MCB
Journals.ASM.org

G-Quadruplex Structures Formed at the *HOX11* Breakpoint Region Contribute to Its Fragility during t(10;14) Translocation in T-Cell Leukemia

Mridula Nambiar, Mrinal Srivastava, Vidya Gopalakrishnan, Sritha K. Sankaran, Sathees C. Raghavan

Department of Biochemistry, Indian Institute of Science, Bangalore, India

The t(10;14) translocation involving the *HOX11* gene is found in several T-cell leukemia patients. Previous efforts to determine the causes of *HOX11* fragility were not successful. The role of non-B DNA structures is increasingly becoming an important cause of genomic instability. In the present study, bioinformatics analysis revealed two G-quadruplex-forming motifs at the *HOX11* breakpoint cluster. Gel shift assays showed formation of both intra- and intermolecular G-quadruplexes, the latter being more predominant. The structure formation was dependent on four stretches of guanines, as revealed by mutagenesis. Circular dichroism analysis identified parallel conformations for both quadruplexes. The non-B DNA structure could block polymerization during replication on a plasmid, resulting in consistent K^+-dependent pause sites, which were abolished upon mutation of G-motifs, thereby demonstrating the role of the stretches of guanines even on double-stranded DNA. Extrachromosomal assays showed that the G-quadruplex motifs could block transcription, leading to reduced expression of green fluorescent protein (GFP) within cells. More importantly, sodium bisulfite modification assay showed the single-stranded character at regions I and II of *HOX11* in the genome. Thus, our findings suggest the occurrence of G-quadruplex structures at the *HOX11* breakpoint region, which could explain its fragility during the t(10;14) translocation.

Chromosomal translocations are genetic hallmarks of several cancers and are associated with different types of leukemias and lymphomas (1, 2). The t(10;14)(q24;q11) translocation occurs in about 10% of T-cell acute lymphoblastic leukemias. This involves a reciprocal translocation between chromosomes 10 and 14, wherein breakage on chromosome 10 occurs at the *HOX11* (*TCL3*) gene, which encodes a transcription factor (3–6). Breaks at *HOX11* occur at the 5′ end of the gene, due to an unknown mechanism, resulting in its juxtaposition to the T-cell receptor (TCR) locus. In the case of chromosome 14, breaks occur within the T-cell receptor δ chain locus at the $D\delta_2$ and $D\delta_3$ segments during V(D)J recombination, a process by which antibody and T-cell receptor diversity is generated (7, 8). Interestingly, a cryptic recombination signal sequence (cRSS) containing a heptamer-like sequence, CACAGCC, was found adjacent to the chromosome 10 breakpoint cluster in the *HOX11* gene (see Fig. S1, blue box, in the supplemental material). The t(10;14) results in an altered expression of *HOX11*, which is generally not expressed in T cells (9) and is believed to be a key event in leukemia genesis (3, 10).

Earlier investigators had proposed a V(D)J recombination-mediated mechanism for *HOX11* breakage, due to the presence of cryptic RSS near the patient breakpoints (3–6). However, the cryptic sequence at the breakpoint region was shown not to be recognized by the recombination-activating gene (RAG) complex (11). Studies using intracellular recombination assay with the *HOX11* cRSS also did not show any recombination with a standard RSS (12, 13). Recently, assays using purified RAGs further confirmed the failure of the cRSS at *HOX11* to get recognized and cleaved (14). Hence, the observed lack of recombination at *HOX11* could be due to the inability of RAGs to cleave at the 5′ end of the cryptic heptamer at the RSS in *HOX11*. Another reason could be the incidence of a non-RAG-mediated, RSS-independent mechanism for the generation of t(10;14) translocation.

One of the major causes of genomic instability being actively investigated is the occurrence of non-B DNA in the genome (15–25). Bioinformatic analyses have revealed the presence of several such non-B DNA structures throughout the genome (26, 27). G-quadruplexes, which are four-stranded DNA molecules, are suggested to be present in the promoters of many genes and at the telomeres, thereby performing functions like gene regulation and chromosomal end protection (28–33). Recent studies have also shown that altered DNA structures like cruciform and G-quadruplexes could play a significant role in the generation of chromosomal translocations in certain leukemias and lymphomas (15, 18, 20, 34–37).

In the present study, we investigated the possibility of occurrence of non-B DNA structures at the *HOX11* breakpoint region. We provide biochemical and biophysical evidence for the presence of two independent G-quadruplex structures, flanking either side of the *HOX11* breakpoints. These structures could block replication in a K^+-dependent manner on both single- and double-stranded DNA. More interestingly, mutations at guanine stretches abolished formation of the altered DNA structures on a double-stranded DNA both *in vitro* and inside cells. Finally, we also provide evidence for the presence of single strandedness pertaining to

Received 3 May 2013 **Returned for modification** 11 June 2013
Accepted 23 August 2013
Published ahead of print 3 September 2013
Address correspondence to Sathees C. Raghavan, sathees@biochem.iisc.ernet.in.
M.N. and M.S. are co-first authors.
Supplemental material for this article may be found at http://dx.doi.org/10.1128/MCB.00540-13.

doi:10.1128/MCB.00540-13

the two G-quadruplex motifs at the *HOX11* fragile region within the genome.

MATERIALS AND METHODS

Enzymes, chemicals, and reagents. Chemical reagents were obtained from Sigma Chemical Co. (United States), Amresco (United States), and SRL (India). DNA-modifying enzymes were from New England BioLabs (United States) and Fermentas (United States). Radioisotope-labeled nucleotides were from BRIT (India). Culture media were from Sera Laboratory International Ltd. (United Kingdom), and fetal bovine serum (FBS) and PenStrep were from Gibco BRL (United States).

Oligomeric DNA. The DNA oligonucleotides used in the current study are listed (see Table S1 in the supplemental material). When required, the oligomers were purified using 15 to 18% denaturing polyacrylamide gels (38).

5′-end labeling of oligomers. 5′-end labeling of oligomeric DNA was carried out using T4 polynucleotide kinase in a buffer containing 20 mM Tris-acetate (pH 7.9), 10 mM magnesium acetate, 50 mM potassium acetate, 1 mM dithiothreitol (DTT), and [γ-P^{32}]ATP at 37°C for 1 h (39). The labeled substrates were purified using a G-25 Sephadex size exclusion column and stored at −20°C until further use.

Plasmid construction. pSKS1 was constructed by cloning the *HOX11* wild-type sequence, PCR amplified from the human genomic DNA, in the BamHI site of pMN4, whereas pSKS3 contains the same sequence at the SalI site of pMN4, in the physiological orientation. pMN18 was constructed by cloning the *HOX11* double-mutant PCR fragment, with both regions I and II mutated by site-directed mutagenesis (see below), into pMN4 (40) at the BamHI site in physiological orientation. pMN20 contains mutant region I and wild-type region II of *HOX11*, cloned into the BamHI site of pMN4, in the physiological orientation. pMN21 contains wild-type region I and mutant region II of *HOX11*, cloned in the BamHI site of pMN4, in the physiological orientation. pMS28 was constructed by amplification of a red fluorescent protein (RFP) expression cassette from the pcDNA3.1RFP vector and cloning it into the AflII site of pIRES2-enhanced green fluorescent protein (EGFP). pMS29 was constructed by cloning regions I and II of the *HOX11* breakpoint region (wild type), while pMS31 contained mutations in both regions I and II. pMS30 contains mutant region I and wild-type region II at the BamHI site of pIRES2-EGFP, while pMS34 contains wild-type region I and mutant region II. Each of these plasmids also contained the complete RFP gene expression cassette at the AflII site driving its expression from an independent promoter. See below for details of mutagenesis.

Site-directed mutagenesis. For construction of mutants at *HOX11* regions I and II, site-directed mutagenesis was performed using specific primers containing the desired mutation as described previously (41). Two sets of primers (SKS9-MN130 and MN131-MN35) were used to amplify *HOX11* region I containing mutations in the first round of PCR. The PCR products from these two reactions were then mixed and used as a template for the second round to get the region I amplification product, using primers SKS9 and MN35. Similarly, the two sets of primers (MN134-MN132 and MN133-SKS10) were used to amplify *HOX11* region II containing the mutations. The wild-type region I was amplified using SKS9 and MN35, and region II was amplified using MN134 and SKS10. These mutant PCR products, along with wild-type regions I and II, were mixed in different combinations to get the single and double mutants for regions I and II. These PCR products were cloned and checked by restriction enzyme digestion, followed by sequencing.

Bioinformatics analysis for G-quadruplex formation. The *HOX11* breakpoint region was analyzed for formation of G-quadruplex structures by using the QuadBase and Non-B DB databases as described earlier (34). The parameter for the number of repeated guanines was kept as either two or three, to assess the possibility of formation of both 2G and 3G plate quadruplex structures.

Gel mobility shift assay. Radiolabeled oligomers were incubated in the presence of 10 to 100 mM KCl (as specified) in Tris-EDTA (TE) buffer (pH 8.0) at 37°C for 1 h. For time and concentration-dependent kinetics, 0 to 30 min and 50 to 800 nM of the substrates were employed, respectively. Different forms of G-quadruplexes were then resolved on 12 to 15% native polyacrylamide gels in the presence of 10 to 100 mM KCl, both in the gel and the buffer, at 150 V at room temperature (RT). The gels were dried and exposed to a screen, and signals were detected using PhosphorImager FLA9000 (Fuji, Japan).

CD spectroscopy. The *HOX11* wild-type, mutant, and C-rich complementary oligomers were incubated either in the presence or absence of 100 mM KCl in TE at 37°C for 1 h as described previously (42). For i-motif structure determination, the wild-type and mutant C-rich complementary oligomers and random oligomers were incubated in the presence of 3 M sodium acetate at different pH (4.8, 5.2, and 6.8) for 10 min at RT. The circular dichroism (CD) spectra were recorded at RT from 220 to 300 nm, and 10 cycles were accumulated for every sample, using a JASCO J-810 spectropolarimeter at a scan speed of 50 nm/min. A separate spectrum was measured for buffer alone for 30 cycles, and this was subtracted from all experimental spectra. For destabilization of the G-quadruplex structure, spectra were recorded for samples incubated with KCl at 90°C. Further, a CD spectrum was recorded after the structure was reformed by incubating the same sample overnight at 37°C. The ellipticity was calculated using the Spectra Manager software and plotted as a function of wavelength.

***Taq* polymerase stop assay.** *Taq* polymerase stop assay was performed as described previously (18). The *HOX11* oligomers for region I (MN150 [wild type], MN151 [single mutant], and MN154 [double mutant]), region II (MN152 [wild type], MN 153 [single mutant], and MN155 [double mutant]), random (MN117) (18), and the *HIF1*α promoter region (MN118) (18) were heat denatured and slowly reannealed in the presence of 100 mM KCl or LiCl (for titration experiments, 50,100, 150, and 200 mM) with radiolabeled MN96 or MN119 (for *HIF1*α). Products were resolved on a 12% denaturing PAGE gel and visualized.

Primer extension. The presence of replication blocks due to quadruplex structure formation at the *HOX11* breakpoint region was studied in plasmids pSKS1, pMN18, pMN20, and pMN21 by primer extension. The reactions were carried out by mixing 100 ng of DNA sample in 1× Thermo polymerase buffer [10 mM KCl, 10 mM $(NH_4)_2SO_4$, 20 mM Tris-HCl (pH 8.8), 4 mM $MgSO_4$ and 0.1% Triton X-100], 4 mM $MgSO_4$, 200 μM deoxynucleoside triphosphates (dNTPs), 0.5 μM end-labeled oligomers, and 1 U Vent (exo-)polymerase. Linear amplification primer extensions were carried out in a PCR machine (25 cycles) under the following conditions: 95°C for 3 min (1 cycle), 94°C for 45 s, 58 to 64°C for 45 s (as specified), 72°C for 45 s, and final extension for 3 min. The annealing temperatures of the primers used were 58°C for MN34 and MN35, 60°C for SKS10, and 64°C for MN134. The reactions were terminated by adding a dye containing formamide, and products were resolved on an 8% denaturing polyacrylamide gel. The gel was dried, and signals were detected using PhosphorImager. Sequencing reactions using respective primers were carried out for the wild-type plasmids and loaded along with the primer extension reactions on the gel.

S1 nuclease assay. The wild-type (pSKS3) and mutant (pMN18) plasmids were heat denatured at 95°C and slowly reannealed in the presence of 100 mM KCl overnight in a water bath. One microgram of plasmid was then treated with either 0, 10, 20, 40, or 80 U of S1 nuclease (Fermentas, USA) in a buffer containing 200 mM sodium acetate (pH 4.5 at 25°C), 1.5 M NaCl, and 10 mM $ZnSO_4$ for 30 min at 37°C, and the DNA was purified by phenol-chloroform extraction. The plasmid was then subjected to primer extension with labeled MN34 and MN91 for the G-rich strands of regions I and II, respectively, and MN35 and SKS10 for the C-rich strand of regions I and II, respectively. The reaction products were resolved on an 8% denaturing PAGE gel and analyzed.

