J. Biochem. **139**, 35–39 (2006)
doi:10.1093/jb/mvj009

A Palindromic Repeat Sequence Adopts a Stable Fold Back Structure under Supercoiling

Arun K. Shukla[1],* and Kunal B. Roy[1],†
[1]*Centre for Biotechnology, Jawaharlal Nehru University, New Mehrauli Road, New Delhi-67, India*

Received August 8, 2005; accepted August 29, 2005

A synthetic deoxyoligonucleotide containing five palindromic repeats of GGATCC self assembles to form a parallel four-stranded structure held together by G-tetrads that shows slower mobility than duplex DNA. This structure is hypersensitive to S1 nuclease and resistant to DMS modification. The same oligonucleotide when cloned in a plasmid forms a different structure under supercoiling that persists stably even in the cleaved out insert. On polyacrylamide gel electrophoresis, the cleaved out insert moves to a position midway between the duplex and parallel four-stranded forms of the oligonucleotide. Upon S1 nuclease treatment, the cleaved out insert shows a discreet band of 18 base pairs, suggesting an unfolded region in the middle. All the guanines in the cleaved out insert are sensitive to DMS modification and produce a positive peak at 285 nM in the circular dichroism spectrum, a signature of fold back tetraplex structures. We propose a fold back quadruplex structure for the insert under supercoiling with only A.T.A.T and G.C.G.C tetrads. This is the first suggestive evidence of a general tetraplex motif without G quartets as that proposed for generalized recombination.

Key words: CD spectra, DMS modification, fold back quadruplex, guanine tetrads, S1 nuclease, supercoiling.

There is renewed interest in DNA quadruplex structures for their possible role in biological processes ranging from telomere function (*1*), gene regulation, homologous recombination (*2*, *3*), and even in therapeutic approaches (*4*, *5*). DNA quadruplex structures are highly polymorphic (monomeric, dimeric or tetrameric) and have been studied mostly with G-rich sequences (*6*). The formation and stability of these quadruplex structures is driven by self association of guanine bases by Hoogstein pairing into G4-tetrad motifs and the coordination of Na^+ or K^+ ions in the axial core of the structure (*7*). Only recently, NMR studies have identified quadruplex structures in which G4-tetrads coexist with A.T.A.T and G.C.G.C tetrads (*8*, *9*). The mixed tetrads form by aligning two Watson-Crick duplexes along the major groove face. This type of major groove association in forming the tetraplex structure was proposed long ago as a model for homologous recombination (*2*, *3*). Despite of discovery of an increasing number of quadruplex binding proteins (*5*), definite demonstration of the presence of a quadruplex structure *in vivo* remains elusive. A.R. Morgan proposed experiments with synthetic palindromic repeats cloned into a supercoiled plasmid to detect putative quadruplexes inside the cell (*3*). We synthesized a 36-mer deoxyoligonucleotide, 5′-AATTC $(GGATCC)_5$G, and cloned it into a plasmid vector to check the structural transition, if any, under supercoiling. We wanted to test Morgan's hypothesis and to detect the putative tetraplex structure under supercoiling. The robust structure formed by this oligonucleotide under supercoiling was probed by CD spectral analysis, S1 nuclease assay and dimethyl sulphate (DMS) modification experiment. We present evidence for a structural transition in RY36 from a linear quadruplex involving G tetrads, *in vitro*, to a stable foldback structure in a plasmid under supercoiling.

MATERIALS AND METHODS

Chemicals and Plasmids—An RY36 oligonucleotide [5′d-AATTC$(GGATCC)_5$G-3′] with *Eco*RI overhangs was purified by 20% denaturing urea PAGE. S1 nuclease, the Klenow fragment of DNA polymerase and restriction endonucleases were purchased from MBI fermentas. Dimethyl sulphate (DMS) was from SIGMA. α^{32}P-ATP was from Amersham Biosciences. Plasmid bluescript (SK^-) was from Invitrogen.

