ISSN 0026-8933, Molecular Biology, 2010, Vol. 44, No. 3, pp. 458–465. © Pleiades Publishing, Inc., 2010.
Original Russian Text © V.K. Gasanova, N.V. Ryadninskaya, C. Gaillard, F. Strauss, G.A. Belitsky, M.G. Yakubovskaya, 2010, published in Molekulyarnaya Biologiya, 2010, Vol. 44, No. 3, pp. 520–528.

STRUCTURAL-FUNCTIONAL ANALYSIS OF BIOPOLYMERS AND THEIR COMPLEXES

UDC 577.123

Invasion of Complementary Oligonucleotides into $(CA/TG)_{31}$ Repetitive Region of Linear and Circular DNA Duplexes

V. K. Gasanova[a], N. V. Ryadninskaya[a], C. Gaillard[b], F. Strauss[b], G. A. Belitsky[a], and M. G. Yakubovskaya[a]

[a] *Carcinogenesis Institute, Blokhin Cancer Research Center, Russian Academy of Medical Sciences, Moscow, 115478 Russia; e-mail: vkgasanova@mail.ru*

[b] *Universite Pierre et Marie Curie, Centre National de la Recherche Scientifique, Paris, 75006 France*

Received July 23, 2009
Accepted for publication October 12, 2009

Abstract—$(CA/TG)_n$ repeats belong to microsatellite DNA. They are the most abundant among the other dinucleotide repeats in mammals, constituting approximately 0.25% of the entire genome. These repeats are recombination hot spots; however, the corresponding mechanisms are yet vague. We postulated that one of the reasons underlying an increase in the recombination frequency in the repetitive region could be the conformational characteristics of duplex resulting from a specific geometry of base-stacking contacts, providing for initiation of a single-stranded DNA invasion in the duplex homologous regions. This work for the first time demonstrates a DNA–DNA interaction of the $d(CA)_{10}$ and $d(TG)_{10}$ oligonucleotides with linear and circular duplexes containing $(CA/TG)_{31}$ repeats during their coincubation in a protein-free water solution at 37°C. Using radioactively labeled oligonucleotides, we demonstrated that the duplex–oligonucleotide interaction intensity depended on the molar ratio of duplex-to-oligonucleotide at a duplex concentration of 30 nM. A decrease in this concentration to 3 nM had no effect on the intensity of oligonucleotide invasion. It was demonstrated that over 1% of the duplexes yet much less than 10% were involved in the interaction with oligonucleotides assuming that one oligonucleotide molecule interacted with one molecule of the duplex. Analysis of the kinetics showed that $d(CA)_{10}$ invasion commenced from the first minute of incubation with duplexes, while $d(TG)_{10}$ interacted with the duplex even at a higher rate. The role of conformational plasticity of CA/TG repeats in the discovered interaction is discussed as well as its biological significance, in particular, the role of CA microsatellites in the initiation of homologous recombination.

DOI: 10.1134/S0026893310030155

Key words: DNA, noncanonical structure, (CA/TG) repeats, duplexes, oligonucleotide invasion, CA microsatellites

Double-stranded DNA is able to form structures differing from the classical B-form, which provides for its involvement in chromatin organization as well as regulation of replication, transcription, and recombination. In particular, the sequences forming noncanonical structures include microsatellite DNA repeats with repetitive units of one to six nucleotides. The most abundant dinucleotide repeats, in particular, of mammals and human, are $(CA/TG)_n$ repeats, which constitute approximately 0.25% of the genome [1]. There are data demonstrating the effect of $(CA/TG)_n$ repeats on the frequencies of homologous recombination in yeast chromosomes and intraplasmid recombination in bacteria as well as the effect of CA repeats on individual stages of RecA-dependent homologous recombination in vitro [2–5]. $(CA/TG)_n$ repeats are also hot spots for human recombination [6]. There are the data suggesting a possible role of purine–pyrimidine sequence in the initiation of chromosome translocations in various types of human lymphoid tumors [7–9]. In addition, CA microsatellites can function as modulators of gene expression via regulation of both the promoter activity [10, 11] and alternative splicing [12–15].

The molecular mechanisms underlying the effect of repeats on recombination has not been yet studied. A possible explanation of the role of microsatellite repeats as recombination hot spots is their ability to form noncanonical DNA structures due to intra- and intermolecular interactions, which influence initiation and(or) termination of homologous recombination [3, 16–18]. This ability is explainable by the fact that the free energy values for the B-DNA and alternative conformation of double-stranded DNA in the repetitive region are close. Correspondingly, study of the characteristics of DNA duplex in the region of CA/TG repeats is a topical problem, in particular, for a better insight into the mechanisms of genetic rearrangements.

Abbreviations: PCR, polymerase chain reaction; NMR, nuclear magnetic resonance.

Earlier, we discovered the phenomenon of a spontaneous interaction between a 20-meric single-stranded oligonucleotide and the complementary terminal region of a strand of a linear duplex formed of a random sequence [19]. In this process, the second strand of the duplex, homologous to the invaded oligonucleotide is successively substituted. It has been also demonstrated that the duplex does not interact with the random sequence oligonucleotide homologous to the inner regions of the double-stranded fragment. The discovered difference between the invasion frequencies of oligonucleotides into the terminal and inner regions of the duplex complies with the modern concept on the structure of DNA duplex, stating that the terminal base pairs melt considerably more frequently as compared with the inner pairs due to weakened base-stacking at the ends of the molecule [20]. Nonetheless, such weakening can also take place in the inner regions of the duplex that contain DNA sequences capable of forming noncanonical structures. The goal of this work was to try to detect possible interactions between the duplex containing CA/TG repetitive sequence and the oligonucleotides homologous to its different regions.