DMS protection assay. Dimethyl sulfate (DMS) protection assay was performed as described previously (18). The substrates and intra- and intermolecular G-quadruplexes were gel purified after electrophoresis in the absence or presence of 25 mM KCl. The purified structures were in-

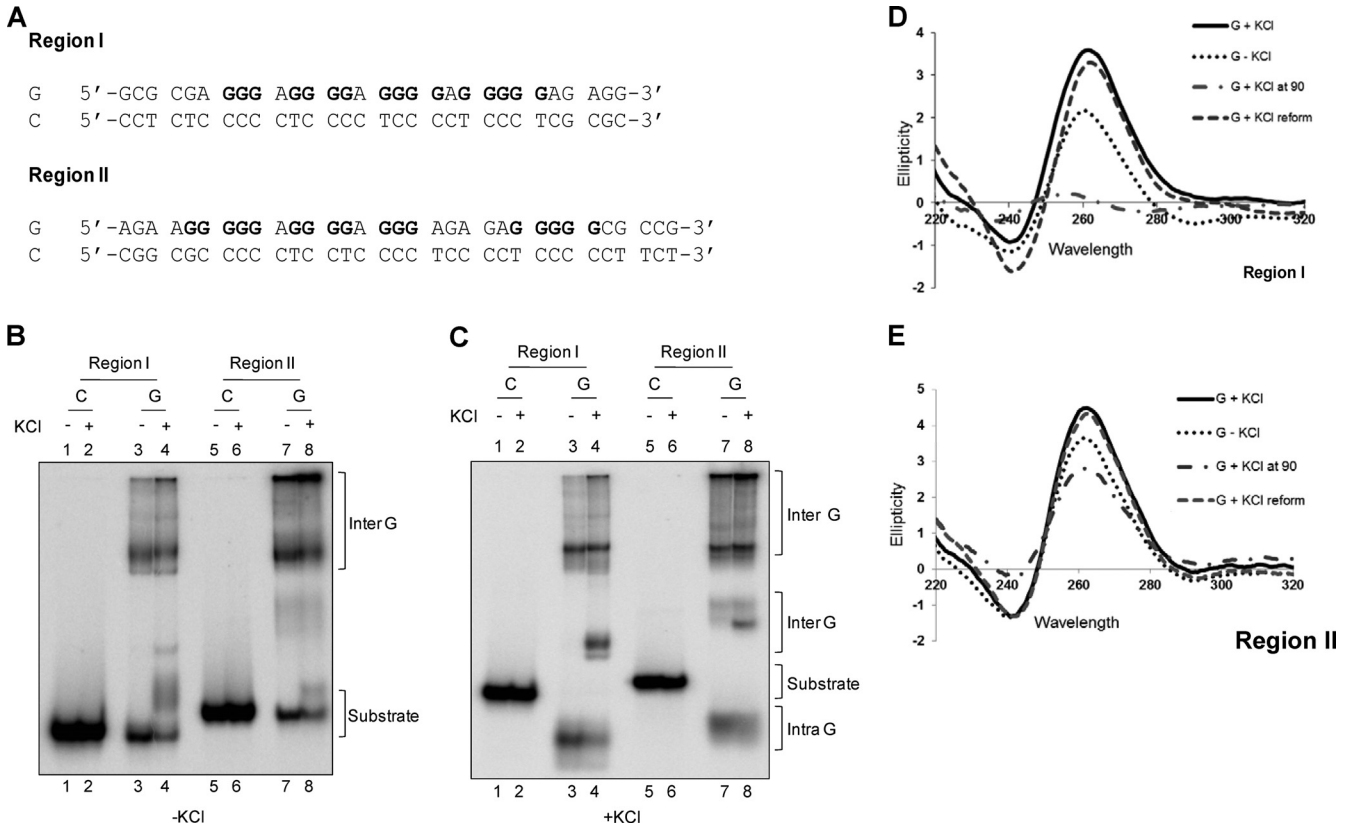

FIG 1 Formation of a G-quadruplex structure in the template strand of regions I and II of the *HOX11* breakpoint region. (A) The oligomeric sequence spanning regions I and II of both G-rich strands and its complementary strands designed from the *HOX11* breakpoint region. (B and C) The G- and C-rich strands from both regions were incubated in the presence of KCl (100 mM) and resolved in the absence (B) or presence (C) of KCl (100 mM), in the gel and running buffer. The substrate, intramolecular (Intra G), and intermolecular (Inter G) quadruplex structures are indicated. (D and E) CD spectra for region I (D) and region II (E) template strands in TE without KCl (G − KCl), with KCl (G + KCl), or with G plus KCl upon heating at 90°C (G + KCl at 90) and then incubation at 37°C overnight (G + KCl reform), in order to reform the structure.

cubated with DMS for 20 min and then treated with piperidine at 90°C (30 min). DNA was purified, and reaction products were resolved on a 15% denaturing polyacrylamide gel, which was further dried and visualized.

Sodium bisulfite modification assay. The sodium bisulfite assay was performed as described previously (19, 43). Briefly, chromosomal DNA was isolated from K562 (chronic myelogenous leukemia) and CEM (T-cell leukemia) cell lines using the nondenaturing method (44). Approximately 5 μg of DNA was incubated in 12.5 μl of 20 mM hydroquinone and 457.5 μl of 2.5 M sodium bisulfite (pH 5.2) for 14 to 16 h at 37°C and purified using the Wizard DNA cleanup kit (Promega, Madison, WI). The bisulfite-modified DNA was desulfonated with 0.3 M NaOH at 37°C for 15 min, ethanol precipitated, and resuspended in 30 μl TE buffer. The *HOX11* breakpoint region from the bisulfite-treated DNA was PCR amplified, resolved on an agarose gel, purified, TA cloned, and sequenced. There is a possibility that the primer sites are also subjected to bisulfite conversion, due to which some molecules lose their ability to serve as PCR templates. However, this basis of background conversion is rare and random.

Test for G-quadruplex formation within cells. REH (B-cell leukemia) cells were transfected with pMS29 (wild type) and pMS31 (mutant for both regions I and II), pMS30 (mutant for region I), or pMS34 (mutant for region II) by electroporation (250 V, 850 μF). The vector containing both GFP and RFP, pMS28, served as the control. The numbers of GFP- and RFP-positive cells were analyzed by flow cytometry in each case. The percentage of GFP expression was normalized to the RFP expression in each case, and the data were analyzed using GraphPad Prism and plotted as a bar diagram with the standard error of the mean. The experiments were repeated a minimum of 3 independent times in every case.

RESULTS

The *HOX11* breakpoint region possesses two G-quadruplex-forming motifs. Previous studies suggested that *HOX11* could not act as a cryptic RSS for V(D)J recombination (12, 13). Thus, other possibilities for its fragility were considered. Analysis using bioinformatic tools such as the Non-B DB database and QuadBase (26, 45) indicated that an approximately 500-bp region in *HOX11*, comprising the patient breakpoints, had sequences with the propensity to form two G-quadruplexes flanking the major patient breakpoint cluster (see Fig. S1 in the supplemental material). The sequence analysis showed two segments of DNA containing 4 stretches of 3 or more guanines (Fig. 1A; see also Fig. S1 in the supplemental material). The first motif was ~20 bp in length and upstream of the patient breakpoint cluster (referred to as region I), while the latter was about 70 bp downstream of the breakpoint cluster region (referred to as region II). A distance of ~150 bp separated both the G-quadruplex-forming regions (see Fig. S1). Interestingly, besides the two potential G-quadruplex-forming regions, we could not find any other such motifs for at least a length of 2 kb on either side of the *HOX11* breakpoint region (see Table S2 in the supplemental material).

The G-rich regions flanking the *HOX11* breakpoints can fold into two independent G-quadruplexes. In order to test whether the two proposed G-quadruplex-forming motifs could fold into non-B DNA structures *in vitro*, oligomeric DNA spanning the regions were designed (Fig. 1A). Equal-length oligomers containing the complementary C-rich sequences were used as controls. [γ-P^{32}]ATP-labeled oligomeric DNA was incubated in the presence of KCl and resolved on a native PAGE gel. Results showed the presence of multiple differentially migrating bands, due to intermolecular G-quadruplexes, in the case of G-rich strands for both the regions (Fig. 1B, lanes 3, 4, 7, 8). However, intramolecular G-quadruplex formation was absent when no KCl was present in the running buffer, and the G-rich strand migrated along with the complementary C-rich strand (Fig. 1B). Interestingly, the addition of KCl in the running buffer facilitated faster migration of a portion of the G-rich strand than of the C-rich strand, suggesting the formation of intramolecular G-quadruplexes (Fig. 1C). In both regions, the presence of KCl in the running buffer had no major effect on the formation of the intermolecular G-quadruplexes (Fig. 1B and C).

Circular dichroism spectroscopy is used for studying the structural conformations of DNA. In both the regions, we observed the characteristic peak at 260 nm and a dip at 240 nm for the wild-type G-rich sequences (Fig. 1D and E), indicating formation of parallel G-quadruplexes. CD spectra following heat denaturation of the G-rich sequence in the presence of KCl at 90°C led to a total collapse of the peak at 260 nm, indicating destruction of the G-quadruplex structure in the case of region I (Fig. 1D). However, in the case of region II, the quadruplex could not be totally destroyed upon heating at 90°C (Fig. 1E). This suggests that G-quadruplexes formed in certain cases can be highly stable in the presence of K^+. Further, CD analysis of the oligomers containing the renatured DNA in KCl following denaturation showed recreation of the respective G-quadruplex structures (Fig. 1D and E).

The formation of both intra- and intermolecular G-quadruplexes was dependent on DNA concentration, with the intensity of each of the products increasing with that of the radiolabeled oligomer (Fig. 2A to C). Time kinetics studies showed that in both the regions, formation of intramolecular G-quadruplexes increased at the initial time points; however, it got saturated immediately. Surprisingly, in the case of the intermolecular G-quadruplexes, the product formation was independent of time and remained at the same level at all the time points studied (Fig. 2D and E).

In order to study the role of stretches of guanines in the potential G-quadruplex-forming regions, mutant oligomers were designed by modifying one or more stretches of guanines (Fig. 3A and D). Upon performing native gel mobility assay, we found that only the wild-type sequences could fold into intramolecular G-quadruplexes, and none of the mutants could do so in the case of region I in the presence of KCl (Fig. 3B and C). Interestingly, the efficiency of formation of the intermolecular species also varied, depending on the extent of the mutation, with only the wild type being able to form a large population, while the mutant oligomers, depending on the number of G stretches mutated, could form G-quadruplexes only up to a limited extent, which was independent of KCl (Fig. 3B and C). Interestingly, despite completely abolishing intramolecular G-quadruplex species, the G_3 mutant could still fold into multiple forms of intermolecular species. Hence, the guanine dependency of the non-B DNA structure at the *HOX11* region I confirms the formation of a G-quadruplex. Comparable results were observed when mutations were introduced in *HOX11* region II (Fig. 3D to F). Thus, these results confirm the requirement of stretches of guanines in the formation of altered DNA structures flanking the *HOX11* breakpoint region.

It has been reported that G-quadruplexes can be formed even by using GNG motifs with bulges in the structures (46). Since the stretches of guanines in *HOX11* region I have GNG motifs both upstream and downstream to the core region, we tested mutants wherein either two of the GNG motifs were mutated individually or together with the G_3 mutations (see Fig. S2A in the supplemental material). Our results showed an absence of bimolecular G-quadruplexes upon GNG mutations to the G_3 oligomer, indicating their possible involvement in the formation of such intermolecular species, at least to a lesser extent (see Fig. S2). Apart from this, there was no other significant change in the formation of higher-order intermolecular quadruplexes. Since the region upstream to core G-stretches contains two GNG motifs, a shorter oligomer containing only one GNG motif was designed and used for the study (see Fig. S3A and C in the supplemental material). Results showed a consistent decrease in the formation of bimolecular G-quadruplexes upon GNG mutation (see Fig. S3). However, mutation of either the upstream or downstream GNG motif alone resulted in the formation of certain new bimolecular species, which were not formed when both motifs were mutated, indicating the importance of GNG motifs in the generation of intermolecular G-quadruplexes at this region. However, more studies are needed to reveal the nature of base pairing in these intermolecular species.