Thermal stability experiment: Oligonucleotide RY36 (50–100 pmol) was dissolved in 20 mM Tris buffer (pH 7.5) without metal ion, incubated at 95°C for 10 min, and allowed to cool to room temperature. This resulted in the RY36 duplex structure, which was then radiolabeled using Klenow polymerase in the presence of α^{32}P-ATP (as per the manufacturer's protocol). This radiolabeled sample was then incubated with 100 mM M^+ (LiCl, NaCl, KCl, CsCl) and 5 mM $MgCl_2$ at 95°C for 10 min, and allowed to cool. Aliquots were collected at different temperatures, flash frozen by plunging into liquid nitrogen quickly, and stored at –80°C until used. Samples were analyzed by electrophoresis in 20% non-denaturing polyacrylamide gels using 1× TB buffer (Tris 89 mM, Boric acid 89 mM, pH 8.0) and used for autoradiography.

Plasmid Construct—RY36 oligonucleotide was cloned into the *Eco*RI site of the Multiple Cloning site of plasmid

*Present address: Department of Molecular Membrane Biology, Max Planck Institute of Biophysics, Frankfurt am main 60438, Germany. Fax: +49-69-63031002, E-mail: arun.shukla@mpibp-frankfurt.mpg.de
†To whom correspondence should be addressed. E-mail: kunalrb@hotmail.com

bluescript to obtain the pKBR[36] construct. The construct was verified by DNA sequencing.

Purification of the Cleaved Out Insert—pKBR[36] was digested with *Eco*RI and the cleaved out insert was separated in an ethidium bromide stained acrylamide gel. A gel piece containing the insert was soaked in extraction buffer (500 mM NH_4Oac,0.1% SDS,0.1 mM EDTA) for one hour at 37°C. The supernatant was removed by centrifugation and then precipitated with 0.1 volume 3 M NaOAc and 2.5 volumes of ethanol. The insert was finally resuspended in Tris buffer and quantified by measuring absorbance at 260 nm for use in further experiments.

S1 Nuclease Assay—DNA samples were radiolabeled by using Klenow polymerase in the presence of α^{32}P-ATP. Radiolabeled DNA samples (20–50 pmol) were treated with S1 nuclease (5–10 units) in reaction buffer (300 mM potassium acetate, pH 4.6, 2.5 mM NaCl, 10 mM $ZnSO_4$ and 50% glycerol) for 2 hours at 37°C. The reaction was stopped by heating the samples at 65°C for 15 min. The DNA samples were extracted once with phenol-chloroform and then recovered by ethanol precipitation. The digested products were resolved in 20% non-denaturing polyacrylamide gels and visualized by autoradiography. The S1 nuclease assay was also performed on supercoiled pKBR[36] followed by *Eco*RI digestion and analysis in ethidium bromide stained polyacrylamide gel.

DMS Modification—1% DMS was used in reaction buffer buffer (60 mM NaCl, 10 mM Tris-Cl, pH 7.0, 10 mM $MgCl_2$, 2 mM EDTA) containing 50–100 pmol radiolabeled DNA (labeled as in the S1 nuclease assay) for 10 min at 37°C. The reaction was stopped by adding stop buffer (1.5 M NaOAc, 1 mM mercaptoethanol) for 15 min on ice. The DMS modified DNA was subjected to piperidine cleavage, and the cleavage products were resolved by 20% sequencing PAGE. In addition, DMS modification was also performed on pKBR[36], and the insert was cleaved out and analyzed as above. Standard Maxam-Gilbert sequencing of the RY36 duplex was also performed as described earlier (*13*).

Circular Dichroism Measurements—Circular dichroism measurements were obtained on a Jasco J-715A spectropolarimeter with a thermostatted cuvette holder and 1 cm quartz cuvette. In typical experiments, 50–100 pmol of duplex RY36 or cleaved out insert samples in Tris buffer (20 mM, pH 7.5) containing 10 mM $MgCl_2$ were used. The self-assembled structure was measured in the same buffer except that it contained 100 mM NaCl. All experiments were performed at room temperature. The data are the averages of three independent sets of measurements. Each spectrum was corrected for the spectral contributions of the buffer. Spectra were plotted as $\Delta\alpha$ [(L-R)] in units of M^{-1} cm^{-1}.

