*10832–10844* *Nucleic Acids Research, 2014, Vol. 42, No. 16* *Published online 19 August 2014*
*doi: 10.1093/nar/gku764*

# A competitive formation of DNA:RNA hybrid G-quadruplex is responsible to the mitochondrial transcription termination at the DNA replication priming site

**Ke-wei Zheng, Ren-yi Wu, Yi-de He, Shan Xiao, Jia-yu Zhang, Jia-quan Liu, Yu-hua Hao and Zheng Tan***

State Key Laboratory of Biomembrane and Membrane Biotechnology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, P.R. China

Received June 18, 2014; Revised August 10, 2014; Accepted August 11, 2014

## ABSTRACT

**Human mitochondrial DNA contains a distinctive guanine-rich motif denoted conserved sequence block II (CSB II) that stops RNA transcription, producing prematurely terminated transcripts to prime mitochondrial DNA replication. Recently, we reported a general phenomenon that DNA:RNA hybrid G-quadruplexes (HQs) readily form during transcription when the non-template DNA strand is guanine-rich and such HQs in turn regulate transcription. In this work, we show that transcription of mitochondrial DNA leads to the formation of a stable HQ or alternatively an unstable intramolecular DNA G-quadruplex (DQ) at the CSB II. The HQ is the dominant species and contributes to the majority of the premature transcription termination. Manipulating the stability of the DQ has little effect on the termination even in the absence of HQ; however, abolishing the formation of HQs by preventing the participation of either DNA or RNA abolishes the vast majority of the termination. These results demonstrate that the type of G-quadruplexes (HQ or DQ) is a crucial determinant in directing the transcription termination at the CSB II and suggest a potential functionality of the co-transcriptionally formed HQ in DNA replication initiation. They also suggest that the competition/conversion between an HQ and a DQ may regulate the function of a G-quadruplex-forming sequence.**

## INTRODUCTION

Mitochondria are cytoplasmic organelles within eukaryotic cells that carry their own genomic materials apart from those in the nucleus. Human mitochondrial DNA (mtDNA) codes 37 genes in a double-stranded closed circular molecule of 16.5 kb. It contains a distinctive guanine-rich (G-rich) motif (GGGGGAGGGGGGGTTTG) denoted conserved sequence block II (CSB II) that directs premature termination of transcription (1). The prematurely terminated transcript serves as a primer to initiate mtDNA replication (1,2). The transition from transcription to primer formation was once proposed by the cleavage of RNA transcript by the mitochondrial RNA processing (RNase MRP) endonuclease (3,4). However, this model is challenged by the fact that the majority of the RNase MRP is localized to the nucleolus (5). The termination is irrespective of the type of RNA polymerase involved and occurs in transcriptions with the mitochondrial RNA polymerase and cofactors or T7 RNA polymerase (6–8). This fact suggests that the transcription termination is primarily determined by the sequence or structural feature of the CSB II (1).

A recent work suggested that the premature termination of transcription at the CSB II is stimulated by G-quadruplex structures formed in RNA transcript (6). G-quadruplexes are four-stranded structures formed by G-rich nucleic acids, in which four G-tracts are held together by Hoogsteen hydrogen bonds in a multi-layered stack of G-quartets (9–11). G-quadruplex formation is stabilized by $K^+$ and $Na^+$, but not by $Li^+$ (12), and involves the 7-nitrogen (N7) in four of the eight Hoogsteen hydrogen bonds in a G-quartet (13). In that work, it was found that transcription using 7-deaza-GTP in place of the normal guanosine triphosphate (GTP) to inhibit RNA from forming G-quadruplex dramatically reduced the CSB II-dependent transcription termination. The termination was more efficient in $K^+$ than in $Li^+$ solution. Single or double G→A mutation in the $G_5AG_7$ core of the CSB II reduced the termination,

*To whom correspondence should be addressed. Tel: +86 10 6480-7259; Fax: +86 10 6480-7099; Email: z.tan@ioz.ac.cn

but those outside of the $G_5AG_7$ did not. All these facts are suggestive of an involvement of G-quadruplex structures that required the participation of RNA within the $G_5AG_7$ tract. A 30-nt RNA oligonucleotide (GAAGCGGGGGAGGGGGGGUUUGGUGGAAAU) covering the CSB II plus the 5 nt upstream and 12 nt downstream of the CSB II formed intramolecular RNA G-quadruplex in an overnight incubation in a $K^+$ or $Na^+$ solution as examined by native gel electrophoresis. It was thus concluded that a predominant unimolecular G-quadruplex uni-G4 whose formation requires guanines within CSB II was the most important quadruplex species to mediate the transcription pretermination (6).