CD analysis was also performed on the G-rich oligomeric DNA containing *HOX11* region I or II, along with their respective complementary strands, random and mutant oligomers (Fig. 3G and H). Unlike the wild-type G-rich sequences, the spectra of the mutant (G_{10} and G_{14} for region I and G3A, G3B, and G9 for region II), random, and complementary strand (C) oligomers resembled that for single-stranded DNA, showing a peak at ~280 nm and a trough at 260 nm (Fig. 3G and H). This further confirmed the hypothesis that both regions I and II could fold into two independent parallel G-quadruplex structures.

The G-rich region upstream of *HOX11* breakpoints blocks replication of the DNA. Earlier studies have suggested that the presence of non-B DNA structures may prevent the progression of DNA polymerases during replication of DNA, resulting in pause sites (47–49). In order to test whether the non-B DNA structure formed at the *HOX11* breakpoint region could also block the DNA polymerase progression, we performed polymerase arrest assay on both single- and double-stranded DNA. Upon using oligomeric DNA containing wild-type sequences of regions I and II, we found two major pause sites on the DNA corresponding to both the regions (Fig. 4A and B, lane 2). Upon partial mutation of the G-quadruplex-forming motif, the pause was still observed, although the intensity was reduced (Fig. 4A and B, lane 3). This could be explained by the fact that partial mutations could disrupt the formation of intramolecular and bimolecular structures; however, the tetramolecular species could still be formed. This was further supported by the disappearance of the pause sites with concomitant reappearance of the full-length extension products, upon complete mutation of the stretches of guanines, in both regions I and II (Fig. 4A and B, lane 4). A similar assay using a random sequence with similar GC content showed no polymerase

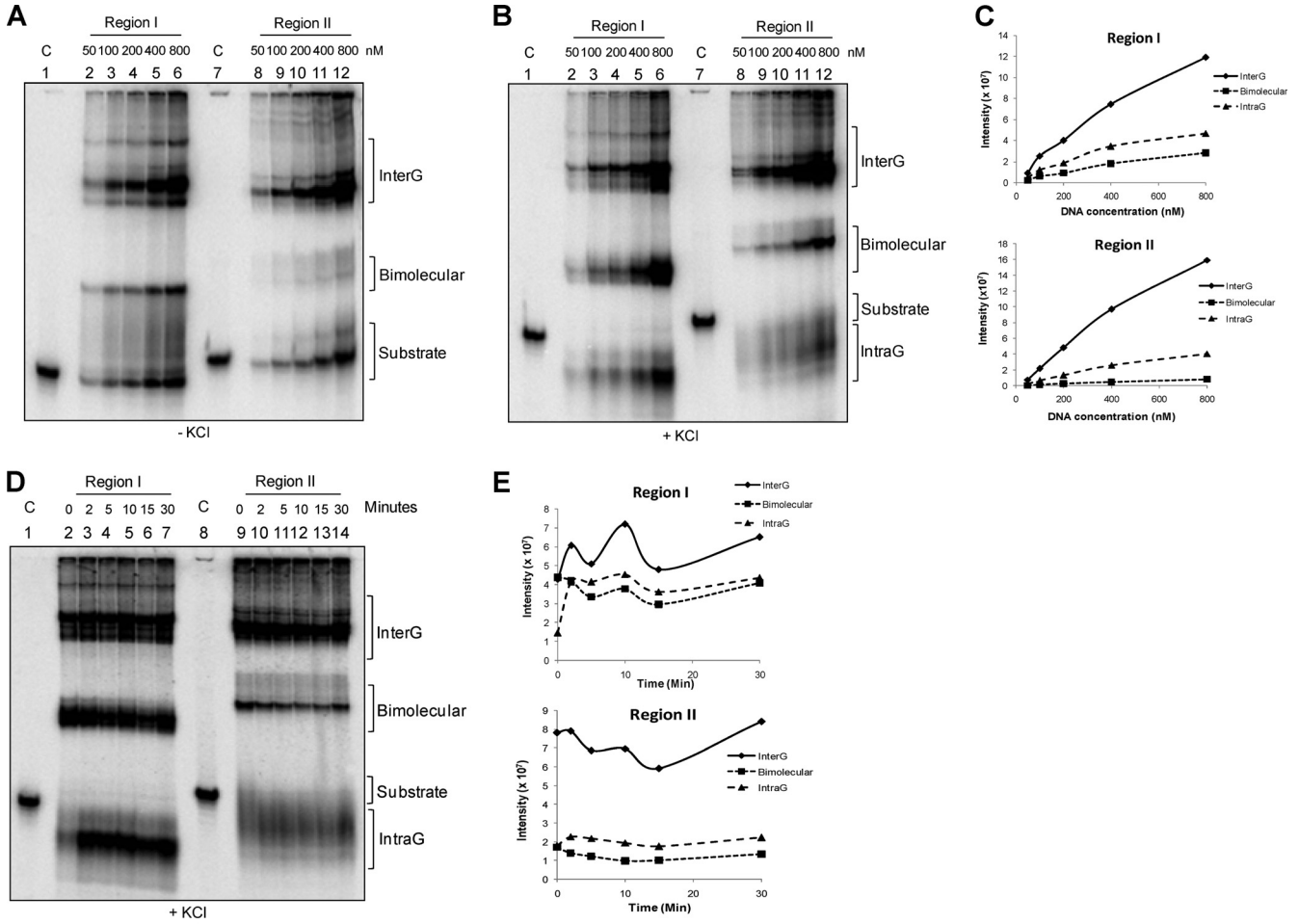

FIG 2 Concentration dependence and time kinetics of G-quadruplexes formed on the template strand of the *HOX11* breakpoint region. (A and B) Increasing concentrations of G-rich strands of regions I and II were incubated in 100 mM KCl for 1 h at 37°C and resolved in the absence (A) or presence (B) of KCl (25 mM) in the gel and running buffer. The substrate, intramolecular (Intra G), and intermolecular (Inter G) quadruplex structures are indicated. (C) Quantification of the rate of formation of G-quadruplexes with increasing concentrations of DNA, shown in panels A and B. (D) G-rich strands from both regions I and II were incubated in the presence of 100 mM KCl at 37°C for increasing time periods, as indicated. After incubation, the substrates were immediately loaded and resolved on a native PAGE gel containing KCl. (E) Line curve representing the rate of formation of intermolecular, intramolecular, and bimolecular G-quadruplexes after 0, 2, 5, 10, 15, and 30 min of incubation. "Intensity" is measured and represented as photostimulated luminescence units (PSLU).

arrest, reiterating the importance of the stretches of guanines in the structure formation and hence stalling of replication (Fig. 4A and B, lane 5). We also performed the *Taq* polymerase arrest using the G-quadruplex-forming motif from the *HIF1*α promoter (50). To further confirm whether the observed polymerase stop was indeed due to formation of a G-quadruplex, the assay was also performed in the presence of increasing concentrations of LiCl. Results showed that a quadruplex at region I still formed, to a limited extent, whereas that of region II was sensitive to even low concentrations of Li ions (Fig. 4C, lanes 2 to 5 and 7 to 10).

DMS protection assay is used for deciphering the nature of base pairing involving guanines during the formation of G-quadruplexes. When guanines are involved in Hoogsteen base pairing (during G-quadruplex formation), the N7 position is not free for methylation and hence does not react with DMS. Since the G-quadruplexes formed at *HOX11* were of both intra- and intermolecular type, we gel purified the different quadruplex structures and substrate. Upon treatment with DMS, the substrate (without any structure) showed modification of all the guanines and further cleavage by piperidine for both regions I and II (Fig. 4D and E, lane 2). However, in the case of the purified intramolecular as well as the intermolecular G-quadruplexes, all the guanines showed protection from modification by DMS, suggesting that all of these were involved in the structure formation (Fig. 4D and E, lanes 2 to 4). In the case of the intramolecular G-quadruplex for region II, protection was limited compared to that for region I (Fig. 4D and E, lanes 2 and 3). This is supported by the gel mobility shift assay, wherein a strong compact band for the intramolecular G-quadruplex could not be seen, as for region I. Interestingly, in the case of the bimolecular structures, there was a clear protection of only two stretches of guanines (Fig. 4D and E, lanes 5 and 6). Therefore, the DMS modification assay confirms that the non-B DNA structures formed by the template strand at *HOX11* are indeed G-quadruplexes.

In order to test whether the G-quadruplex structure can be formed in the context of double-stranded supercoiled DNA, a

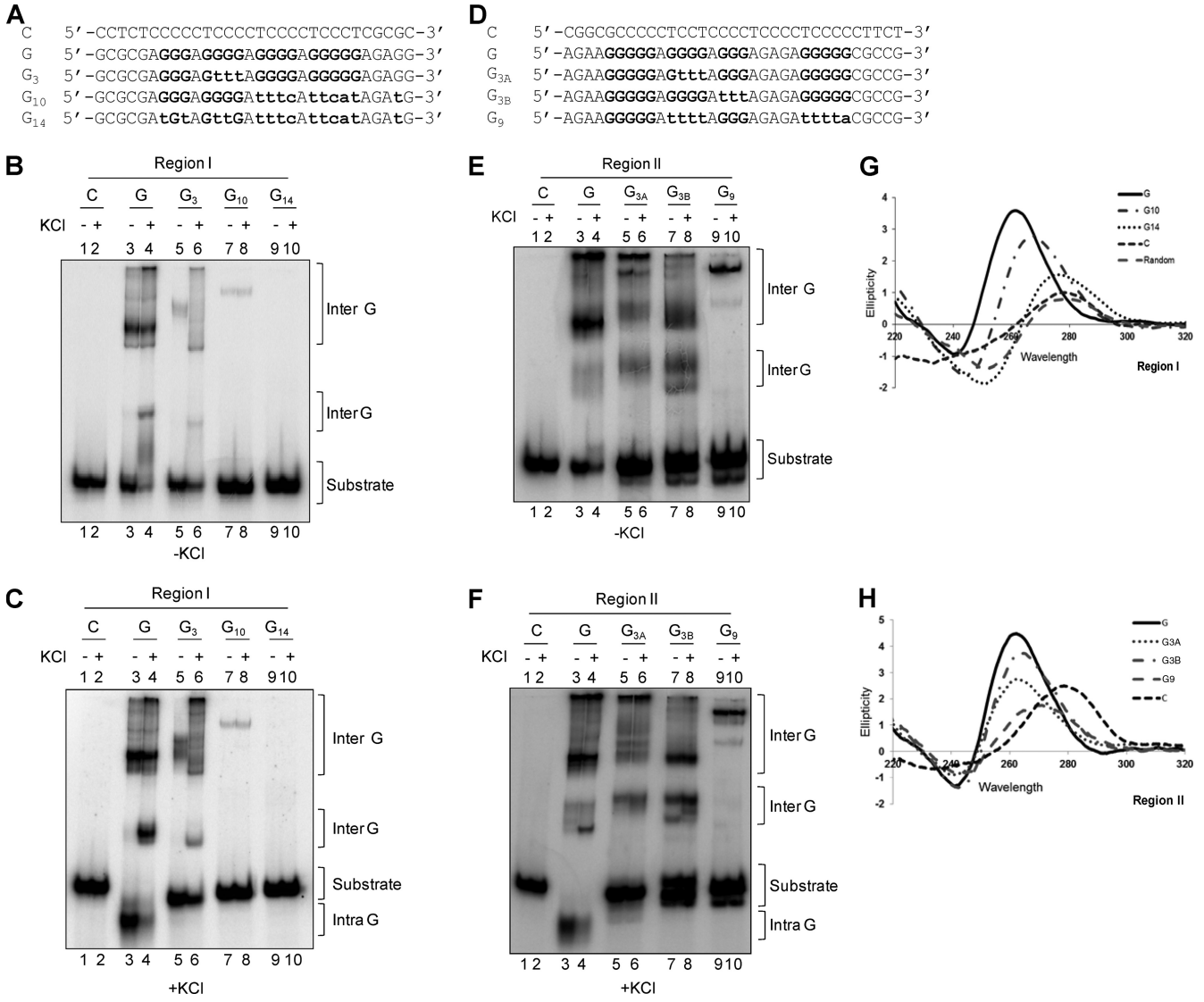

FIG 3 Mutations in the G-quadruplex-forming sequences at the *HOX11* breakpoint region abolish structure formation. (A) Mutations in the G-strand of region I of the *HOX11* breakpoint were generated such that 3 (G_3), 10 (G_{10}), and 14 (G_{14}) guanines were altered to other bases (in bold and lowercase). (B and C) Different oligomeric DNA was incubated either in the presence or absence of KCl and resolved on a 15% native PAGE gel in the absence (B) or presence (C) of KCl in the electrophoresis buffer. (D) Mutations in the G-rich strand of region II were generated such that 3 (G_{3A} and G_{3B}) and 9 (G_9) guanines were altered to other bases. (E and F) Oligomeric DNA was resolved in the absence (E) or presence (F) of KCl on a native PAGE gel. (G and H) CD spectra showing G-quadruplex formation in region I (G) and region II (H) of the *HOX11* breakpoint region. Wild-type (G), 10-nt (G_{10}), and 14-nt (G_{14}) mutant G-rich strands, complementary C-rich (C) strand, and a random sequence oligomer were incubated with 100 mM KCl and TE at 37°C, and the CD spectra were recorded in the case of region I. In the case of region II, besides the wild type (G), two 3-nt (G3A and G3B) mutant strands, a 9-nt (G9) mutant strand, and a complementary C-rich (C) strand were used for the analysis.