EXPERIMENTAL

Used in this work was the plasmid pE10 (X96980 in the EMBL database), which is a derivative of the plasmid pUC19 [21]. The plasmid was isolated from *E. coli* strain XL1 blue [*recA1 endA1 gyrA96 thi*-1 *hsdr17 supE44 relA1 lac* (F9 *proAB lacI* q *ZM15 Tn10*)] (Stratagene) by alkaline lysis and conventional purification from proteins and RNA [22]. A linearized plasmid was also used; it was produced by digesting pE10 with the restriction endonuclease *Pvu*II (SibEnzim, Russia), hydrolyzing phosphodiester bond at a distance of 217 bp from the beginning of $(CA/TG)_{31}$ repeats, followed by protein removal using chloroform–isoamyl alcohol mixture. DNA from aqueous phase was precipitated with 96% ethanol and dissolved in TE buffer (10 mM Tris-HCl and 1 mM EDTA, pH 7.5).

PCR amplification of the fragment with a length of 1129 bp containing $(CA/TG)_{31}$ repeats was performed with the primers containing heptameric AT-rich sequences noncomplementary to the template at their 5'-ends (Sintol, Russia): 5'atattttcccgactggaaagcggg3' (forward, 5'-CA-1129) and 5'taattaaggaagctagagtaagtag3' (reverse, 3'-CA-1129). *Taq* polymerase (Biomaster, Russia) and pE10 plasmid as a template were used in the reaction. PCR was conducted according to the following mode: 1 min at 95°C; 30 cycles of 40 s at 95°C, 1 min at 56°C, and 1 min at 72°C; 10 min at 72°C, and cooling to 4°C.

To study the interaction of oligonucleotides with duplexes, the following synthetic oligonucleotides (Sintol, Russia) were used: $d(CA)_{12}$; $d(CA)_{10}$; $d(TG)_{10}$; 5'-cccggggatcctctagagtc-3' (20-meric 5-R_{1s}-20), complementary to a random sequence of TG containing pE10 strand; 5'-ctggcacgacaggtttcccga-3' (21-meric E10-*Pvu*II), complementary to the 3'-end of TG-containing strand on the linearized pE10; 5'-cgagctcgccgggggatcct-3' (20-meric 5'-*Eco*RI), complementary to a random sequence of the TG-containing plasmid strand; and 5'-aaagaatggtcctgcacc-3' (18-meric 3'-18-*Ki-ras*), displaying no homology to pE10 plasmid. Figure 1a schematizes the localization of the listed oligonucleotides relative to the $(CA/TG)_{31}$ repeats in the plasmid.

DNA was separated by electrophoresis at a room temperature in 1% agarose gel at a voltage of 60 V in 0.5× Tris–borate buffer pH 7.5 (0.089 M Tris–borate and 0.002 M EDTA). Gel was stained with ethidium bromide (Sigma, United States). The dye SYBR Gold (Molecular Probes, United States) was used for staining preparative gels.

PCR products were purified by electrophoresis in 1% agarose gel; the target fragment was extracted from the gel with the help of DNA-binding matrix of fine silicon dioxide (Sigma, United States) according to the earlier described protocol [23] followed by DNA precipitation with ethanol to remove SYBR Gold and dissolution in TE buffer.

Radioactive labeling of oligonucleotides. The oligonucleotides dephosphorylated at 5'-ends were labeled using γ-[^{32}P]ATP (10 mCi/ml) and T4 polynucleotide kinase (SibEnzim, Russia). The mixture was incubated at 37°C for 30 min. T4 polynucleotide kinase was inactivated by heating at 65°C for 10 min. Labeled oligonucleotides were purified on a column with Sephadex G-25 (Sigma, United States).

DNA incubation. The purified duplexes with a final concentration of 30 nM (except for the experiment on the effect of various concentrations of duplexes on the oligonucleotide invasion) were mixed at different molar ratios with synthetic radioactively labeled oligonucleotides. The mixture was incubated in the buffer containing 10 mM Tris-HCl, 1 mM EDTA, and 300 mM NaCl (pH 7.5). The total volume of the mixture was 5 μl (5 to 50 μl in the experiment with various concentrations of duplexes). Since DNA freezing could cause structural changes and, correspondingly, distortion of the observed phenomenon of oligonucleotide invasion, the DNA solutions in kinetic experiments were initially incubated for the longest period (24 h) with its subsequent reduction. The interaction of duplexes with oligonucleotides was stopped by placing the DNA sample into ice with subsequent loading on gel. The mixture of duplexes with the corresponding oligonucleotide prepared immediately before loading on gel was used as a zero point. DNA electrophoresis was performed in TBE buffer previously cooled to 4°C. In the subsequent kinetic experiment, the incubated samples were loaded on gel without cooling in ice.

Autoradiography. After electrophoresis, DNA was blotted from agarose gel onto a Hybond N+ membrane according to a standard protocol [22] with sub-