One key question remaining unaddressed in that work is that whether the G-quadruplex seen in the incubation actually forms in transcription, which is important for establishing a definitive connection between the structure and the transcription termination. Recently, we found that DNA bearing two or more G-tracts on the non-template strand readily forms DNA:RNA hybrid G-quadruplex (HQ) structures in transcription by recruiting G-tracts from both the non-template DNA strand and RNA transcript. This is a general phenomenon instead of being characteristic of specific sequences. Such HQs can in turn modulate gene expression under both *in vitro* and *in vivo* conditions (14). We further showed that putative HQ-forming sequences are evolutionally selected and present in the vast majority of genes in warm-blooded animals (15), suggesting a constitutional nature and roles of the HQs in physiological processes. These findings prompted us to investigate whether HQ structures are formed in the transcription of CSB II and involved in the correspondent premature termination of transcription. Using a reconstituted T7 transcription model, here we show that an HQ and an intramolecular DNA G-quadruplex (DQ) of three G-quartet layers are formed at the CSB II upon transcription of the human mtDNA. The two structures compete with each other for the $G_5AG_7$ tract. However, the HQ is more stable and the dominant structure that is responsible to the vast majority of the premature termination of transcription at the CSB II. More importantly, we provide evidence that the type of the G-quadruplex, i.e. HQ or DQ, is crucial in determining the transcription termination.

## MATERIALS AND METHODS

### Plasmid and double-stranded DNA

The plasmid was prepared by inserting into the pMD-19T-simple vector a fragment containing a 247-bp human mtDNA starting from the light strand promoter (LSP) between the T7 promoter and T7 terminator. This plasmid was further constructed by inserting a SP6 promoter and terminator for the experiments in Figure 1C. Mutated plasmid was prepared from the wild type using the Site-Directed Mutagenesis Kit (NEB). Linear double-stranded DNA (dsDNA) in Figures 2 and 4 carried a T7 promoter and the human mitochondrial CSB II sequence. It was prepared by overlap polymerase chain reaction (PCR), as described previously (14). Linear dsDNA in Figure 8B was prepared by PCR using the plasmid in Figure 8A as a template with either normal deoxyguanosine triphosphate (dGTP) or 7-deaza-dGTP/dGTP (4/1, molar) and the three other deoxynucleotide triphosphates (dNTPs).

### *In vitro* transcription and analysis of RNA transcripts

Transcription and analysis of RNA transcripts were carried out as described previously (14). When a T7 transcription was followed by a SP6 transcription, the T7 transcription was stopped by 2 mM competitive dsDNA (5'-GAAATTAATACGACTCACTATA-3')(16,17). Then 0.02 mM fluorescein-UTP (Roche) and 2 U/ml SP6 RNA polymerase (Fermentas, Thermo Scientific) were added and the samples were incubated at 37°C for 60 min, followed by a treatment with 0.04 U/ml DNase I (Fermentas, Thermo Scientific) at 37°C for 15 min and an extraction with phenol/chloroform. RNA products were denatured and resolved on an 8% denaturing polyacrylamide gel.

### Dimethyl sulfate (DMS) footprinting

DMS footprinting was performed as described previously (14).

### Circular dichroism spectroscopy

For intramolecular structures, oligonucleotide was dissolved in 10-mM Lithium Cocadylate buffer (pH 7.4) containing 50-mM LiCl or KCl, heated to 95°C for 5 min and cooled down to 25°C at a rate of 0.03°C/s. For multimeric structures, mixture of oligonucleotides was treated as above, but in buffer containing only 30-mM LiCl. Then 50-mM LiCl or KCl was added and the sample was incubated at 37°C for 1 h. Circular dichroism (CD) spectra were collected from 320 to 220 nm on a CD spectropolarimeter (Chirascan Plus, Applied Photophysics, UK) at 25°C with 0.5-mm pathlength at 1-nm bandwidth. Buffer blank correction was made for all samples.