486-bp region containing the *HOX11* breakpoint region was cloned into a plasmid, pSKS1. To perform the primer extension on pSKS1, multiple radiolabeled primers positioned at different lengths from the *HOX11* G-rich region I were used (Fig. 5A). Results showed the presence of a very prominent pause site, whose position shifted in tandem with the position of the primers (Fig. 5B, lanes 1 to 8, and C, lanes 1 to 12). The band corresponding to the pause site matched G-rich region I, indicating formation of the G-quadruplex structure (Fig. 5B, lanes 1 to 8; see also Fig. S1, region I, marked using a dotted box, in the supplemental material). Importantly, upon primer extension through the complementary C-rich strand, we could not observe any truncated products resulting due to pause sites (Fig. 5B, lanes 9 to 12, and C, lanes 13 to 16). Thus, our results show that the G-quadruplex structure is formed at *HOX11* region I, even on a double-stranded plasmid DNA.

The replication block at *HOX11* breakpoint G-rich region I is dependent on K^+. In order to confirm that the observed pause sites at region I of *HOX11* were indeed due to a G-quadruplex, we tested the sensitivity of these structures to different cations (Fig. 5D). Results showed that the presence of the pause sites was dependent on KCl (Fig. 5D, lanes 1 to 4). Consistently, in the absence of K^+ ions or in the presence of Li^+ or Na^+, intensity of the pause was reduced dramatically (Fig. 5D, lanes 5 to 10). This suggests

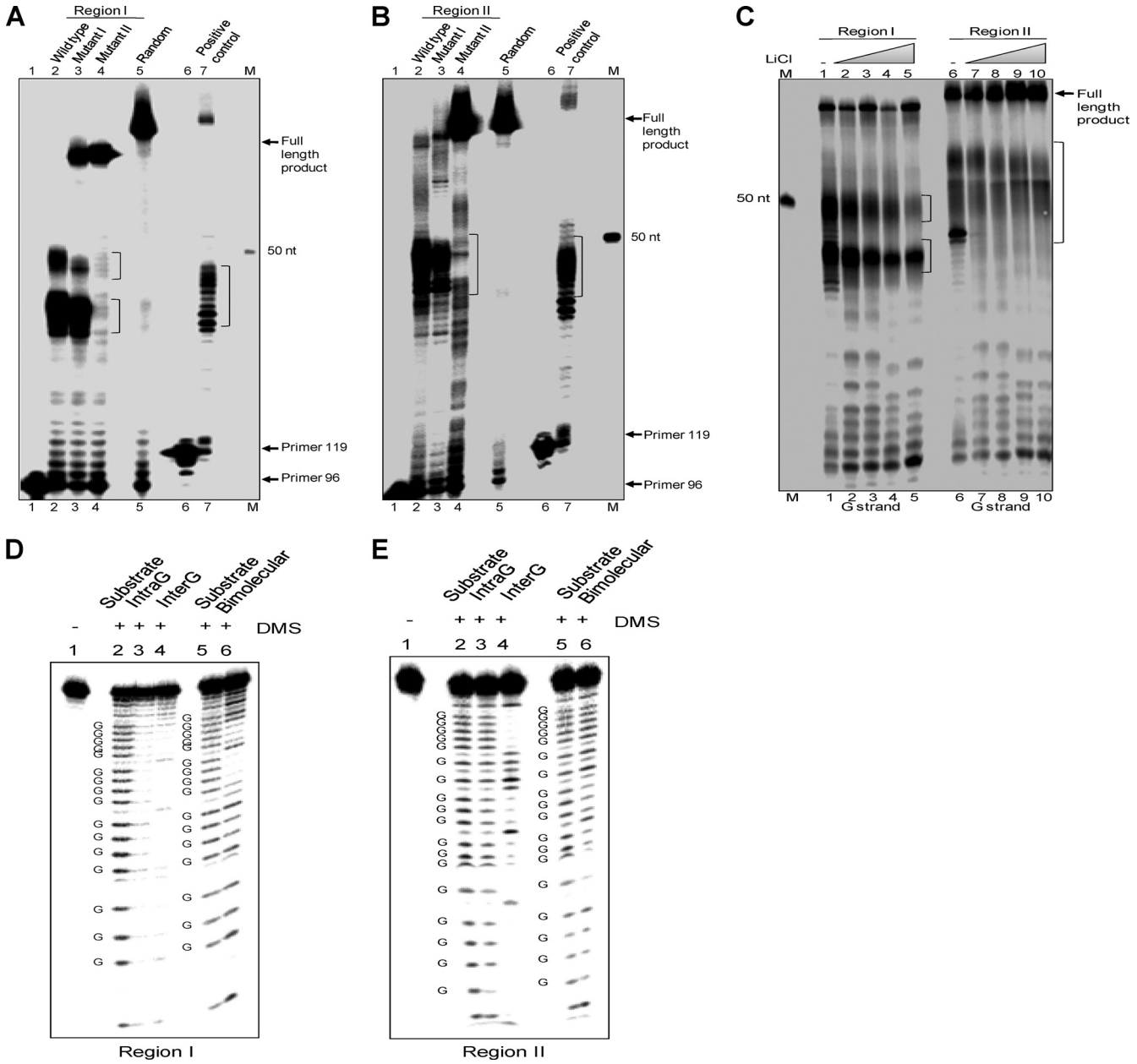

FIG 4 *Taq* polymerase arrest and DMS protection assay for the G-quadruplex-forming region at the *HOX11* breakpoint region. Oligomers spanning region I and II were annealed with respective radiolabeled primers in the presence of 50 mM KCl, and extension reaction was performed using *Taq* polymerase. (A) The gel profile following primer extension on wild-type, mutant I (two stretches, 8 Gs), and mutant II (four stretches, 17 Gs) template DNA. (B) The gel profile after extension for the wild type, mutant I (two stretches, 7 Gs), and mutant II (four stretches, 14 Gs) of *HOX11* region II. In both panels A and B, full-length extension is represented by the arrow, and pause sites are indicated with brackets. A random oligomer with a GC content and sequence similar to those of the *HIF1*α promoter was used as negative and positive controls, respectively. Primer MN96 was used for *HOX11* and random sequence, whereas MN119 was used for *HIF1*α region extension. (C) Evaluation of the effect of increasing concentrations of LiCl on primer extension through region I and region II of *HOX11*. (D and E) Intra G, Inter G, and bimolecular forms of G-quadruplexes were purified from the gel and were treated with DMS, followed by cleavage with piperidine, and were resolved on a 15% denaturing PAGE gel. Substrate indicates the gel-purified DNA from the reaction incubated without KCl. For both regions I and II, the positions of guanines are indicated.

that the altered DNA structure at *HOX11* breakpoint region I is a G-quadruplex, since it was stabilized only in the presence of K^+. In the presence of Ca^{2+} (10 mM), the primer extension itself was affected, and therefore a conclusion could not be drawn (Fig. 5D, lanes 11 and 12). When primer extension was performed using increasing concentrations of KCl, an elevated intensity of the band due to polymerase arrest was observed in a concentration-dependent manner (Fig. 5E, lanes 9 to 16). Since Li^+ does not support the formation of such structures, primer extension reactions were performed in the presence of increasing concentrations of LiCl. Results showed no or limited polymerase arrest in the presence of Li^+ (Fig. 5E, lanes 3 to 8). Upon performing primer extension in

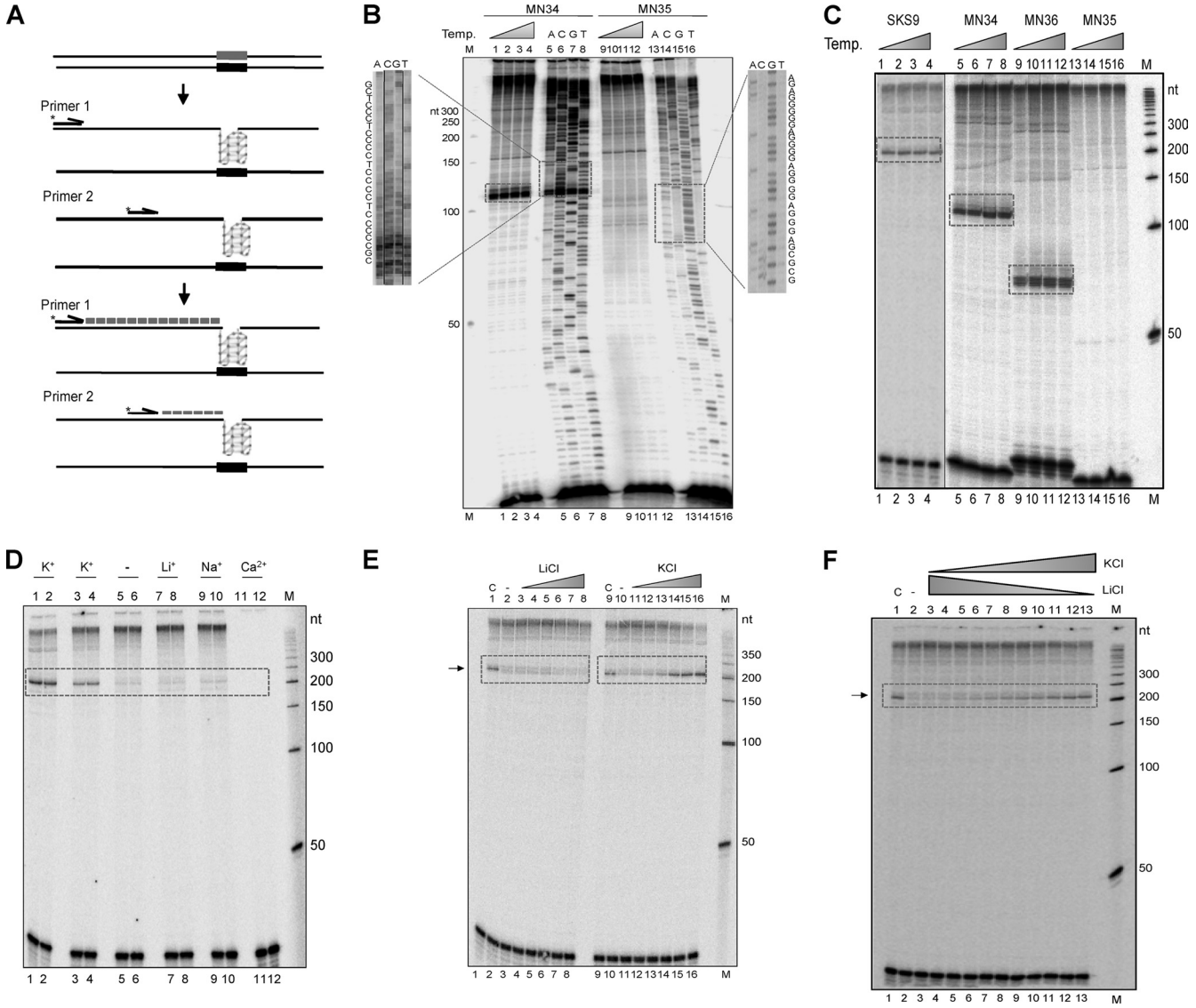

FIG 5 Evaluation of G-quadruplexes induced pause sites on region I of the *HOX11* breakpoint in a plasmid DNA. (A) Schematic diagram showing the primer extension across a template DNA containing a G-quadruplex. A primer binds to one of the strands of the plasmid DNA upon heat denaturation. During extension, the primer can extend only till a non-B DNA structure like a G-quadruplex, as it blocks the progression of the polymerase. The truncated products (dotted lines) are then resolved on a denaturing PAGE gel. (B) The pSKS1 plasmid, containing the *HOX11* breakpoint region, was used for primer extension reactions for region I using radiolabeled primers for the G-rich (MN34) and C-rich (MN35) strands using a temperature gradient (50, 53, 56, and 59°C). The extension reactions were done for 25 cycles in a buffer containing KCl. Sequencing ladder was prepared for region I, using MN34 and MN35 by the chain termination method of sequencing. A, C, G, and T denote the corresponding ddNTP-mediated chain termination reaction. The products were resolved on 8% denaturing PAGE gels, and signals were detected. Pause sites are indicated with dotted rectangles. Sequences corresponding to the pause site and G-quadruplex-forming motifs are indicated. Lines are used to indicate separation of merged lanes. "M" denotes the 50-nt ladder. (C) Primer extension on region I of pSKS1 using primers (SKS9, MN34, and MN36) that bind at different positions with respect to the G-quadruplex-forming motif. Primer extension was performed on the C-rich strand using MN35. The temperature gradient used was 52, 56, 58, and 61°C. (D) Evaluation of effect of different cations (K^+, Li^+, Na^+, and Ca^{2+} ions) on G-quadruplex-mediated primer extension pause sites on pSKS1. Lanes 1 and 2 contain reactions performed in the presence of the standard thermo polymerase buffer, while reactions in lanes 3 and 4 were performed in reconstituted buffer supplemented with 10 mM KCl. Lanes 5 and 6 contain no ions. (E) Effect of increasing concentrations of Li^+ and K^+ (0, 1, 2, 5, 10, 20, and 50 mM) on G-quadruplex formation in pSKS1. (F) The primer extension reactions were performed in increasing concentrations of KCl and decreasing concentrations of LiCl in the same reaction. The range of concentrations used is 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10 mM for both Li^+ and K^+. "C" denotes the primer extension reaction in the presence of buffer containing KCl, while the rest are in the presence of a buffer without KCl but supplemented with the appropriate concentrations of the same. In panels D, E, and F, SKS9 was used for primer extension. In all panels, "M" denotes the radiolabeled 50-nt ladder.

the presence of both Li^+ and K^+ in the same reaction mix, we found that with an increase in K^+ and decrease in Li^+, there was a concentration-dependent enhancement in the occurrence of the pause site (Fig. 5F, lanes 3 to 13).

