol α, resembled the patterns obtained in the
corresponding reactions catalysed by the Klenow polymerase (Fig. 7). Both the
locations and the symmetrical distribution of the arrest sites were similar.
However, unlike the Klenow polymerase, pol α was preferentially arrested at A
residues when (GA)n tracts were in the templates, and at G residues when (TC)n
tracts were in the templates. These differences indicate that the interactions
of these two polymerases with the (GA)n and (TC)n sequences are not entirely
equivalent. On the other hand, similar assays conducted with the chemically
modified T7 DNA polymerase designated sequenase (33; obtained from USB
corporation), which shares structural features with the Klenow polymerase (34),
gave patterns of arrest at (GA)n and (TC)n tracts that were the same as those
found in assays of the Klenow polymerase (Baran, N., Lapidot, A. and Manor, H.,
Unpublished Results).

As mentioned in the Introduction, (TC)n.(GA)n tracts are rather abundant
and highly dispersed in vertebrate genomes, particularly in rodents. Their wide
distribution does not necessarily imply that these elements perform a useful
function in vivo. However, our recent discovery of the (TC)₂₇.(GA)₂₇ tract at
the end of an amplified region of DNA in Py transformed rat cells, has indicated
that such tracts might serve as pause or arrest signals for DNA amplification, or for normal replicon replication (18,20). The observation that a segment including this sequence causes a pause in SV40 DNA replication (21), and the results presented in this article, lend further support to this hypothesis.

The existence of termination signals for replication in mammalian cells is compatible with a widely discussed model for gene amplification. According to this model, multiple rounds of replication are initiated at some replicons during a single cell cycle. It is assumed that several replication forks accumulate in these replicons, and that homologous and illegitimate recombination events occur in the vicinity of the forks at an enhanced rate. These recombination events generate multiple sequence repeats (35-37). It is evident that fork accumulation, and hence the amplification process, may be dependent on the presence of sites which slow down, or arrest the movement of the forks. The wide dispersion of (GA)n.(TC)n tracts in vertebrate genomes suggests that these elements might serve as pause or arrest signals for amplification at many chromosomal regions. It is also possible that amplification of large segments of DNA could be lethal to the cells, and that, therefore, the (GA)n.(TC)n elements might prevent cell death by limiting the size of amplified regions.

It is interesting to note that homopurine-homopyrimidine sequences, which are not strictly alternating, are also prevalent in mammalian cells (7). Such sequences have been observed in rat genomes at the ends of the highly dispersed long repeat elements, designated LINES, and have been also shown to arrest single stranded DNA replication in vitro (38). It has been suggested that these homopurine-homopyrimidine sequences might play a role in transposition of the LINE elements by virtue of their ability to arrest DNA replication (38).

It could also be advantageous to cells to have termination signals for bidirectional replication at the boundaries of normal replicons. For example, if segregation of newly replicated DNA occurred only at specific sites, movement of replication forks beyond such sites would be deleterious to the cells. The presence of pause sites, or termination signals, at or next to segregation sites, will ensure that the forks will not move beyond these sites. It is interesting to note, in this connection, that pause signals for replication have been identified in SV40 DNA next to its segregation site (39). In bacterial chromosomes, DNA replication arrest sites were found to be located on both sides of the segregation point. Furthermore, these sites are asymmetrical, that is, each of them blocks the movement of forks that cross the segregation point, but not the movement of forks that have not reached this point (40,41). Mammalian pause sites, or arrest signals, might also act asymmetrical on forks which originate at two adjacent replicons and advance in opposite directions. In view of our data on the differential ability of (GA)n and (TC)n tracts to arrest DNA replication by pol a (see Fig. 7), and the asymmetrical mode of DNA replication, double stranded (TC)n. (GA)n tracts might be such asymmetrical termination signals; for in one direction the (GA)n sequence would be the template for the retrograde chain and the (TC)n sequence would be the template for the leading strand, while in the other direction these roles will be reversed. It is conceivable, for example, that the replication would only be arrested at a given (TC)n. (GA)n sequence when the (GA)n tract, and not the (TC)n tract, would be in the templates for the retrograde chain. Therefore, (TC)n.(GA)n sequence may only cause arrest of DNA replication in one of the two orientations.

ACKNOWLEDGEMENTS
We thank Drs. L. Loeb and F. Perrino for the generous gift of calf thymus DNA polymerase a. We also thank Drs. R.G. Martin, A.V. Furano and T. Miles for helpful discussions and Yaffa Bot for technical assistance. This work was
supported by a grant from the National Council for Research and Development, Israel and the Deutsches Krebsforschungszentrum, Heidelberg, Federal Republic of Germany, and by grants from the United States - Israel Binational Science Foundation and the Israel Cancer Research Fund.

REFERENCES