### Fluorescence resonance energy transfer melting

Oligonucleotides labeled at the 3′ end with a fluorescent donor FAM and 5′ with an accepter TAMRA were purchased from Takara Biotechnology (Dalian). They were dissolved at 0.5 μM in 10-mM lithium cacodylate buffer (pH 7.4) containing 50-mM KCl, denatured at 95°C for 5 min and slowly cooled down to 22°C. Fluorescence resonance energy transfer (FRET) melting was carried out as described (18) by monitoring FAM fluorescence on a CFX96 Touch Real-Time PCR Detection System (Bio-Rad). Samples were initially equilibrated at 22°C for 10 min followed by a stepwise increase in temperature of 0.5°C until 99°C. Derivative of fluorescence over temperature ($dF/dT$) was obtained from the original melting profile to determine the melting temperature.

### Exonuclease I digestion

Oligonucleotide (wild, M3G) was individually dissolved at 1 μM in 10-mM Tris-HCl (pH 7.4), 1-mM ethylenediaminetetraacetic acid (EDTA), 50-mM LiCl, heated at 95°C for 5

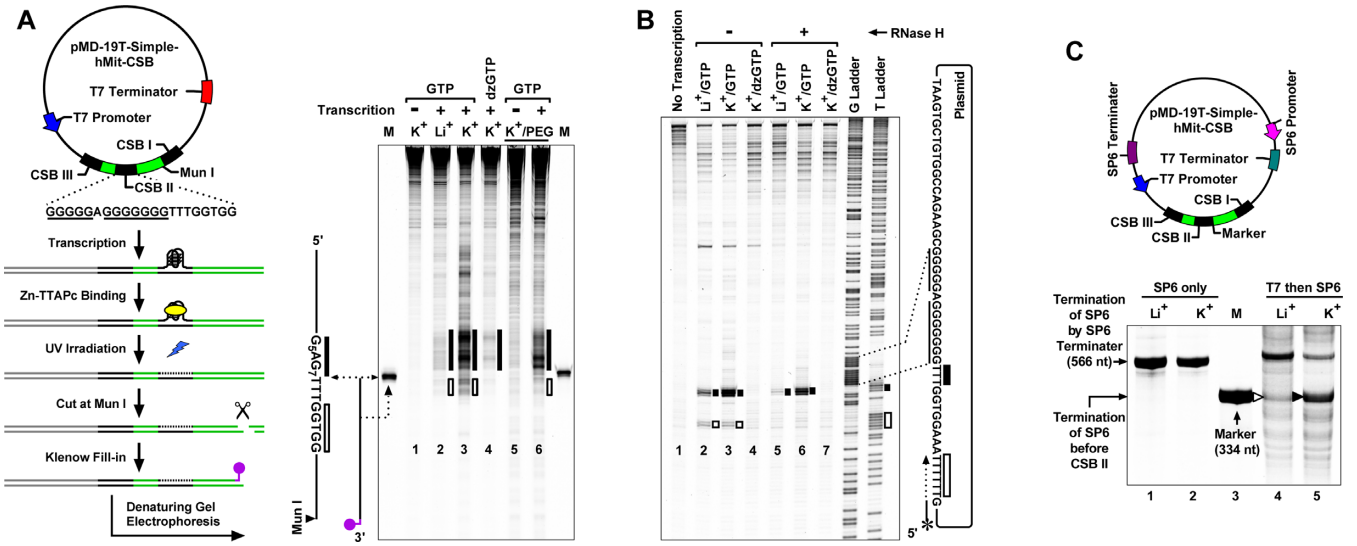


**Figure 1.** HQ formation in transcribed plasmid containing a human mtDNA fragment starting from the light strand promoter (LSP) and containing the CSB II and CSB I. (**A**) Scheme of the plasmid and detection of G-quadruplex formation by ligand-induced photocleavage. Plasmid transcribed in 50 mM of $K^+$ or $Li^+$ solution in the presence or absence of 40% (w/v) PEG 200 with GTP or dzGTP was subjected to Zn-TTAPc-mediated photocleavage, then cut at the Mun I restriction site, and filled in at the recessive 3′ end with a dATP followed by a fluorescein-dUTP, before being resolved on a denaturing gel. The marker (M) was a single-stranded synthetic DNA equivalent to the fragment between the Mun I site and the 3′ end of the $G_5AG_7$ motif. Filled and open bars indicate G-quadruplex-specific cleavage signals. (**B**) Detection of RNA in HQ by photo-crosslinking. Transcription was conducted using normal GTP or dzGTP and 4S-UTP in solution containing 50-mM $K^+$ or $Li^+$. With or without a prior RNase H digestion, transcribed plasmid was crosslinked and precipitated. Then a 5′-FAM-labeled primer (5′-CCAGCCTGCGGCGAGTG-3′) was annealed to the non-template DNA strand downstream of CSB II, followed by extension with DNA sequenase. Extension products were resolved on a denaturing gel. G and T ladders were obtained by primer extension on the non-template DNA strand with ddCTP and ddATP, respectively. Filled and open bars indicate crosslinking sites. (**C**) Detection of G-quadruplex formation by RNA polymerase arrest assay. A plasmid containing convergent T7 and SP6 promoters and the correspondent terminators (top scheme) was transcribed by SP6 RNA polymerase in 50-mM $K^+$ or $Li^+$ solution without (lanes 1 and 2) or with (lanes 4 and 5) a prior transcription with T7 RNA polymerase in the same solution. The T7 transcription was stopped by competitive DNA specific to the T7 polymerase before the SP6 transcription was initiated. Fluorescein-UTP was supplied with SP6 RNA polymerase. RNA transcripts were resolved on a denaturing gel and visualized by the incorporated fluorescein-UTP. The marker represents SP6 transcript terminated right before the CSB II obtained by transcription of a linear DNA amplified from the plasmid.