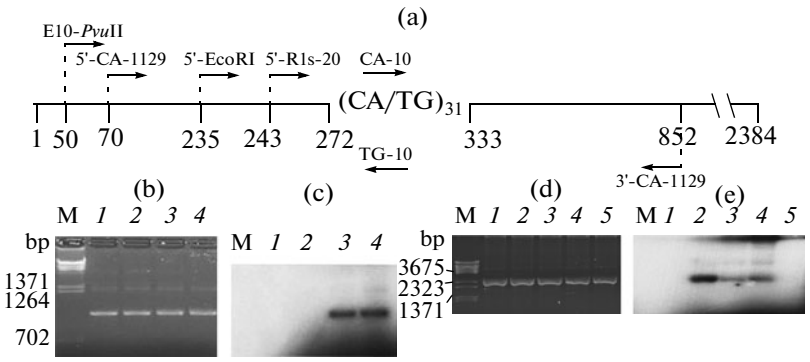

Fig. 1. (a) Schematic localization of the oligonucleotides used in experiments relative to the $(CA/TG)_{31}$ repeats in pE10 plasmid (arrows denote oligonucleotides in 5'⟶3' direction; base pairs in the plasmid are numbered). (b) Agarose gel electrophoresis of the mixtures of PCR products with the following oligonucleotides: (*1*) purified PCR product incubated without oligonucleotides; (*2–4*) purified PCR product incubated at a ratio of 1 : 50 with (*2*) nonhomologous oligonucleotide, (*3*) oligonucleotide 5'-CA-1129, and (*4*) oligonucleotide $d(CA)_{12}$. (c) Autoradiogram of gel in (b). (d) Agarose gel electrophoresis of the mixtures of linearized plasmid with the following oligonucleotides: (*1*) linearized plasmid incubated without oligonucleotides; (*2–5*) linearized plasmid incubated at a ratio of 1 : 50 with (*2*) oligonucleotide E10-*Pvu*II, (*3*) oligonucleotide $d(CA)_{10}$, (*4*) oligonucleotide $d(TG)_{10}$, and (*5*) oligonucleotide 5-R_{1s}-20. (e) Autoradiogram of gel in (d). M, molecular weight marker (λ/*Bst*EII).

sequent drying and exposure with a Kodak X-OMAT (Sigma, United States) film at –70°C. The gels with linearized and circular plasmids were dried in a vacuum drier followed by exposure with the film.

RESULTS

Demonstrating Interaction of the Linear Duplexes Containing $(CA/TG)_{31}$ Repeats with the Oligonucleotides Homologous to Repetitive Region

To analyze the interactions of oligonucleotides with repetitive sequence of linear duplexes, we used purified PCR product with a length of 1129 bp containing $(CA/TG)_{31}$ repeats in the region 208–269 bp from the 5'-end of the fragment. The purified PCR product was mixed at a molar ratio of 1 : 50 with one of the following radioactively labeled oligonucleotides: (1) 26-meric oligonucleotide 5'-CA-1129, complementary to the 3'-end of the TG-containing duplex strand (forward primer); (2) $d(CA)_{12}$, complementary to the TG sequences of the duplex; and (3) oligonucleotide 3'-18-*Ki-ras*, nonhomologous to the fragment. The samples were incubated at 37°C for 24 h. After the incubation, the samples were loaded on 1% agarose gel. Autoradiography of the gel demonstrated invasion of the oligonucleotides complementary to the 3'-end of the duplex strand and its inner repetitive sequence (Fig. 1b, c).

Invasion of the oligonucleotides with repetitive sequence into the linearized plasmid pE10 was also analyzed. For this experiment, the plasmid pE10 (length, 2384 bp) isolated from cells and purified was linearized at the site for *Pvu*II restriction endonuclease and mixed at a molar ratio of 1 : 50 with one of the following radioactively labeled oligonucleotides: (1) oligonucleotide E10-*Pvu*, complementary to the 3'-end of the TG strand in the duplex; (2) oligonucleotide $d(CA)_{10;}$ (3) oligonucleotide $d(TG)_{10}$; and (4) oligonucleotide 5-R_{1s}-20, complementary to a pE10 random sequence with its 3'-end located at a distance of 11 bp from the beginning of the repetitive regions. The gel autoradiogram demonstrates invasion of both the oligonucleotide complementary to the 3'-end of duplex strands and the oligonucleotides $d(CA)_{10}$ and $d(TG)_{10}$, whereas invasion of the oligonucleotide complementary to the inner nonrepetitive sequence is unobservable (Fig. 1d, e).

Dependence of the Interaction Intensity of Oligonucleotides $d(CA)_{10}$ and $d(TG)_{10}$ with Linear Duplexes on the Duplex-to-Oligonucleotide Molar Ratio

It was of interest to study the interaction intensity of oligonucleotides with the duplex depending on the amount of oligonucleotides. For this purpose, the linearized plasmid was mixed with $d(CA)_{10}$ at molar ratios of 1 : 10, 1 : 5, and 1 : 1, respectively, and with $d(TG)_{10}$ at molar ratios of 1 : 50, 1 : 25, 1 : 10, 1 : 5, and 1 : 1, respectively. After incubation, all samples were loaded on agarose gel to compare the intensities of interaction between oligonucleotides and the duplex (Fig. 2a). The gel autoradiogram shows that the invasion intensity of $d(TG)_{10}$ increased only insignificantly in the case of its 50-fold molar excess relative to the duplex as compared with the samples containing a lesser amount of this oligonucleotide (Fig. 2b, wells *4–8*).

The effect of the duplex-to-oligonucleotide ratio on the interaction intensity was also analyzed using the model PCR product of the plasmid pE10 with a length of 1129 bp. It was demonstrated that $d(TG)_{10}$ invaded the PCR product even at its 10-fold molar deficiency relative to the duplex (Fig. 2c, d; wells *2* and *4*). Unlike

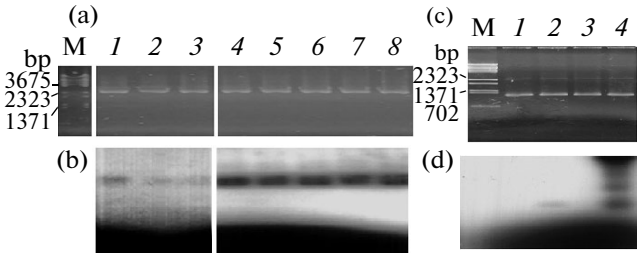

Fig. 2. Dependence of the interaction intensity of $d(CA)_{10}$ and $d(TG)_{10}$ oligonucleotides with linearized plasmid on the duplex-to-oligonucleotide molar ratio. (a) Agarose gel electrophoresis of the mixtures of duplexes with oligonucleotides (*1–3*) $d(CA)_{10}$ and (*4–8*) $d(TG)_{10}$ at a duplex-to-oligonucleotide ratios of (*1* and *6*) 1 : 10, (*2* and *7*) 1 : 5, (*3* and *8*) 1 : 1, (*4*) 1 : 50, and (*5*) 1 : 25. (b) Autoradiogram of gel in (a). (c) Agarose gel electrophoresis of the mixtures of duplexes with oligonucleotides (*1* and *3*) $d(CA)_{10}$ and (*2* and *4*) $d(TG)_{10}$ at a duplex-to-oligonucleotide ratios of (*1* and *2*) 10 : 1 and (*3* and *4*) 5 : 1.(d) Autoradiogram of gel in (c). M, molecular weight marker (λ/*Bst*EII).