RESULTS AND DISCUSSION

RY36 was dissolved in 20 mM Tris buffer, pH 7.5 containing 100 mM NaCl. However, when heated to 95°C and slowly annealed to room temperature, it migrated more slowly than a normal duplex DNA of approximately the same size. Interestingly, when RY36 was dissolved in the same buffer but without NaCl and annealed to room temperature, it migrated like a normal duplex DNA. This indicates that, probably, in presence of Na^+, RY36 is adopting a structure by itself during slow annealing. Since the RY36 oligonucleotide is guanine rich, it is predicted that it can self-assemble through G tetrads that are stabilized by the Na^+ present in the buffer. In order to check this possibility, a systematic thermal stability experiment was performed, in which characteristic of structures involving guanine tetrads were found.

RY36 Self-Assembles into a Structure That is Stabilized by Metal Ion—The thermal stability of the self-assembled structure of RY36 was checked in the presence of different metal ions (Fig. 1). Maximum thermal stability of the structure was seen in the presence of K^+ ion, where it maintains its structure to 70°C. The presence of Na^+ had a moderate effect while Cs^+ and Li^+ had little or no effect. This kind of metal ion dependence is characteristic of structures containing guanine tetrads; therefore, the data indicate that RY36 self-assembles via guanine tetrads. This rules out the possibility of any cruciform structure in self assembled RY36. For guanine tetrads, K^+ seems to fit best in the coordination sphere while Li^+ and Cs^+ ions are either too small or too big for coordination complex formation. We observed the appearance of single stranded DNA at higher temperatures which corresponds to the loss of duplex RY36 and self assembled structure. In Fig. 1 single stranded DNA is not seen as the gel electrophoresis was performed for a longer time in order to acheive better resolution.

RY36 Adopts a Different Structure under Supercoiling—RY36 was then cloned into the pBS (SK^-) vector at the *Eco*RI site, resulting in the pKBR[36] construct. *E. coli* cells

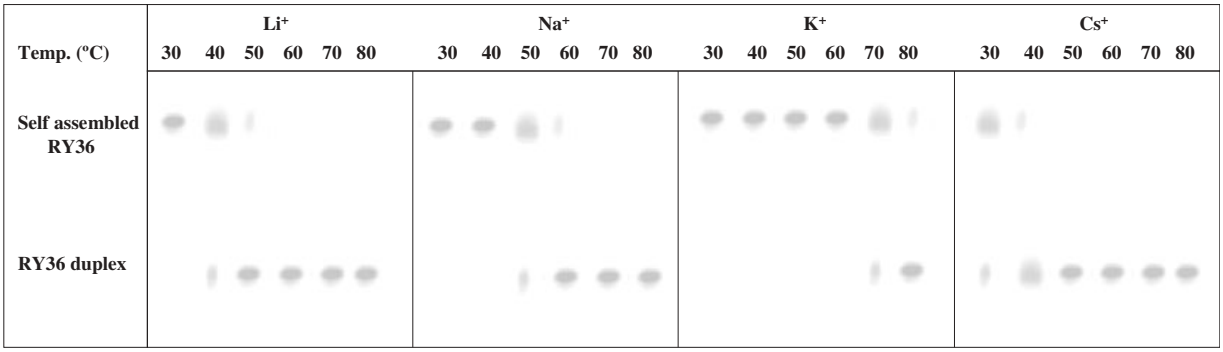

Fig. 1. **Thermal stability of the RY36 oligomer in the presence of 100 mM of various monovalent cations.** The upper bands correspond to the self-assembled structure while the lower bands represent the duplex.