min and cooled down to room temperature. Two oligonucleotides for the dimeric HQ were mixed and treated in the same way. They were then individually diluted to 0.15 μM into buffer of 40-mM Tris-HCl (pH 7.9), 6-mM $MgCl_2$, 10-mM dithiothreitol (DTT), 2-mM spermidine and 50-mM KCl and maintained at 37°C for 1 h. The dimeric HQ sample was then digested with 0.3-U/μl RNase H (Fermentas, Thermo Scientific) at 37°C for 20 min to remove the RNA at the duplex region. All the three samples (Wild, M3G and HQ) were mixed, digested with 0.67-U/μl Exonuclease I (Fermentas, Thermo Scientific) at 37°C. Aliquot was taken at the indicated time, mixed with three volumes of stop buffer of 95% formamide, 15-mM EDTA, 40-mM NaOH and heated at 95°C for 5 min. Samples were resolved on a 12% denaturing polyacrylamide gel, scanned on a Typhoon 9400 (GE Healthcare) imager and quantitated with the ImageQuant 5.2 software.

## Photo-crosslinking

Photo-crosslinking of transcribed plasmid was carried out as described (14) with modifications. Briefly, after transcription with 4-thio-UTP (TriLink BioTechnologies), samples were treated with or without 0.2-U/μl RNase H (Fermentas, Thermo Scientific) at 37°C for 15 min. They were then transferred to a 24-well microtiter plate (Greiner Bio-One, Germany), placed on ice in a UVP CL-1000 Ultraviolet Cross-linker (UVP), and irradiated for 20 min with 365-nm ultraviolet (UV) light at a distance of 4–5 cm. DNA was recovered with phenol/chloroform extraction and ethanol precipitation. Next, primer extension was performed with a 5′-FAM-ATTACAGGCGAACATACTTAC-3′ primer and Thermo Sequenase from the Cy5 Dye Terminator Cycle Sequencing Kit (GE Healthcare). Guanine and thymine ladders were prepared by primer extension in the presence of ddCTP or ddATP, in a 1/100 molar ratio to dCTP or dATP, respectively.

## Ligand-induced photocleavage

Photocleavage of transcribed plasmid was conducted as described previously (14,19) with modifications. After photocleavage and purification, the plasmid was digested with 50-U Mun I (Fermentas, Thermo Scientific) at 37°C for 1 h, followed by heating at 65°C for 10 min to inactivate Mun I. The digested DNA was purified with the desalting column of the mini Quick Spin Oligo Columns (Roche, Germany). Purified DNA was labeled at the recessive 3′ end in a fill-in extension reaction at 37°C for 30 min with klenow exopolymerase (Fermentas, Thermo Scientific), 20 μM-dATP and 1-μM fluorescein-12-dUTP (Fermentas, Thermo Scientific). The reaction was terminated by adding EDTA to a final concentration of 20 mM. To precipitate the DNA, 60-μl samples were thoroughly mixed with 30 μl of 7.5-M