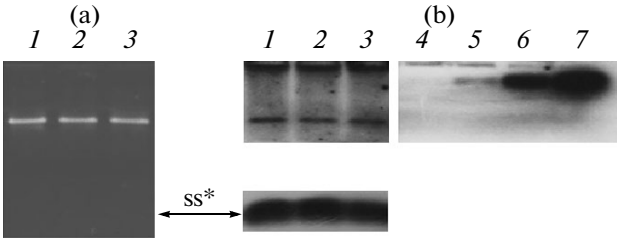

Fig. 3. (a) Agarose gel electrophoresis of the linearized pE10 plasmid with $d(CA)_{10}$ at a ratio of 1 : 50 at the following concentrations of the duplex: (*1*) 30 nM, (*2*) 10 nM, and (*3*) 3 nM. (b) Autoradiogram of gel in (a). Dilutions of free $d(CA)_{10}$ oligonucleotide relative to the amount of duplex in sample: (*4*) 1000-fold, (*5*) 100-fold, (*6*) tenfold, and (*7*) twofold; ss* is radioactively labeled (single-stranded) oligonucleotides not bound to duplex.

$d(TG)_{10}$, the interaction between linear duplex and $d(CA)_{10}$ under molar excess was less intensive in the case of linearized plasmid (Fig. 2b, wells *1–3*); moreover, no invasion was observed when incubating the PCR product at a molar deficiency of $d(CA)_{10}$, that is at duplex-to-oligonucleotide ratios of 5 : 1 and 10 : 1 (Fig. 2d, wells *1* and *3*).

The Effect of Concentration of Linear Duplexes on the Intensity of Interaction with Oligonucleotides

To clarify the effect of duplex concentration on the invasion intensity of homologous oligonucleotides, the linearized plasmid pE10 was mixed as a molar ratio of 1 : 50 with radioactively labeled oligonucleotide $d(CA)_{10}$; the concentration of duplex in the mixture was 30, 10, or 3 nM. As is evident from the gel autoradiogram, no changes in the intensity of oligonucleotide invasion into the duplex was observed at its different concentrations (Fig. 3a, b; wells *1–3*). To determine a percent ratio of the duplexes involved in the interaction with oligonucleotides, several dilutions of free radioactively labeled $d(CA)_{10}$ in an amount of 0.1, 1, 10, and 50% of the total duplex quantity in the sample were also loaded on the gel with the above indicated samples (Fig. 3b, wells *4–7*). Assuming that one molecule of oligonucleotide interacts with one molecule of duplex, the radioactivity in the duplex band in autoradiogram suggests that over 1% yet much less than 10% of the duplexes are involved in the interaction. When the samples containing a mixture of the duplexes with $d(CA)_{10}$ and duplexes with $d(TG)_{10}$ are loaded on the same gel, the invasion intensities of these oligonucleotides are approximately equal and even somewhat higher for $d(TG)_{10}$ (see Fig. 1); this suggests that the fraction of the duplexes interacting with $d(TG)_{10}$ falls into the same range.

Kinetics of Oligonucleotide Invasion in Linear Duplexes

Then we demonstrated the dependence of the invasion intensity of $d(CA)_{10}$ oligonucleotides in linear duplexes on the time of mixture incubation. For this purpose, linearized plasmid pE10 was mixed at a molar ratio of 1 : 50 with the radioactively labeled oligonucleotide and incubated at 37°C for various time periods. The incubation was stopped by placing DNA samples into ice. It was found that $d(CA)_{10}$ commenced to interact from the very first minute of mixture incubation with the duplex reaching the maximum at 6 h of incubation (Fig. 4a, b). Analysis of the interaction between the duplex and $d(TG)_{10}$ demonstrated that its invasion was so rapid that its maximal amount appeared in the duplex during the first seconds of incubation. One of the reasons can be the formation of stable quadruplexes with TG repeats by analogy with the process demonstrated by Kalyuzhny et al. [24] in the case of single-stranded oligonucleotides $d(TG)_n$ at 1–3°C. Correspondingly, we studied the invasion kinetics for $d(CA)_{10}$ and $d(TG)_{10}$ without cooling DNA samples in ice after incubation and performing electrophoresis at room temperature. Under such experimental conditions, $d(CA)_{10}$ displayed a more rapid interaction with the duplex, reaching the intensity maximum after 1-min incubation (Fig. 4c, d). On the other hand, $d(TG)_{10}$ still invaded the duplex at a very high rate, preventing from tracing the invasion kinetics at the selected time intervals of incubation (Fig. 5).