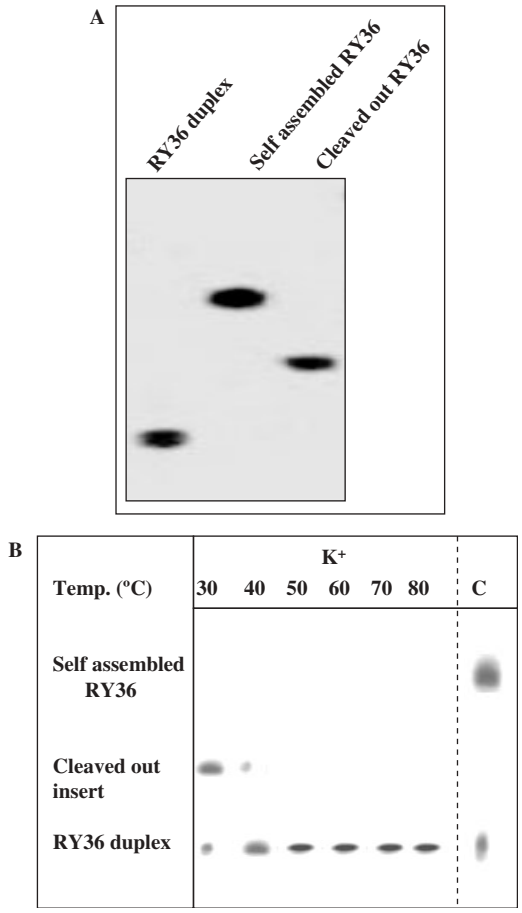

Fig. 2. **A: Unusual mobility of self-assembled RY36 and cleaved out RY36 in polyacrylamide gel electrophoresis. B: Thermal stability of the cleaved out insert in the presence of 100 mM K^+.** Lane "C" represents self-assembled structure in the presence of 100 mM K^+ at 70°C as a control. Other cations (Li^+, Na^+ and Cs^+ also did not stabilize the cleaved out insert).

containing this construct were grown in the presence of chloramphenicol and used to check if the insert undergoes any structural transition under supercoiling. When RY36 was cleaved out from the plasmid, it again migrated differently than normal duplex DNA. Surprisingly, its mobility was faster than that of the self-assembled structure *in vitro*, but slower than normal duplex DNA (Fig. 2A). Additionally, this structure was not stabilized by metal ions (Fig. 2B, thermal stability of cleaved out insert in the presence of K^+ is shown), and, therefore, indicates that it is significantly different from the self assembled structure.

CD Spectra Support a Foldback Structure for the Cleaved Out Insert but a Linear Quadruplex for Self-Assembled RY36—In order to determine the secondary structures of the cleaved out insert and self-assembled RY36, circular dichroism spectroscopy was used. The spectrum of the self-assembled structure showed a positive peak at around 265 nm and a negative peak at around 236 nm (Fig. 3). This is a characteristic of linear quadruplex structures with G tetrads (*14*, *15*) and thus supports a linear quadruplex formation within the self-associated RY36 involving G tetrads. This further rules out any

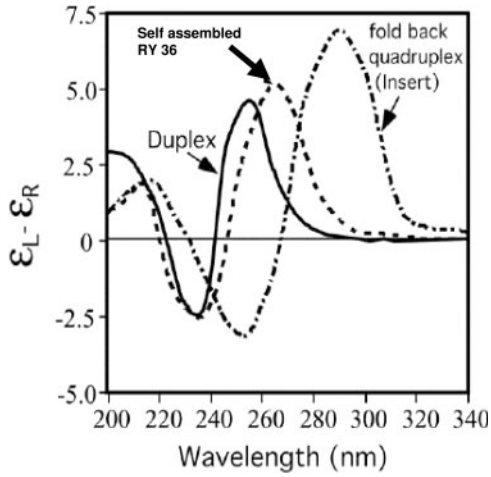

Fig. 3. **CD spectra corresponding to the cleaved out insert, self-assembled structure of RY36 and RY36 duplex.**

possibility of a cruciform structure. On the other hand, the CD spectrum of the cleaved out insert showed a positive peak at around 285 nm and a negative peak at around 255 nm, which is a signature of fold back quadruplex structures (*15*). Both spectra are quite different from the spectrum obtained for the duplex RY36, suggesting the presence of different complexes in the same sequence under different experimental conditions. At this point, we hypothesized that the self-assembled structure of RY36 would resemble a "bouquet" in which the strands would probably remain single and unassociated between the G-tetrads. This bouquet shaped structure should be hypersensitive to S1 nuclease. However, if the cleaved out insert forms a fold back structure, it will contain unpaired bases only in the folding region.