The G-rich region downstream of the *HOX11* breakpoints can fold into a G-quadruplex. The ability of G-rich region II to block the DNA replication was tested by primer extension on pSKS1 using primers positioned upstream of the potential G-qua-

druplex-forming sequence and downstream of region I. Results showed the occurrence of two prominent bands corresponding to the second G-quadruplex-forming region at *HOX11* (Fig. 6A). Further, the position of pause sites moved according to the primer binding sites, when different primers were used (Fig. 6B and C). Interestingly, primer extension across the complementary C-rich strand in region II resulted in a pause, unlike region I (Fig. 6A, marked by an asterisk). Further, the intensity of the pause sites in the G-rich strand increased in the presence of K^+, while it was minimal in the absence of K^+, and the addition of Li^+ also did not support structure formation (see Fig. S4A and B in the supplemental material). Besides, in the presence of both K^+ and Li^+, a pattern similar to that of region I was observed (see Fig. S4C). Overall, these results provide evidence for the formation of two independent G-quadruplex structures on a plasmid DNA at the *HOX11* breakpoint region.

Mutations to the stretches of guanines flanking the *HOX11* breakpoint region can prevent polymerase arrest on a plasmid DNA. Since formation of G-quadruplex structures flanking the *HOX11* breakpoint region could block the progression of polymerases, we wondered whether mutations to the guanines at regions I and II could restore the passage of polymerase. In order to generate mutations, site-directed mutagenesis was performed using sets of primers having the desired nucleotide modifications, as described in Materials and Methods. A plasmid, pMN18, containing mutations both in regions I and II and plasmids with mutations either in region I (pMN20) or region II (pMN21) were generated (see Fig. S5A and B in the supplemental material). The effect of mutations on the polymerase arrest was studied by performing primer extension on the wild-type and mutant plasmids in the presence of K^+. Since it is difficult to resolve the pause sites resulting from both regions I and II simultaneously in the same gel using a single primer, two independent primers, MN34 and MN91, were used for regions I and II, respectively. Results showed that the mutations to either one or both of the G-rich motifs led to a remarkable abrogation of the pause sites due to polymerase arrest in these regions (Fig. 6D, lanes 2 to 5 and 12 to 15; see also Fig. S5C and D in the supplemental material). Interestingly, the pause observed for the complementary C-rich strand in region II disappeared upon double mutation and mutation in the region I sequence but not in region II (Fig. 6D, lanes 17 to 20; see also Fig. S5). This indicates that both the regions can form two independent non-B DNA structures individually and simultaneously, since mutation at one motif does not hinder the formation of the other structure. Overall, mutations at the two G-rich sequences present in the *HOX11* breakpoint region can prevent the formation of the two independent G-quadruplexes, demonstrating an important role of these motifs in imparting fragility to this region.

Single strandedness exists at the G-rich regions, near the *HOX11* breakpoints, on a plasmid DNA. G-quadruplex structure formation on a double-stranded DNA would result in a certain amount of single strandedness on both the strands of DNA, particularly the C-rich strand. In order to test this, S1 nuclease assay was performed on the pSKS1 plasmid, which contains both regions I and II of *HOX11*. Results showed a significant increase in the intensity of bands upon the addition of increasing concentrations of S1 nuclease (0, 10, 20, 40, and 80 U), and cleavage positions were at close proximity to the previously observed pause sites due to G-quadruplexes in both the regions (Fig. 7A and C, lanes 2 to 5). Interestingly, we could also observe S1-specific nicks, corresponding to region II, during primer extension across region I (Fig. 7A, lanes 2 to 5, marked by asterisks). More importantly, multiple weak S1-specific bands were observed in the region complementary to the G-quadruplex-forming motifs (Fig. 7B and D, lanes 2 to 5). It was noted that the number of bands in the complementary region increased upon titration of S1 nuclease, as expected, and at the highest concentration, the bands decreased due to excessive S1-specific cleavage. The limited S1 sensitivity on the complementary strand may be due to the involvement of additional bonds by cytosines, such as i-motifs (see below). In addition, there was no S1 sensitivity observed when the double-mutant plasmid, pMN18, was used in either of the regions (Fig. 7E). These results confirm formation of G-quadruplexes at regions I and II flanking the *HOX11* breakpoint cluster region, imparting a certain amount of single-stranded character to the G-rich and their complementary strands.

The complementary strand of the *HOX11* G-quadruplexes may form i-motifs. Limited S1 sensitivity on a cytosine-rich strand of the *HOX11* breakpoint region containing plasmid DNA indicated the possibility of formation of cytosine-specific additional bonds, such as those found in i-motif structures. To test this, we used CD spectroscopy to determine the ability of the two C-rich strands at the *HOX11* breakpoint region to fold into i-motifs. Since formation of such structures generally requires protonation of cytosines, we incubated the complementary strands from regions I and II in buffers with differing pH (4.8, 5.2, and 6.8). Interestingly, we found that both these regions exhibited the spectra characteristic of i-motifs, with absorption maxima of ~285 nm and minima of 240 nm (see Fig. S6A and D in the supplemental material). Importantly, sequences with mutations at the cytosines or random nucleotides of identical lengths, under similar conditions, did not show such a spectra, showing the possible relevance of cytosine stretches in the formation of potential i-motifs (see S6B, C, E and F).

G-quadruplexes at the *HOX11* breakpoint region affect transcription within the cells. Further, we wanted to examine the ability of the *HOX11* breakpoint region to form G-quadruplex structures *in vivo*. For this, an extrachromosomal assay system was used following cloning of the *HOX11* breakpoint regions I and II between promoter and coding regions of the EGFP gene (pMS29) (Fig. 8A). Expression of RFP was used as a measure of transfection efficiency. Appropriate mutants of regions I and II were also generated (Fig. 8B). The constructs were transfected into REH cells, and GFP/RFP expression was scored by flow cytometry. The results showed a significant decrease in the GFP expression in the cells transfected with the episome containing the wild-type breakpoint region (Fig. 8C). However, when either of the regions was mutated, a significant increase was observed in the GFP expression (Fig. 8C). Interestingly, GFP expression was comparable to that of the vector control when both the regions were mutated, suggesting that the structures formed at *HOX11* could block transcription *in vivo* (Fig. 8C). Taken together, these results suggest that both the regions can fold into G-quadruplexes independent of each other, since single mutants express higher GFP than the wild type. However, neither of the single mutants expressed GFP as high as the vector control or the double mutants, suggesting a possible interaction between the two structures. These results suggest that G-quadruplexes at the *HOX11* breakpoint region can be formed within mammalian cells and could have an effect on physiological processes *in vivo*.

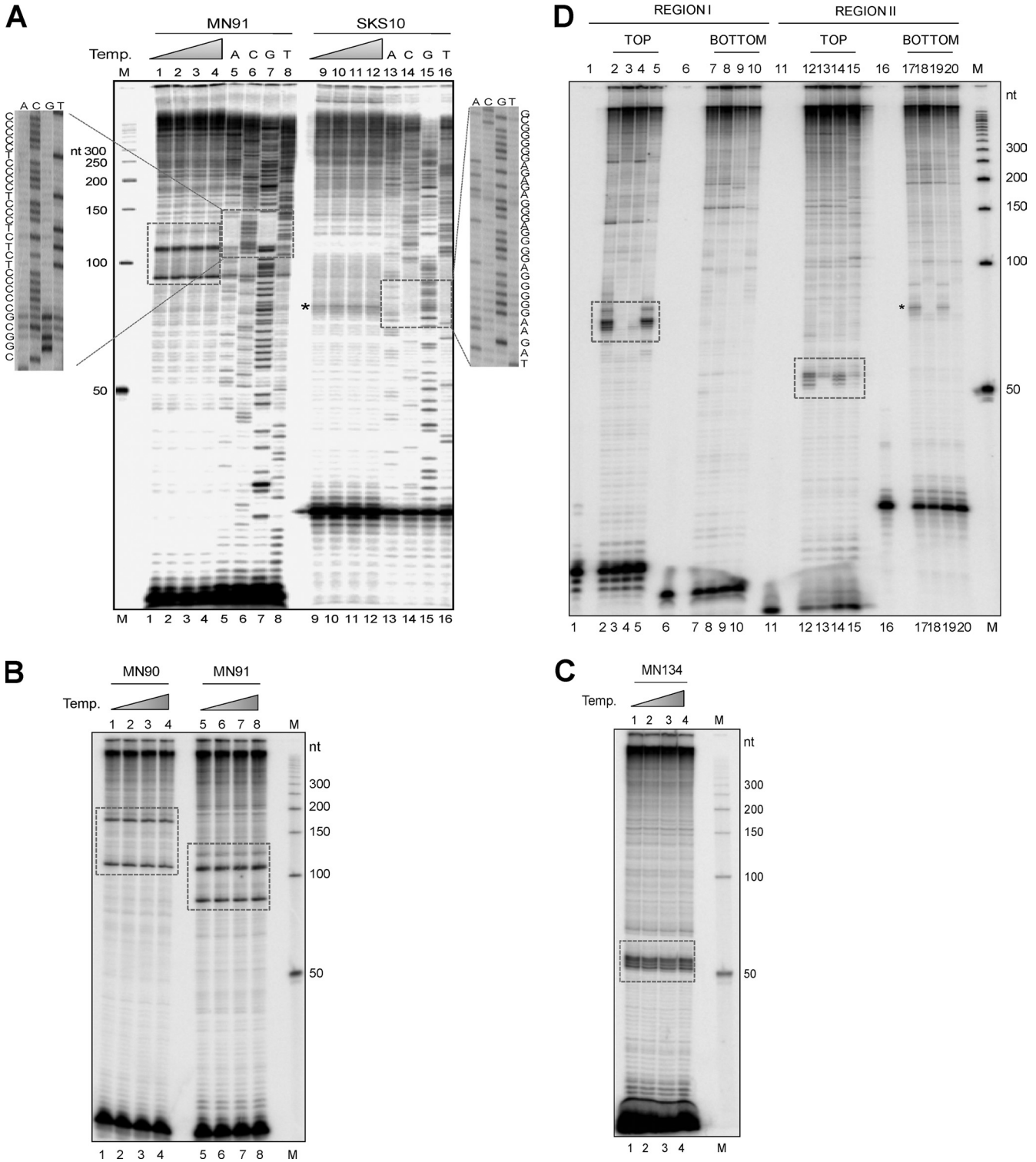

FIG 6 Evaluation of the effect of mutations on G-quadruplex-forming motifs in the *HOX11* breakpoint region. (A) Primers, MN91 for the G-rich strand and SKS10 for the C-rich strand, were used for primer extension across the second G-rich region in *HOX11* in pSKS1 at different temperatures (48, 50, 52, 54°C for MN91 and 56, 59, 60, 62°C for SKS10). Sequencing ladders were prepared using the appropriate primers (MN91 and SKS10) as described. Regions corresponding to pause sites are highlighted with the sequence of the region based on the ladder. (B and C) Primers MN90 and MN91 (at 48, 50, 52, and 54°C) and MN134 (at 60, 62, 64, and 66°C) were used for primer extension across the G-motif of region II in pSKS1. (D) Evaluation of effect of mutations at regions I and II on observed pause sites. Primer extension across pSKS1 (wild type; lanes 2, 7, 12, 17), pMN18 (double mutant; lanes 3, 8, 13, 18), pMN20 (mutant region I and wild-type region II; lanes 4, 9, 14, 19), and pMN21 (wild-type region I and mutant region II; lanes 5, 10, 15, 20) containing the *HOX11* breakpoint region on plasmid DNA using specific radiolabeled primers. Extension was performed for both the regions as well as the top and bottom strands. In all panels, the pause sites in the G-rich strands are marked in dotted boxes, while the pause at the C-rich strand is marked by an asterisk. "M" denotes a 50-nt ladder.