Invasion of Oligonucleotides $d(CA)_{10}$ and $d(TG)_{10}$ in Circular DNA Molecules

It was of interest to study the interaction of oligonucleotides not only with linear, but also with circular DNA molecules. For this purpose, we mixed pE10 plasmid at a molar ratio of 1 : 5 or 1 : 10 with each of the following radioactively labeled oligonucleotides: (1) oligonucleotide E10-*Pvu*II, complementary to the

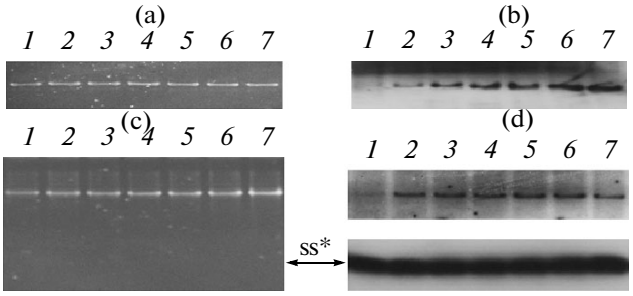

Fig. 4. (a) Agarose gel electrophoresis of the linearized pE10 plasmid incubated with $d(CA)_{10}$ at 37°C for the following time (with placing the samples in ice before loading): (*1*) 0 min, (*2*) 1 min, (*3*) 15 min, (*4*) 30 min, (*5*) 1 h, (*6*) 6 h, and (*7*) 24 h. (b) Autoradiogram of gel in (a). (c) Agarose gel electrophoresis of the linearized pE10 plasmid incubated with $d(CA)_{10}$ at 37°C for the following time (without cooling the samples in ice before loading): (*1*) 0 min, (*2*) 1 min, (*3*) 15 min, (*4*) 30 min, (*5*) 1 h, (*6*) 6 h, and (*7*) 24 h. (b) Autoradiogram of gel in (c); ss* is radioactively labeled (single-stranded) oligonucleotides not bound to duplex.

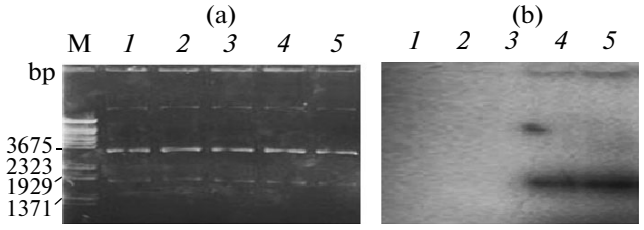

Fig. 6. Oligonucleotide invasion in the circular plasmid containing $(CA/TG)_{31}$ repeats. (a) Agarose gel electrophoresis of the mixtures of pE10 plasmid with the following radioactively labeled oligonucleotides: (*1*) nonhomologous oligonucleotide, (*2*) oligonucleotide 5'-*Pvu*II, (*3*) oligonucleotide 5'-*Eco*RI. (*4*) oligonucleotide $d(CA)_{10}$, and (*5*) oligonucleotide $d(TG)_{10}$. (b) Autoradiogram of gel in (a). M is molecular weight marker (λ/*Bst*EII).

inner nonrepetitive sequence of the plasmid, whose 3'-end is at a distance of 197 bp from the beginning of repeat; (2) oligonucleotide 5'-*Eco*RI, complementary to the inner nonrepetitive sequence of the plasmid, whose 3'-end is at a distance of 41 bp from the beginning or repeat (see scheme in Fig. 1); (3) oligonucleotide $d(CA)_{10}$; (4) oligonucleotide $d(TG)_{10}$); and (5) a nonhomologous oligonucleotide. The samples were incubated at 37°C for 24 h to load on 1% agarose gel. The plasmid in all these samples was present as several species, namely, circular, supercoiled, linear, and circular relaxed DNA (Fig. 6a). Gel autoradiogram demonstrates a predominant invasion of both $d(CA)_{10}$ and $d(TG)_{10}$ oligonucleotides into the circular supercoiled DNA species (Fig. 6b). No invasion of the oligonucleotides homologous to nonrepetitive plasmid sequences was observed.

Thus, this work for the first time demonstrates the interaction of single-stranded oligonucleotides $d(CA)_{10}$ and $d(TG)_{10}$ with the duplexes containing $(CA/TG)_{31}$ repeats in both circular DNA molecules and within linear DNA fragments.

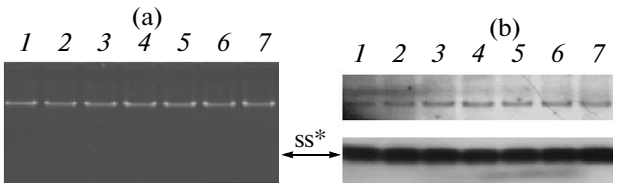

Fig. 5. (a) Agarose gel electrophoresis of the linearized pE10 plasmid incubated with $d(TG)_{10}$ at 37°C for the following time (without cooling the samples in ice before loading): (*1*) 0 min, (*2*) 1 min, (*3*) 15 min, (*4*) 30 min, (*5*) 1 h, (*6*) 6 h, and (*7*) 24 h. (b) Autoradiogram of gel in (a); ss* is radioactively labeled (single-stranded) oligonucleotides not bound to duplex.

DISCUSSION

Earlier studies of our laboratory have demonstrated the phenomenon of interaction between a linear DNA duplex and a single-stranded oligonucleotide complementary to the terminal region of one of the strands forming the duplex; this interaction brings about a triple-stranded complex with invasion of the oligonucleotide [19]. In this process, the oligonucleotide homologous to the inner region of the double-stranded fragment is not observed. In this work, we have found the specific features in invasion of oligonucleotides in the linear and circular duplexes containing $(CA/TG)_{31}$ repeats.

Similar to the situation with fragments of a random sequence, analysis of the interaction between oligonucleotides and the linear duplexes carrying $(CA/TG)_{31}$ repeats in their inner region has demonstrated the invasion of the single-stranded oligonucleotides complementary to the 3'-end of the duplex strand. Invasion is observed into both the fragments amplified by PCR and linearized plasmid. However, unlike the oligonucleotides with a random sequence, the oligonucleotides with repetitive sequence are able to invade the inner homologous region within the linear duplex containing $(CA/TG)_{31}$ repeats. These data suggest that DNA in the region of CA/TG repeats has specific conformational features.