A single stranded region is present around the middle of the cleaved out insert: To further confirm if the cleaved out insert exists as a foldback structure and to identify the folding region, we used S1 nuclease mapping. As shown in Fig. 4, S1 nuclease digestion of cleaved out insert resulted in a band of 18 bp (lane 2), while the self-assembled structure was almost completely degraded (lane 6). These data indicate that a folded, and thus unpaired, region is present around the middle of the cleaved out insert. However, in the self-assembled structure, unpaired regions are present between each of five G-tetrad pairs, leading to it being completely digested by S1 nuclease. A very faint band parallel to the 18 bp band was seen in lane 6 also. We think that the palindromic repeat has a propensity to fold by itself but has a kinetic disadvantage to form a fold back quadruplex. Probably, a miniscule population of fold back quadruplex might have formed but escaped detection. Duplex DNA did not show any S1 nuclease activity as expected. In addition, the S1 nuclease assay was performed on intact plasmids that were subsequently digested by *Eco*RI. This also resulted in the appearance of an 18 bp fragment on the acrylamide gel (Fig. 4B, lane 3). This suggests that the cleaved out insert has the same folding pattern as under supercoiling. We think that the structure is robust and can maintain itself once it is formed.

Guanines are accessible to DMS modification in the cleaved out insert: We also did DMS modification studies

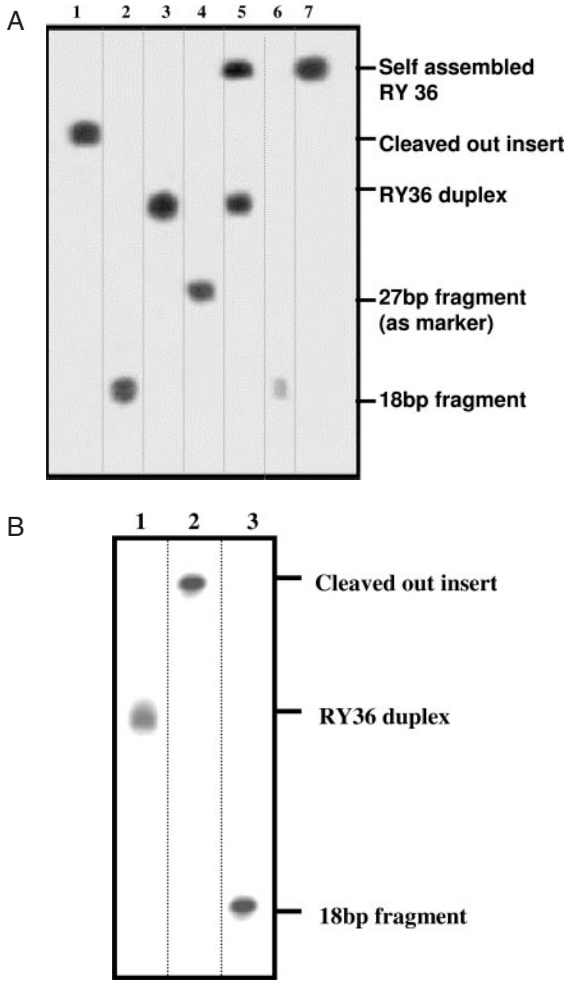

Fig. 4. **A: S1 nuclease assay performed on different forms of RY36.** Lane 1: Cleaved out insert, Lane 2: S1 digestion of the cleaved out insert, Lane 3: Duplex RY36, Lane 4: 27 bp DNA fragment as a marker, Lane 5: Self assembled RY36 at Tm, Lane 6: S1 digestion of self assembled RY36, Lane 7: Self assembled RY36. **B: S1 nuclease assay on supercoiled pKBR[36]**. Lane1: RY36 duplex, Lane 2: pKBR[36] digested with *Eco*RI, Lane 3: pKBR[36] treated with S1 nuclease followed by *Eco*RI digestion.