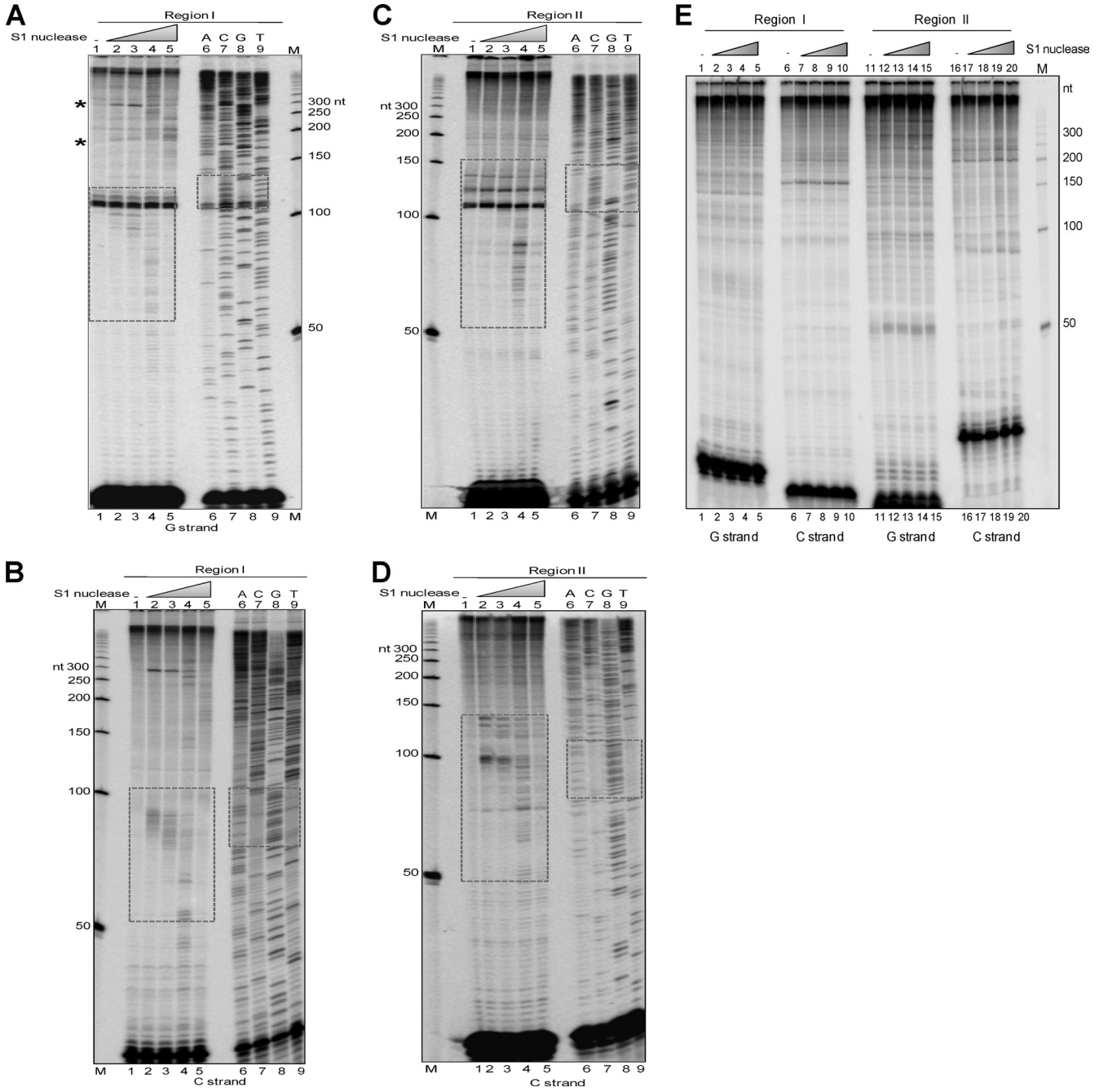

FIG 7 S1 nuclease sensitivity pattern at the wild-type and mutant *HOX11* breakpoint region, when present on a plasmid DNA. The plasmid, pSKS3 (wild type), was heat denatured and reannealed in the presence of 100 mM KCl overnight. The plasmids were then treated with increasing concentrations of S1 nuclease (0, 10, 20, 40, and 80 U). The DNA was then subjected to primer extension using radiolabeled primers specific for region I, G-rich strand (MN34) (A) and C-rich strand (MN35) (B), and for region II, G-rich strand (MN91) (C) and C-rich strand (SKS10) (D), of the *HOX11* breakpoint region. "−" represents 0 U, and the S1 nuclease-sensitive regions are marked by dotted boxes. Two specific S1-sensitive bands are marked by asterisks. "M" is a 50-nt ladder. For the sequencing ladders, the regions corresponding to the S1 nuclease sensitivity are marked by dotted boxes. A, C, G, and T denote the corresponding ddNTP chain termination reaction. (E) The plasmid harboring mutations in both the regions, pMN18, was treated with increasing concentrations of S1 nuclease, 0, 10, 20, 40, and 80 U, followed by primer extension using radiolabeled primers specific for both regions in the *HOX11* breakpoint region.

Genomic loci containing the *HOX11* breakpoint region possess single strandedness. Since we could observe single strandedness in the regions flanking the *HOX11* breakpoints on a plasmid, we wondered whether such a single-stranded character could be detected in the genome. In order to test this, genomic DNA was isolated from the CEM and K562 mammalian cell lines under nondenaturing conditions and subjected to sodium bisulfite modification. The *HOX11* breakpoint region was then PCR amplified, cloned, and sequenced. If an unmethylated cytosine is present in the single-stranded DNA, it can react with sodium

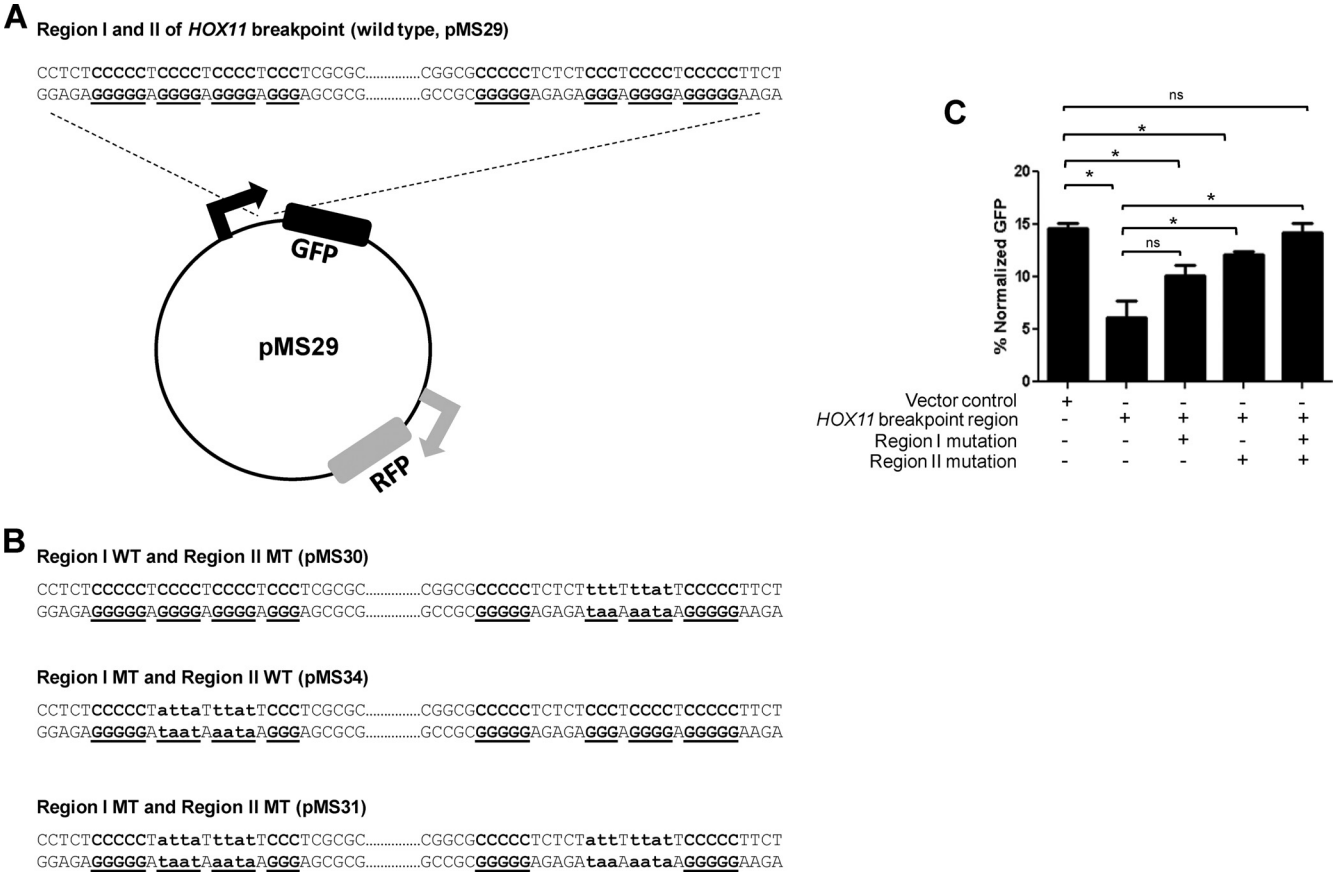

FIG 8 Evaluation of the effect of G-quadruplexes on expression of GFP using an extrachromosomal episome containing the *HOX11* breakpoint region within the cells. (A) Schematic presentation of episome; pMS29 containing region I and II of *HOX11* breakpoints (wild type) cloned between promoter and coding sequence of EGFP. pMS29 also harbors an RFP expression cassette. (B) The nucleotide sequences of *HOX11* breakpoint regions containing mutations in G-quadruplex motifs, used for cloning into the episomes, are indicated. The stretches of guanines are in bold and underlined and the mutations are represented in bold and lowercase. (C) Comparison of GFP expression in the presence of wild type and mutants of *HOX11* breakpoint region. REH cells were transfected with the pMS28 (vector control), pMS29, pMS30, pMS31, or pMS34. After 48 h of transfection, GFP and RFP expression was analyzed by flow cytometry. In each case, the percentage of GFP-positive cells for wild type and mutants was normalized against their respective RFP expression and plotted as a bar graph. The bar graph shows means ± standard errors of the means (SEM). ns, not significant; *, $P < 0.05$.

bisulfite, leading to deamination and formation of uracil. Such a cytosine-to-uracil change can be read as a thymine conversion upon PCR amplification followed by DNA sequencing. We sequenced 15 clones from K562 genomic DNA and 30 clones from CEM cells. Out of the total 45 DNA molecules which were sequenced, 37 molecules exhibited C-to-T conversions to different extents. Interestingly, ~10% of molecules showed conversion of all the cytosines opposite the guanines involved in quadruplex formation. The cumulative frequency of C-to-T conversion suggested that only the cytosines which were present in regions I and II of the *HOX11* breakpoint region were single stranded, while the other cytosines upstream or downstream of this region were paired (Fig. 9). These results suggest that G-quadruplex structures could be formed at the G-motifs identified on either side of the *HOX11* breakpoint region within the genomic DNA as well. Besides, we noted that ~65% of molecules showed partial single-stranded character in the regions corresponding to the G-quadruplex motifs, suggesting the possibility of secondary interactions among the cytosines even in the genome. It is also important to note that among the 45 molecules sequenced, 8 did not have any conversion, suggesting that the G-quadruplex structures may not form in all the molecules in the genome but could be present in a fraction of the molecules at the *HOX11* fragile region, which supports the low frequency of the t(10;14) translocations in cancer.