It is known that the error rate during DNA replication is by one order of magnitude higher in the region of microsatellite repeats as compared with a random nucleotide sequence [25–27]. The presence of duplexes with different number of repeats in PCR solution can lead to their interaction with one another due to DNA denaturation and renaturation cycles with equiprobable formation of both homoduplexes, containing the same copy number of repeats, and he-

teroduplexes, containing different number of repeats in complementary strands. Therefore, the discovered invasion of $d(CA)_{12}$ oligonucleotides in linear double-stranded fragments obtained during PCR can result from the interaction between the oligonucleotide in question and the unpaired TG repeats of heteroduplexes. The in vivo mutation frequency in the region of microsatellite repeats is also higher than in the nonrepetitive sequence [28].

When isolating plasmid DNA from cells, we also can obtain a mixture of the plasmids with different numbers of repeats. As we have earlier demonstrated, linearized DNA duplexes can interact with one another through the association of terminal base pairs with subsequent strand exchange between duplexes [29, 30]. This suggests that heteroduplexes can be also formed in solutions of linear plasmid DNA fragments. However, it has been shown that in the absence of any defects in the mismatch repair, the frequency of deletions/insertions involving (CA/TG) repeats during the replication of *E. coli* strain XL1 blue is by one order of magnitude lower as compared with an in vitro amplification [31]. Thus, the fraction of heteroduplexes in solution of linearized plasmid should be considerably smaller, while the invasion intensity of $d(CA)_{10}$ and $d(TG)_{10}$ oligonucleotide in this DNA fragment is comparable with the intensity of oligonucleotide invasion in the fragment synthesized by PCR. This assumption also fails to explain the interaction between oligonucleotides and circular supercoiled plasmid.

The hypothesis on a partial denaturation of linear duplexes in the repetitive region, which enhances the invasion of complementary oligonucleotides in this region, also has not found confirmation in the earlier studies of the structure of DNA carrying $(CA/TG)_{31}$ repeats. In particular, study of chemicatenanes, nonlinear intramolecular structures capable of forming under certain conditions from pE10 restriction fragments, a linear 120-bp fragment of this plasmid was used as a control for chemical and enzymatic testing [32]. Unlike chemicatenanes, the linear DNA duplex treated with S1 and P1 nucleases displayed no stable regions of single-stranded loops in both the repetitive region and beyond it. However, it is quite possible that addition of single-stranded oligonucleotides to the duplex-containing solution weakens the base-stacking in the repetitive region, thereby providing further invasion.

Another explanation for the oligonucleotide invasion is that CA/TG repeats form a structure distinct from the classical right-handed helical structure (B-DNA). One of such structures, left-handed Z-DNA, is formed, in particular, by the sequences with alternating pyrimidines and purines. Note here that the left-handed DNA helix containing $(CA/TG)_{31}$ repeats considerably differs in its chemical reactivity and thermodynamic characteristics from the classical Z-DNA formed of $(CG)_n$ repeats [34]. Since the physical parameters of Z- and B-DNAs considerably differ, the base pairing at the B–Z interface is damaged [34]. Therefore, the invasion of $d(CA)_{10}$ and $d(TG)_{10}$ oligonucleotides in the duplex can be faster at the B–Z interface. However, this assumption is adequate only for a circular supercoiled plasmid, because Z-form under physiological ionic conditions requires a negative supercoiling [35]. This assumption can explain a considerably more intensive invasion of oligonucleotides in a circular supercoiled plasmid as compared with a linear variant (see Fig. 6).

The possibility of oligonucleotide invasion in CA/TG repeats within linear duplexes can be associated with a conformational plasticity of these dinucleotides. According to crystallography data and molecular simulation, the largest variation range of helix parameters is characteristic of the pyrimidine–purine dinucleotides and the most pronounced is for d(C-A)/d(T-G) [36, 37]. It has been also demonstrated by NMR that the base stacking in d(C-A-C) sequence in DNA is somewhat weakened [38]. Nonetheless, these data fail to explain the fact of different intensities for $d(CA)_{10}$ and $d(TG)_{10}$ oligonucleotides in their interaction with linear duplex depending on their molar ratio. Results of this work demonstrate a very high affinity of $d(TG)_{10}$ for the CA/TG repetitive region in double-stranded DNA, which is explainable by the ability of TG repeats to form noncanonical structures, whereas the CA strand remains unpaired and interacts with a complementary oligonucleotide. An example of noncanonical structures is intramolecular quadruplexes containing guanine quartets; recently, Kalyuzhny et al. [24] demonstrated the formation of such quartets by $d(GT)_n$ sequences. Such structures are most stable at a temperature of 1 to 3°C. When studying the kinetics of oligonucleotide invasion, we discovered that $d(TG)_{10}$ invaded linear duplex during the very first seconds of mixing in ice directly before loading on gel with subsequent electrophoretic separation in the cold. An explanation of the observed effect is formation of stable TG quadruplexes in the linearized plasmid at a low temperature, since they enhance the interaction of $d(TG)_{10}$ with the CA strand of the duplex. However, the interaction pattern of this oligonucleotide in the experiment without placing DNA samples in ice was analogous to the previous pattern. Therefore, in the future experiments, we will select the conditions that would inhibit the interaction of $d(TG)_{10}$ with the duplex, so that it could, possibly, allow us to speak about the formation of a particular structure by CA/TG repeats.

All the above described considerations about the specific structural features of CA/TG repeats that enhance oligonucleotide invasion are of hypothetical character and imply further studies with involvement of physical and chemical methods.

Our results are of great interest from biological standpoint, because they are related to the mechanism underlying the function of CA repeats as recombina-

tion hot spots, in particular, their role in recombination initiation.