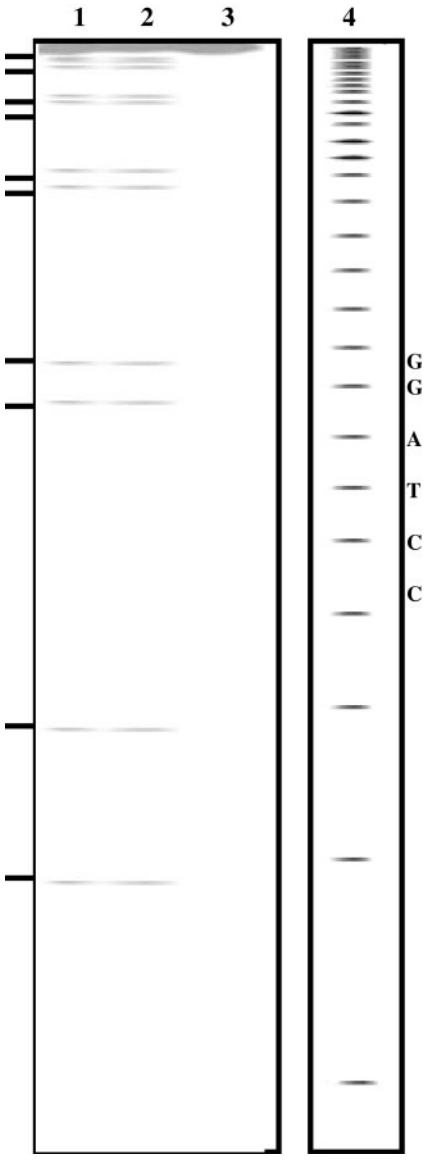

Fig. 5. **DMS modification pattern of RY36 duplex (Lane 1), cleaved out insert (Lane 2) and self-assembled RY36 (Lane 3).** Lines on the left represent guanine bases. Maxam Gilbert sequencing of the RY36 duplex is shown in Lane 4. One repeat has been marked.

on the cleaved out insert and self assembled structure. DMS is a small molecule that can penetrate through the major groove of DNA and reach the Nitrogen at the N7 ring position of guanine and methylate it. However, if guanines are involved in Hoogstein base pairing, they will not be susceptible to N7 modification by DMS. In this experiment, the normal RY36 duplex was used as a control. As shown in Fig. 5, all ten guanines in the duplex showed DMS sensitivity. This is expected, as in the normal Watson Crick duplex structure, the N7 positions of all guanines are accessible to DMS modification. Interestingly, the modification pattern for the cleaved out insert (lane 2) is exactly the same as that of the duplex . This demonstrates that none of the guanines in the fold-back structure of the cleaved out insert is involved in Hoogstein base pairing. Instead, the Watson Crick base pairings are intact in the structure, indicating that it does not involve any G-quartet or GxG.C triplet. It is important to mention here that a similar DMS modification pattern was observed when pKBR[36] supercoiled DNA was used as a template and the insert was cleaved out with *Eco*RI, followed by piperidine treatment and gel analysis. This DMS modification pattern, in addition to the S1 nuclease sensitivity and CD spectrum, therefore, suggests that the structure of the cleaved out insert can be a stable fold back quadruplex, in which two W.C duplexes associate through either their major groove face (*9*) or minor groove face (*11*). The latter type of quadruplex structure stabilized by A.T.A.T and G.C.G.C tetrads has been reported earlier in both crystalline (*12*) and solution states (*9*). We think that superhelical stress forces the RY36 insert in the supercoiled plasmid to fold back. This brings guanines and cytosines on the opposite strands into close proximity and leads to the formation of G.C.G.C. tetrads. Under these conditions, the distance

constraints allow adenines and thymines also to get involved in A.T.A.T. tetrads. However, as guanines are arranged in *syn-syn-anti-anti* conformations, this does not allow the formation of guanine tetrads. Therefore, this structure is insensitive to the presence of metal ions.