DISCUSSION

The reason for fragility of the *HOX11* gene during t(10;14) chromosomal translocation in T-cell leukemia has remained elusive so far. Flanking the *HOX11* breakpoint region, we identified two G-rich motifs that can independently adopt parallel G-quadruplex structures. The formation of these structures was dependent on K^+, while they were not supported by Li^+. Mutations in even a single stretch of guanines abolished the intramolecular G-quadruplex formation but not the intermolecular forms in both the regions, indicating the critical requirement of guanines to form the structures. The importance of the sequence and K^+ for the structure formation was further reiterated on double-stranded DNA, where a major affect was seen on the formation of replication blocks. The structure formation could also block transcription inside the cells. Evidence for the structure formation on genomic DNA was provided by sodium bisulfite modification assay, where several molecules showed the presence of full or partial single

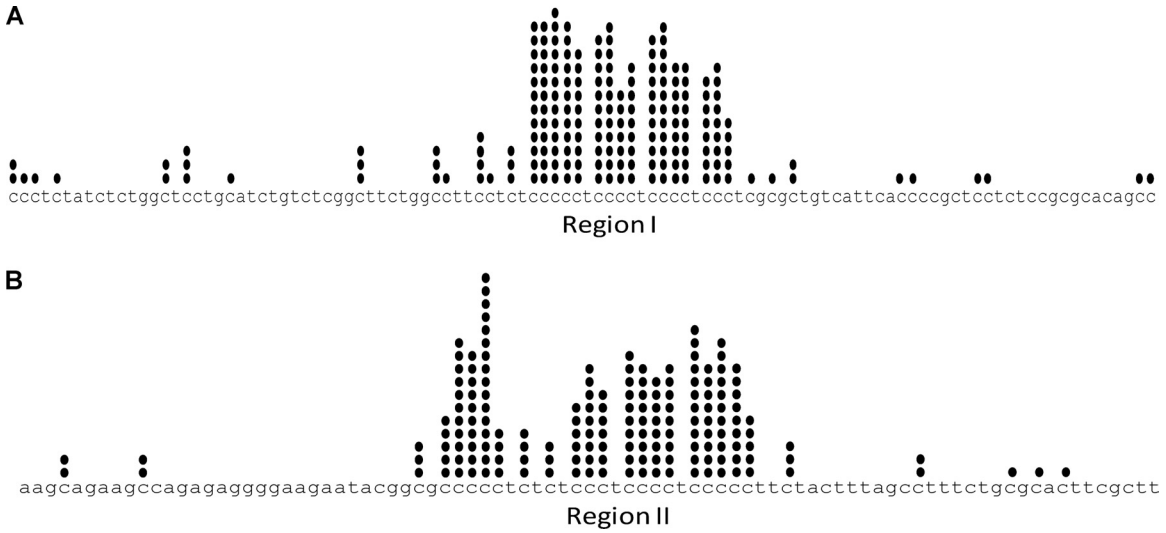

FIG 9 Sodium bisulfite modification assay on chromosomal DNA extracted from mammalian cells. The genomic DNA was extracted using nondenaturing conditions from K562 and CEM cells and subjected to sodium bisulfite modification assay. Each dark circle represents the number of times the respective cytosine in the *HOX11* breakpoint regions I (A) and II (B) is converted to thymine after deamination in the presence of sodium bisulfite, followed by PCR. A total of 45 DNA molecules were sequenced, out of which around 8 molecules were completely unreactive to sodium bisulfite treatment.

strandedness, due to the quadruplex formation on one strand. This also rules out a possible YRY triplex formation, as the pyrimidine third strand would engage with duplex DNA, which would have prevented the observed bisulfite reactivity of the complementary cytosine strand. Bioinformatics analysis using the Non-B database also did not predict any triplex DNA motif in both regions I and II. Thus, our data demonstrates the existence of two independent G-quadruplexes in the *HOX11* breakpoint region.

Since the genome may be interspersed with several G-quadruplex-forming motifs, it was imperative to determine whether there were other such G-quadruplex-forming motifs in *HOX11*, particularly in the vicinity of the region of interest (30–32, 49). This was tested using databases and bioinformatics tools to identify such altered forms of DNA (26, 45). It was observed that apart from the two regions studied, there were no other possible motifs nearby, at least within 2 kb flanking the *HOX11* breakpoint region.

Does the formation of G-quadruplexes in the *HOX11* breakpoint region explain its fragility during generation of the translocation? Previous studies have linked non-B DNA structures like G-quadruplexes, cruciforms, and G-loops to the fragility of genomic loci involved in chromosomal translocations (15, 16, 18, 20, 37, 51, 52). In the case of t(14;18) translocation, it was suggested that the RAG complex could recognize a G-quadruplex structure, formed at the *BCL2* major breakpoint region on chromosome 18, leading to the translocation (18, 19). Cruciform structures have also been shown to form at palindromic AT-rich repeat (PATRR) sequences on chromosome 11. A well-studied example of such a translocation is the t(11;22)(q23;q11), where such a structure was found to form *in vitro*. In bacteria and yeast, chemical modification assays have shown the formation and resolution of such structures inside the cells (53, 54). The breakpoint region in the promoter of the c-*MYC* gene, involved in the t(8;14) translocation in Burkitt's lymphoma, has also been shown to form G-loop structures (52).

In the case of *HOX11*, since the main cluster of breakpoints lies in between the two regions but nearer to region I, it is possible that DNA breaks in this region are a direct or an indirect effect of the observed non-B DNA formation. It appears from the sodium bisulfite modification assay that the intervening region is not single stranded and exists majorly in the duplex form. Overall, our results indicate that the two G-quadruplex structures could independently act as targets for structure-specific/single-strand-specific nucleases and thereby cause fragility in this region (19, 41, 55) (Fig. 10A). Alternatively, such structures once formed may block DNA replication and transcription inside the cells, leading to the generation of single- and double-strand breaks. The resulting breaks could undergo various levels of processing before joining with the partner chromosome. One of the intriguing and unique properties of both the G-rich motifs in the *HOX11* fragile region is their higher propensity to form intermolecular G-quadruplexes. Therefore, there is a possibility that the two independent structures formed at regions I and II interact, thereby contributing toward the fragility of the intervening sequences (Fig. 10B). Although it is plausible, based on the results from extrachromosomal assays, such an interaction may be seen only in a fraction of DNA molecules bearing G-quadruplex structures. Hence, such an interaction could bring the two regions close together, thereby introducing a DNA bend in the intervening region, which could be recognized and cleaved by endonucleases. However, this aspect needs to be investigated further.

Several studies have suggested the possible formation of i-motif structures in C-rich strands, especially when complementary G-rich strands fold into G-quadruplexes. Such i-motif structures can be formed due to intercalated cytosine · protonated cytosine base pairs, generally at slightly acidic conditions (56, 57). In the case of the second G-rich sequence at the *HOX11* fragile region, we could observe a pause site when primer extension was carried out for the C-rich strand, which was abolished upon introducing mutations at the other G-quadruplex-forming motif. This possibly indicates some kind of unusual pairing in the C-rich strand of region II, which could also act as a replication block. Circular dichroism studies also indicated the possibility of i-motif struc-

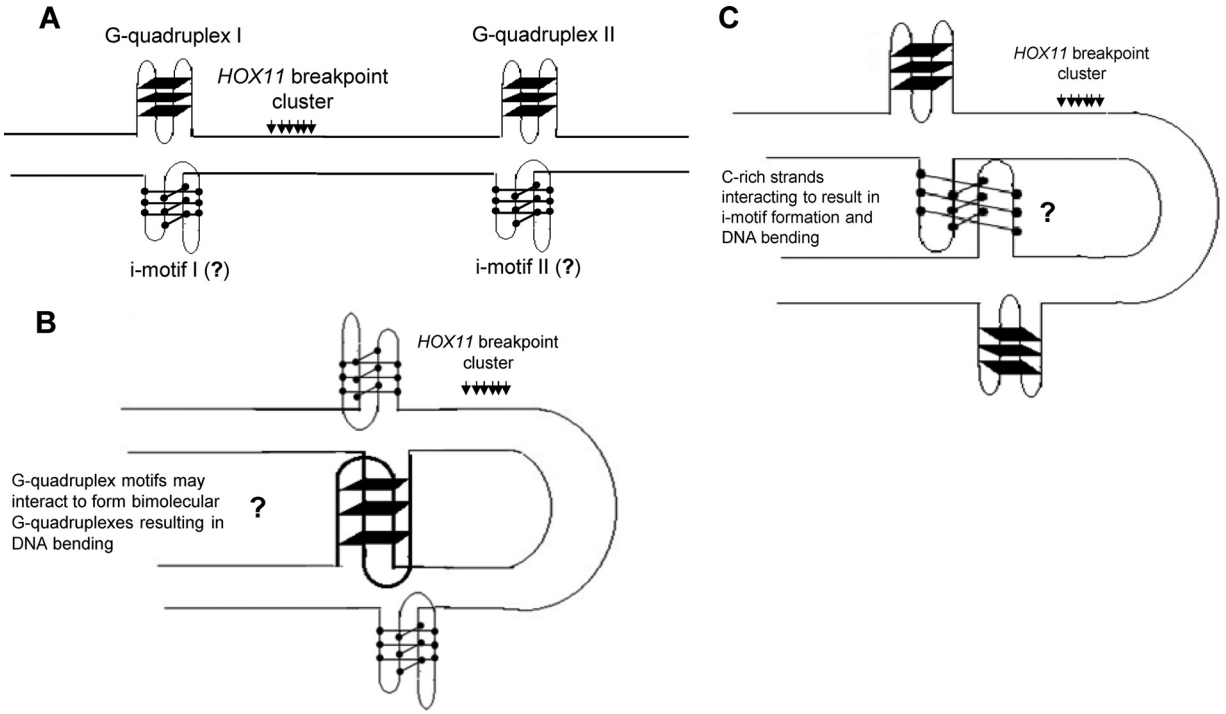

FIG 10 Proposed mechanism of fragility of the *HOX11* breakpoint region. (A) The two G-quadruplexes are formed independently, flanking the patient breakpoint cluster region at *HOX11*, and the complementary C-strands may form i-motifs, under appropriate conditions. These non-B DNA structures can act as targets for structure-specific endonucleases, thereby leading to fragility in this region. (B) A second possibility is that two G-quadruplex motifs interact and form bimolecular G-quadruplex species as shown, when they are in close proximity. (C) In a third scenario, complementary single-stranded cytosine stretches from regions I and II could interact and form an i-motif structure. In both panels B and C, the resulting bending could increase the susceptibility and accessibility of this intervening DNA to nucleases, thereby resulting in breaks and hence translocation.

tures *in vitro*. The bisulfite modification assay showed that several C-to-T conversions at the complementary region corresponding to the G-motifs were noncontinuous, suggesting that even at the genomic level, some secondary interactions may be possible, making the cytosines resistant to bisulfite reactivity. Hence, another plausible hypothesis could be that the complementary C-rich strands could interact with each other probably due to formation of structures like i-motifs (Fig. 10C).

What could aid the formation of such G-quadruplexes inside cells? In order to form the structures, the double-stranded DNA has to unwind, which can happen during processes like replication and transcription. It is possible that once such structures are formed, they are highly stable and do not get immediately resolved. The higher stability of these structures, especially the one formed in region II in the presence of KCl, was evident by the lack of effect of heat on this structure. In addition, previously it has been reported that G-quadruplexes containing shorter or single-nucleotide loops are extremely stable (58, 59). Moreover, our primer extension studies showed that these structures could block progression of polymerases, leading to replication stalling. Hence, nicks could be introduced which might result in double-strand breaks, if the structures remain unresolved, leading to translocations. The presence of these structures could also make the immediate neighboring regions sensitive to certain nucleases. This was suggested by the S1 nuclease assay, which showed that the surrounding sequences were particularly sensitive to the single-strand DNA-specific enzyme.

The role of G-quadruplex structures in protecting and stabilizing the ends of the human genome has been well characterized. However, their role in causing genomic instability is only beginning to be understood. Studies have predicted that G-quadruplex-forming motifs occur once in 10,000 bases in our genome (31, 32). The propensity of occurrence of G-quadruplex motifs at or near many translocation breakpoint regions in lymphoid cancers was shown recently, suggesting that they could be responsible for the generation of many more chromosomal translocations, which have not been studied thus far (34). Hence, future studies will reveal the role of G-quadruplex structures in the generation of many more chromosomal abnormalities in cancer.

ACKNOWLEDGMENTS

We thank B. Choudhary, S. M. Javadekar, V. K. Katapadi, M. Nishana, and members of the SCR laboratory for discussions, help, and comments on the manuscript. pcDNA3.1RFP was a kind gift from Arun Kumar. We acknowledge Krishnamurthy and C-CAMP for help in FACS studies.

This work was supported by grants from CSIR, India [37(1400)/10/EMR-II:2010], to S.C.R. M.N. is supported by a research associate fellowship, IISc, Bangalore, and M.S. and V.G. are supported by SRF and JRF, respectively, from CSIR, India.

We disclose that there are no conflicts of interests.

S.C.R., M.N., and M.S. designed experiments; M.N., M.S., V.G., and S.K.S. conducted the experiments; S.C.R., M.N., and M.S. analyzed the data and wrote the manuscript.