Among the major proteins involved in homologous recombination are the bacterial RecA and the RecA-like homologs of eukaryotes. At the initial stages of recombination, this protein provides for interaction of single-stranded end of the DNA molecule with the homologous double-stranded molecule to form a structure analogous to D-loop [39]. This structure was for the first time isolated from *S. cerevisiae* by Hunter et al. [40] when studying homologous recombination and was referred to as a single-end invasion. This structure is formed with transition from double-stranded DNA breaks to double Holliday junctions. Despite a complexity and multifactorial character of recombination, with both DNA conformation and DNA interactions with manifold proteins as major players, the discovered in vitro invasion of single-stranded DNA in the duplex can be an in vivo DNA characteristic important for initiating Holliday structures during recombination.

Numerous data suggest the existence of hot spots for recombination—the regions where recombination takes place considerably more frequently than expected. Such hot spots have been discovered in bacteriophages, bacteria, yeasts, and mammals, including humans [6, 41]. The DNA sequences (including repetitive sequences) able to form various noncanonical structures belong to the genome regions that increase the recombination frequency in both in vitro and in vivo experiments. $(CA/TG)_n$ repeats influence the frequency of homologous recombination as well as the localization of recombination events predominantly in the vicinity of or within microsatellite regions [2, 3]. Several studies have demonstrated the role of other microsatellite repeats capable of forming hairpins, H- and Z-DNAs, in the induction of double-stranded breaks and genetic instability in mammals, both in the experiments with cell lines and in vivo [17, 18, 42]. However, the mechanisms of the observed processes still remain vague. The discovered phenomenon of invasion of single-stranded DNA in the homologous region of CA/TG repeats within a duplex can be useful for further studies into the mechanisms underlying the genetic rearrangements involving microsatellites.

ACKNOWLEDGMENTS

The work was supported by the Russian Foundation for Basic Research (project no. 07-04-92171).

REFERENCES

1. Lander E.S., Linton L.M., Birren B., et al. 2001. Initial sequencing and analysis of the human genome. *Nature*. **409**, 860–921.
2. Treco D., Thomas B., Arnheim N. 1985. Recombination hot spot in the human beta-globin gene cluster: Meiotic recombination of human DNA fragments in *Saccharomyces cerevisiae*. *Mol. Cell Biol.* **5**, 2029–2038.
3. Gendrel C.-G., Boulet A., Dutreix M. 2000. $(CA/GT)_n$ microsatellites affect homologous recombination during yeast meiosis. *Genet. Dev.* **14**, 1261–1268.
4. Murphy K.E., Stringer R. 1986. RecA independent recombination of poly d(GT)-d(CA) in pBR322. *Nucleic Acids Res.* **14**, 7325–7340.
5. Dutriex M. 1997. (GT)n repetitive tracts affect several stages of RecA-promoted recombination. *J. Mol. Biol.* **273**, 105–113.
6. Majewski J., Ott J. 2000. GT repeats are associated with recombination on human chromosome 22. *Genet. Res.* **10**, 1108–1114.
7. Adachi M., Tsujimoto Y. 1990. Potential Z-DNA elements surround the breakpoints of chromosome translocation within the 5' flanking region of *bcl*-2 gene. *Oncogene*. **5**, 1653–1657.
8. Ford A.M., Bennet C.A., Price C.M., Bruin M.C.A., van Wering E.R., Greaves M. 1998. Fetal origins of the *TEL-AML* fusion gene in identical twins with leukemia. *Proc. Natl. Acad. Sci. USA*. **95**, 4584.
9. Thandla S.P., Ploski J.E., Raza-Egilmez S.Z., Chhalliyil P.P., Block A.M.W., de Jong P.J., Aplan P.D. 1999. *ETV6-AML1* translocation breakpoints cluster near a purine/pyrimidine repeat region in the *ETV6* gene. *Blood*. **93**, 293–299.
10. Naylor L.H., Clark E.M. 1990. $d(TG)_n \cdot d(CA)_n$ sequences upstream of the rat prolactin gene form Z-DNA and inhibit gene transcription. *Nucleic Acids Res.* **18**, 1595–1601.
11. Shimajiri S., Arima N., Tanimoto A., Murata Y., Hamada T., Wang K.Y., Sasaguri Y. 1999. Shortened microsatellite $d(CA)_{21}$ sequence down-regulates promoter activity of matrix metalloproteinase 9 gene. *FEBS Lett.* **455**, 70–74.
12. Hui J., Hung L.H., Heiner M., Schreiner S., Neumüller N., Reither G., Haas S.A., Bindereif A. 2005. Intronic CA-repeat and CA-rich elements: A new class of regulators of mammalian alternative splicing. *EMBO J.* **24**, 1988–1998.
13. Lorenz M., Hewing B., Hui J., Zepp A., Baumann G., Bindereif A., Stangl V., Stangl K. 2007. Alternative splicing in intron 13 of the human *eNOS* gene: A potential mechanism for regulating eNOS activity. *FASEB J.* **21**, 1556–1564.
14. Yu J., Hai Y., Liu G., Fang T., Kung S.K., Xie J. 2009. The heterogeneous nuclear ribonucleoprotein L is an essential component in the Ca^{2+}/calmodulin-dependent protein kinase IV-regulated alternative splicing through cytidine-adenosine repeats. *J. Biol. Chem.* **284**, 1505–15013.
15. Rossbach O., Hung L.H., Schreiner S., Grishina I., Heiner M., Hui J., Bindereif A. 2009. Auto- and cross-regulation of the hnRNP L proteins by alternative splicing. *Mol. Cell. Biol.* **29**, 1442–1451.
16. Benet A., Azorin F. 1999. The formation of triple-stranded DNA prevents spontaneous branch-migration. *J. Mol. Biol.* **294**, 851–857.
17. Wang G., Vasquez K.M. 2004. Naturally occurring H-DNA-forming sequences are mutagenic in mammalian cells. *Proc. Natl. Acad. Sci. USA*. **101**, 13448–13453.