On the other hand, the guanines in the self-assembled structure of RY36 did not show sensitivity to DMS. This suggests their involvement in Hoogstein base pairing making the N7 position inaccessible to DMS modification. These data, taken together with the metal ion dependence, S1 nuclease hypersensitivity and CD spectrum, establish the linear quadruplex structure of self-assembled RY36.

The structure we propose here has serious biological implications. A quadruplex structure stabilized by A.T.A.T and G.C.G.C base tetrads only paves a general way for bringing two homologous DNA sequences into register, as a first step in the recombination process prior to strand exchange through the Holiday junction. The observation of four stranded structures poly (CA). poly (TG) fragments and their recognition by HMG proteins (*13*) further suggests that such a tetraplex structural motif is very general and ought to exist *in vivo*. In our current study, we have shown that such a robust foldback quadruplex structure can be formed under supercoiling. Another important point of this study is that we have obtained strong suggestive evidence for Morgan's hypothesis of a fold back tetraplex structure with only A.T.A.T and G.C.G.C tetrads in supercoiled plasmid.

We are thankful to Dr. Tapas Saha for providing RY36 oligo. We also thank Dr. Shyamal Goswami and Sindhu K.V. for technical support in some experiments.

REFERENCES

1. Sen, D. and Gilbert, W. (1988) Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature* **334**, 364–366
2. McGavin, S. (1971) Models of specifically paired like (homologous) nucleic acid structures. *J. Mol. Biol.* **55**, 293–298
3. Morgan, A.R. (1993) Base mismatches and mutagenesis: how important is tautomerism *Trend Biochem. Sci.* **37**, 244–247
4. Shafer, R.H. and Smirnov, I. (2000) Effect of loop sequence and size on DNA aptamer stability. *Biochemistry* **39**, 1462–1468
5. Shafer, R.H. and Smirnov, I. (2001) Biological aspects of DNA/RNA quadruplexes. *Biopolymers* **56**, 209–227
6. Williamson, J.R. (1994) G-quartet structures in telomeric DNA. *Annu. Rev. Biophys. Bimol. Struct.* **23**, 703–730
7. Gilbert, D.E. and Feigon J. (1999) Multistranded DNA structures. *Curr. Opin. Struct. Biol.* **9**, 305–314
8. Kettani, A., Kumar, R.A., and Patel, D.J. (1995) Solution structure of a DNA quadruplex containing the fragile X syndrome triplet repeat. *J. Mol. Biol*. **254**, 636–656
9. Zhang, N., Gorin, A., Majumdar, A., Kettani, A., Chernichenko, N., Stripkin, E., and Patel, D.J. (2001) Dimeric DNA quadruplex containing major groove-aligned A-T-A-T and G-C-G-C tetrads stabilized by inter-subunit Watson-Crick A-T and G-C pairs. *J. Mol. Biol.* **312**, 1073–1088
10. Maxam, A.M. and Gilbert, W. (1977) A method for determining DNA sequence by labeling the end of the molecule and cleaving at the base. *Meth. Enzym.* **65**, 499–560
11. Salisbury, S.A., Wilson, S.E., Powell, H.R., Kennard, O., Lubini, P., and Sheldrick, G. (1997) The bi-loop, a new general four-stranded DNA motif. *Proc. Natl. Acad. Sci. USA* **94**, 5515–5518
12. James, H.T., Susana, C.M., Teixeira, B.C., and Christine, J.C. (2003) Crystal structure of the complementary quadruplex formed by d(GCATGCT) at atomic resolution. *Nucleic Acids Res.* **31**, 844–849
13. Gaillard, C. and Strauss, F. (1994) Association of poly(CA). poly(TG) DNA fragments into four-stranded complexes bound by HMG1 and 2. *Science* **264**, 433–436
14. Vorlickova, M., Zimulova, M., Kovanda, J., Fojtik, P., and Kypr, J. (1998) Conformational properties of DNA dodecamers containing four tandem repeats of the CNG triplets. *Nucleic Acids Res.* **26**, 2679–2685
15. Saha, T. and Usdin, K. (2001) Tetraplex formation by the progressive myoclonus epilepsy type-1 repeat: implications for instability in the repeat expansion diseases. *FEBS Lett.* **491**, 184–187