REFERENCES

1. **Korsmeyer SJ.** 1992. Chromosomal translocations in lymphoid malignancies reveal novel proto-oncogenes. Annu. Rev. Immunol. **10:**785–807.

2. **Nambiar M, Kari V, Raghavan SC.** 2008. Chromosomal translocations in cancer. Biochim. Biophys. Acta. **1786:**139–152.
3. **Kagan J, Finger LR, Letofsky J, Finan J, Nowell PC, Croce CM.** 1989. Clustering of breakpoints on chromosome 10 in acute T-cell leukemias with the t(10;14) chromosome translocation. Proc. Natl. Acad. Sci. U. S. A. **86:**4161–4165.
4. **Salvati PD, Watt PM, Thomas WR, Kees UR.** 1999. Molecular characterization of a complex chromosomal translocation breakpoint t(10;14) including the HOX11 oncogene locus. Leukemia **13:**975–979.
5. **Park JK, Le Beau MM, Shows TB, Rowley JD, Diaz MO.** 1992. A complex genetic rearrangement in a t(10;14)(q24;q11) associated with T-cell acute lymphoblastic leukemia. Genes Chromosomes Cancer **4:**32–40.
6. **Kagan J, Joe YS, Freireich EJ.** 1994. Joining of recombination signals on the der 14q- chromosome in T-cell acute leukemia with t(10;14) chromosome translocation. Cancer Res **54:**226–230.
7. **Raghavan SC, Lieber MR.** 2006. DNA structures at chromosomal translocation sites. Bioessays **28:**480–494.
8. **Marculescu R, Vanura K, Montpellier B, Roulland S, Le T, Navarro JM, Jager U, McBlane F, Nadel B.** 2006. Recombinase, chromosomal translocations and lymphoid neoplasia: targeting mistakes and repair failures. DNA Repair (Amst.) **5:**1246–1258.
9. **Hecht F, Morgan R, Hecht BK, Smith SD.** 1984. Common region on chromosome 14 in T-cell leukemia and lymphoma. Science **226:**1445–1447.
10. **Kirsch IR.** 1993. The causes and consequences of chromosomal translocations. CRC Press, Boca Raton, FL.
11. **Hatano M, Roberts CW, Minden M, Crist WM, Korsmeyer SJ.** 1991. Deregulation of a homeobox gene, HOX11, by the t(10;14) in T cell leukemia. Science **253:**79–82.
12. **Raghavan SC, Kirsch IR, Lieber MR.** 2001. Analysis of the V(D)J recombination efficiency at lymphoid chromosomal translocation breakpoints. J. Biol. Chem. **276:**29126–29133.
13. **Marculescu R, Le T, Simon P, Jaeger U, Nadel B.** 2002. V(D)J-mediated translocations in lymphoid neoplasms: a functional assessment of genomic instability by cryptic sites. J. Exp. Med. **195:**85–98.
14. **Zhang M, Swanson PC.** 2008. V(D)J recombinase binding and cleavage of cryptic recombination signal sequences identified from lymphoid malignancies. J. Biol. Chem. **283:**6717–6727.
15. **Nambiar M, Raghavan SC.** 2011. How does DNA break during chromosomal translocations? Nucleic Acids Res. **39:**5813–5825.
16. **Bacolla A, Jaworski A, Larson JE, Jakupciak JP, Chuzhanova N, Abeysinghe SS, O'Connell CD, Cooper DN, Wells RD.** 2004. Breakpoints of gross deletions coincide with non-B DNA conformations. Proc. Natl. Acad. Sci. U. S. A. **101:**14162–14167.
17. **Bacolla A, Wells RD.** 2004. Non-B DNA conformations, genomic rearrangements, and human disease. J. Biol. Chem. **279:**47411–47414.
18. **Nambiar M, Goldsmith G, Moorthy BT, Lieber MR, Joshi VM, Choudhary B, Hosur RV, Raghavan SC.** 2011. Formation of a G-quadruplex at the BCL2 major breakpoint region of the t(14;18) translocation in follicular lymphoma. Nucleic Acids Res. **39:**936–948.
19. **Raghavan SC, Swanson PC, Wu X, Hsieh CL, Lieber MR.** 2004. A non-B-DNA structure at the Bcl-2 major breakpoint region is cleaved by the RAG complex. Nature **428:**88–93.
20. **Kurahashi H, Inagaki H, Yamada K, Ohye T, Taniguchi M, Emanuel BS, Toda T.** 2004. Cruciform DNA structure underlies the etiology for palindrome-mediated human chromosomal translocations. J. Biol. Chem. **279:**35377–35383.
21. **Kurahashi H, Inagaki H, Ohye T, Kogo T, Kato T, Emanuel BS.** 2006. Palindrome-mediated chromosomal translocations in humans. DNA Repair (Amst). **5:**1136–1145.
22. **Mirkin SM, Lyamichev VI, Drushlyak KN, Dobrynin VN, Filippov SA, Frank-Kamenetskii MD.** 1987. DNA H form requires a homopurine-homopyrimidine mirror repeat. Nature **330:**495–497.
23. **Mirkin SM.** 2007. Expandable DNA repeats and human disease. Nature **447:**932–940.
24. **Mirkin SM.** 2008. Discovery of alternative DNA structures: a heroic decade (1979–1989). Front. Biosci. **13:**1064–1071.
25. **Wells RD.** 2007. Non-B DNA conformations, mutagenesis and disease. Trends Biochem. Sci. **32:**271–278.
26. **Cer RZ, Bruce KH, Mudunuri US, Yi M, Volfovsky N, Luke BT, Bacolla A, Collins JR, Stephens RM.** 2011. Non-B DB: a database of predicted non-B DNA-forming motifs in mammalian genomes. Nucleic Acids Res. **39:**D383–D391.
27. **Damas J, Carneiro J, Goncalves J, Stewart JB, Samuels DC, Amorim A, Pereira F.** 2012. Mitochondrial DNA deletions are associated with non-B DNA conformations. Nucleic Acids Res. **40:**7606–7621.
28. **Biffi G, Tannahill D, McCafferty J, Balasubramanian S.** 2013. Quantitative visualization of DNA G-quadruplex structures in human cells. Nat. Chem. **5:**182–186.
29. **Fox KR, Waring MJ, Brown JR, Neidle S.** 1986. DNA sequence preferences for the anti-cancer drug mitoxanthrone and related anthraquinones revealed by DNase I footprinting. FEBS Lett. **202:**289–294.
30. **Neidle S, Balasubramanian S.** 2006. Quadruplex nucleic acids. RSC Biomolecular Sciences, London, United Kingdom.
31. **Todd AK, Johnston M, Neidle S.** 2005. Highly prevalent putative quadruplex sequence motifs in human DNA. Nucleic Acids Res. **33:**2901–2907.
32. **Huppert JL, Balasubramanian S.** 2005. Prevalence of quadruplexes in the human genome. Nucleic Acids Res. **33:**2908–2916.
33. **Huppert JL, Balasubramanian S.** 2007. G-quadruplexes in promoters throughout the human genome. Nucleic Acids Res. **35:**406–413.
34. **Katapadi VK, Nambiar M, Raghavan SC.** 2012. Potential G-quadruplex formation at breakpoint regions of chromosomal translocations in cancer may explain their fragility. Genomics **100:**72–80.
35. **Maizels N.** 2006. Dynamic roles for G4 DNA in the biology of eukaryotic cells. Nat. Struct. Mol. Biol. **13:**1055–1059.
36. **Inagaki H, Ohye T, Kogo H, Kato T, Bolor H, Taniguchi M, Shaikh TH, Emanuel BS, Kurahashi H.** 2009. Chromosomal instability mediated by non-B DNA: cruciform conformation and not DNA sequence is responsible for recurrent translocation in humans. Genome Res. **19:**191–198.
37. **Kogo H, Inagaki H, Ohye T, Kato T, Emanuel BS, Kurahashi H.** 2007. Cruciform extrusion propensity of human translocation-mediating palindromic AT-rich repeats. Nucleic Acids Res. **35:**1198–1208.
38. **Kumar TS, Kari V, Choudhary B, Nambiar M, Akila TS, Raghavan SC.** 2010. Anti-apoptotic protein BCL2 down-regulates DNA end joining in cancer cells. J. Biol. Chem. **285:**32657–32670.
39. **Sharma S, Choudhary B, Raghavan SC.** 2011. Efficiency of nonhomologous DNA end joining varies among somatic tissues, despite similarity in mechanism. Cell. Mol. Life Sci. **68:**661–676.
40. **Nambiar M, Raghavan SC.** 2012. Mechanism of fragility at BCL2 gene minor breakpoint cluster region during t(14;18) chromosomal translocation. J. Biol. Chem. **287:**8688–8701.
41. **Raghavan SC, Swanson PC, Ma Y, Lieber MR.** 2005. Double-strand break formation by the RAG complex at the bcl-2 major breakpoint region and at other non-B DNA structures in vitro. Mol. Cell. Biol. **25:**5904–5919.
42. **Raghavan SC, Chastain P, Lee JS, Hegde BG, Houston S, Langen R, Hsieh CL, Haworth IS, Lieber MR.** 2005. Evidence for a triplex DNA conformation at the bcl-2 major breakpoint region of the t(14;18) translocation. J. Biol. Chem. **280:**22749–22760.
43. **Raghavan SC, Tsai A, Hsieh CL, Lieber MR.** 2006. Analysis of non-B DNA structure at chromosomal sites in the mammalian genome. Methods Enzymol. **409:**301–316.
44. **Sambrook J, Russell DW.** 2001. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
45. **Yadav VK, Abraham JK, Mani P, Kulshrestha R, Chowdhury S.** 2008. QuadBase: genome-wide database of G4 DNA—occurrence and conservation in human, chimpanzee, mouse and rat promoters and 146 microbes. Nucleic Acids Res. **36:**D381–D385.
46. **Mukundan VT, Phan AT.** 2013. Bulges in G-quadruplexes: broadening the definition of G-quadruplex-forming sequences. J. Am. Chem. Soc. **135:**5017–5028.
47. **Voineagu I, Surka CF, Shishkin AA, Krasilnikova MM, Mirkin SM.** 2009. Replisome stalling and stabilization at CGG repeats, which are responsible for chromosomal fragility. Nat. Struct. Mol. Biol. **16:**226–228.
48. **Voineagu I, Narayanan V, Lobachev KS, Mirkin SM.** 2008. Replication stalling at unstable inverted repeats: interplay between DNA hairpins and fork stabilizing proteins. Proc. Natl. Acad. Sci. U. S. A. **105:**9936–9941.
49. **Sinden RR (ed).** DNA structure and function, 1st ed. Academic Press, San Diego, CA.
50. **De Armond R, Wood S, Sun D, Hurley LH, Ebbinghaus SW.** 2005. Evidence for the presence of a guanine quadruplex forming region within a polypurine tract of the hypoxia inducible factor 1alpha promoter. Biochemistry **44:**16341–16350.

51. **Kurahashi H, Shaikh TH, Hu P, Roe BA, Emanuel BS, Budarf ML.** 2000. Regions of genomic instability on 22q11 and 11q23 as the etiology for the recurrent constitutional t(11;22). Hum. Mol. Genet. **9:**1665–1670.
52. **Duquette ML, Pham P, Goodman MF, Maizels N.** 2005. AID binds to transcription-induced structures in c-MYC that map to regions associated with translocation and hypermutation. **24:**5791–5798.
53. **Dayn A, Malkhosyan S, Mirkin SM.** 1992. Transcriptionally driven cruciform formation in vivo. Nucleic Acids Res. **20:**5991–5997.
54. **Cote AG, Lewis SM.** 2008. Mus81-dependent double-strand DNA breaks at in vivo-generated cruciform structures in S. cerevisiae. Mol. Cell **31:** 800–812.
55. **Santagata S, Besmer E, Villa A, Bozzi F, Allingham JS, Sobacchi C, Haniford DB, Vezzoni P, Nussenzweig MC, Pan ZQ, et al.** 1999. The RAG1/RAG2 complex constitutes a 3′ flap endonuclease: implications for junctional diversity in V(D)J and transpositional recombination. Mol. Cell **4:**935–947.
56. **Phan AT, Mergny JL.** 2002. Human telomeric DNA: G-quadruplex, i-motif and Watson-Crick double helix. Nucleic Acids Res. **30:**4618–4625.
57. **Gehring K, Leroy JL, Gueron M.** 1993. A tetrameric DNA structure with protonated cytosine.cytosine base pairs. Nature **363:**561–565.
58. **Risitano A, Fox KR.** 2004. Influence of loop size on the stability of intramolecular DNA quadruplexes. Nucleic Acids Res. **32:**2598–2606.
59. **Phan AT, Modi YS, Patel DJ.** 2004. Propeller-type parallel-stranded G-quadruplexes in the human c-myc promoter. J. Am. Chem. Soc. **126:** 8710–8716.