18. Wang G., Vasquez K.M. 2006. Non-B DNA structure induced genetic instability. *Mutat. Res*. **598**, 103–119.
19. Gasanova V.K., Neschastnova A.A., Belitskii G.A., Yakubovskaya M.G. 2006. Specific oligonucleotide invasion into an end of a DNA duplex. *Mol. Biol*. **40**, 132–138.
20. Saenger W. 1984. *Principles of Nucleic Acid Structure*. N.Y.: Springer.
21. Gaillard C., Strauss F. 1994. Association of poly(CA).poly(TG) DNA fragments into four-stranded complexes bound by HMG1 and 2. *Science*. **264**, 433–436.
22. Sambrook J., Russell D.W. 2001. *Molecular Cloning: A Laboratory Manual*, 3rd ed., Cold Spring Harbor, NY: Cold Spring Harbor Lab. Press, vol. 1.
23. Yakubovskaya M.G., Neschastnova A.A., Popenko V.I., Lipatova Zh.V., Belitsky G.A. 1999. Holliday junctions are formed in concentrated solutions of purified products of DNA amplification. *Biokhimiya*. **64**, 1550–1554.
24. Kaluzhny D., Shchyolkina A., Livshits M., Lysov Y., Borisova O. 2009. A novel intramolecular G-quartet-containing fold of single-stranded d(GT)(8) and d(GT)(16) oligonucleotides. *Biophys. Chem*. **143**, 161–165.
25. Hite J.M., Eckert K.A., Cheng K.C. 1996. Factors affecting fidelity of DNA synthesis during PCR amplification of $d(C\text{-}A)_n \cdot d(G\text{-}T)_n$ microsatellite repeats. *Nucleic Acids Res*. **24**, 2429–2434.
26. Eckert K., Mowery A., Hile S. 2002. Misalignment-mediated DNA polymerase beta mutations: Comparison of microsatellite and frame-shift error rates using a forward mutation assay. *Biochemistry*. **41**, 10490–10498.
27. Shinde D., Lai Y., Sun F., Arnheim N. 2003. *Taq* DNA polymerase slippage mutation rates measured by PCR and quasi-likelihood analysis: $(CA/GT)_n$ and $(A/T)_n$ microsatellites. *Nucleic Acids Res*. **31**, 974–980.
28. Debrauwere H., Gendrel C.G., Lechat S., Dutreix M. 1997. Differences and similarities between various tandem repeat sequences: Minisatellites and microsatellites. *Biochimie*. **79**, 577–586.
29. Yakubovskaya M.G., Neschastnova A.A., Humphrey K.E., et al. 2001. Interaction of linear homologous DNA duplexes via Holliday junction formation. *Eur. J. Biochem*. **268**, 7–14.
30. Markina V.K., Danilova O.A., Neschastnova A.A., Belitskii G.A., Yakubovskaya M.G. 2002. The role of duplex ends in the spontaneous interaction of homologous linear DNA fragments. *Mol. Biol*. **36**, 693–697.
31. Bichara M., Pinet I., Schumacher S., Fuchs R.P. 2000. Mechanisms of dinucleotide repeat instability in *Escherichia coli*. *Genetics*. **154**, 533–542.
32. Gaillard C., Shlyakhtenko L.S., Lyubchenko Y.L., Strauss F. 2002. Structural analysis of hemicatenated DNA loops. *BMC Struct. Biol*. **2**, 7.
33. Kladde M.P., Kohwi Y., Kohwi-Shigematsu T., Gorski J. 1994. The non-B-DNA structure of $d(CA/TG)_n$ differs from that of Z-DNA. *Proc. Natl. Acad. Sci. USA*. **91**, 1898–1902.
34. Ha S.C., Lowenhaupt K., Rich A., Kim Y.G., Kim K.K. 2005. Crystal structure of a junction between B-DNA and Z-DNA reveals two extruded bases. *Nature*. **437**, 1183–1186.
35. Singleton C.K., Klysik J., Stirdivant S.M., Wells R.D. 1982. Left-handed Z-DNA is induced by supercoiling in physiological ionic conditions. *Nature*. **299**, 312–316.
36. Suzuki M., Yagi N. 1995. Stereochemical basis of DNA bending by transcription factors. *Nucleic Acids Res*. **23**, 2083–2091,
37. Lankas F., Sponer J., Langowski J., Cheatham T.E.3rd. 2003. DNA basepair step deformability inferred from molecular dynamics simulations. *Biophys. J*. **85**, 2872–2883.
38. Cheung S., Arndt K., Lu P. 1984. Correlation of *lac* operator DNA imino proton exchange kinetics with its function. *Proc. Natl. Acad. Sci. USA*. **81**, 3665–3669.
39. Kowalczykowski S.C., Dixon D.A., Eggleston A.K., Lauder S.D., Rehraurer W.M. 1994. Biochemistry of homologous recombination in *Escherichia coli*. *Microbiol. Rev*. **58**, 401–465.
40. Hunter N., Kleckner N. 2001. The single-end invasion: An asymmetric intermediate at the double-strand break to double-Holliday junction transition of meiotic recombination. *Cell*. **106**, 59–70.
41. Wahls W.P. 1998. Meiotic recombination hotspots: Shaping the genome and insights into hypervariable minisatellite DNA change. *Curr. Top. Dev. Biol*. **37**, 37–75.
42. Wang G., Carbajal S., Vijg J., DiGiovanni J., Vasquez K.M. 2008. DNA structure-induced genomic instability *in vivo*. *J. Natl. Cancer Inst*. **100**, 1815–1817.