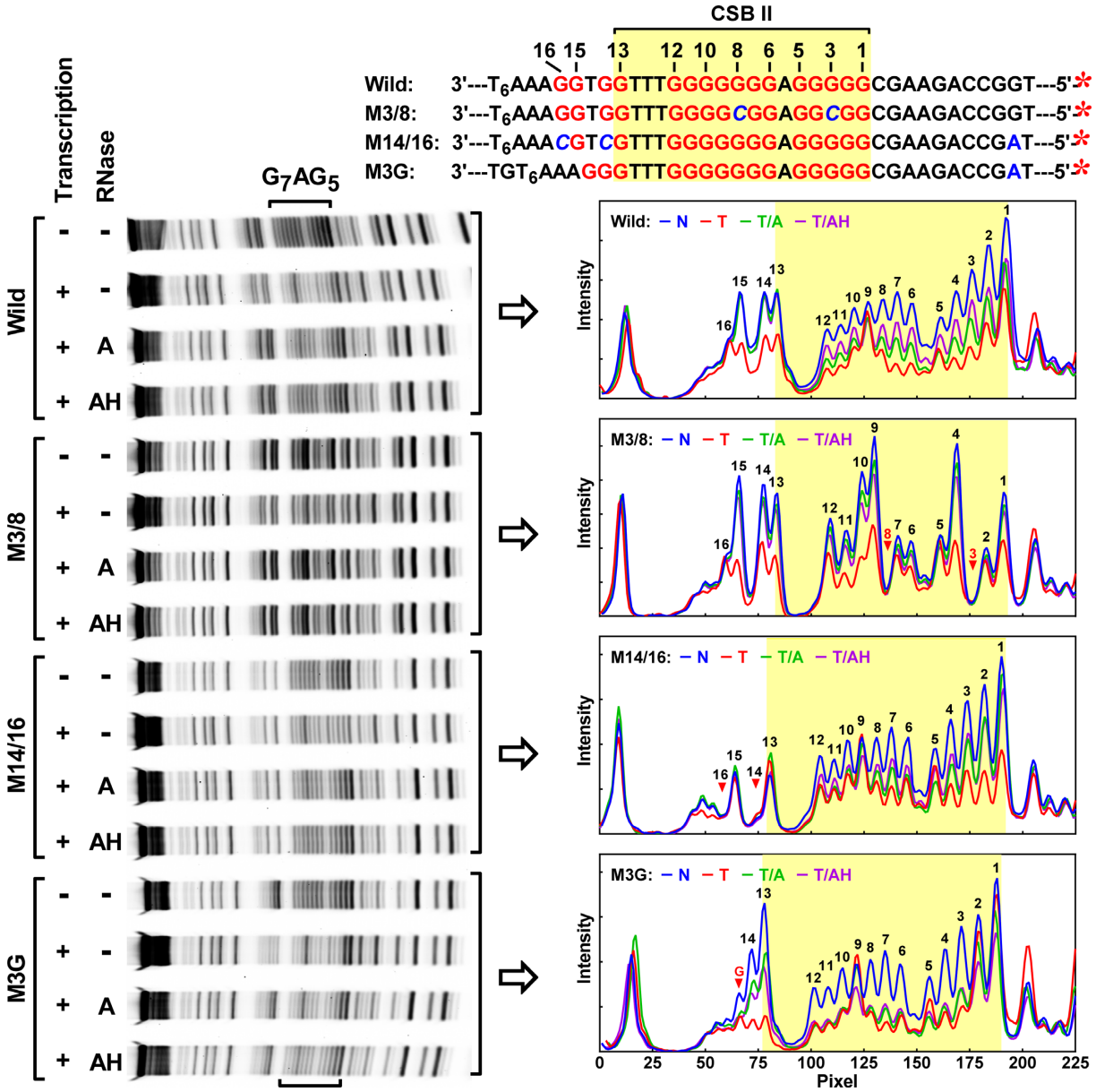


**Figure 2.** Identification of G-quadruplexes in linear dsDNA at the wild and mutated CSB II by DMS footprinting. Linear dsDNA was prepared by PCR using an FAM-labeled upstream primer. The DNA was transcribed in 50mM $K^+$ solution with normal NTP and followed by a treatment with RNase A (A) or RNase A and H (AH) before DMS footprinting. Mutation is indicated by blue character in the sequence. Footprinting cleavage fragments were resolved on a denaturing gel (left) and digitized for graphing (right).

ammonium acetate, 2-μl glycogen (10 mg/ml) and 300-μl 100% ethanol, and the mixture was left at −20°C overnight. Precipitated DNA was dissolved, denatured and resolved on a denaturing 12% polyacrylamide gel (19). The gel was scanned on a Typhoon 9400 (GE Healthcare) imager and processed with the ImageQuant 5.2 software.

## RESULTS

### HQ formation in transcribed plasmid containing mitochondrial CSB II

To inspect the formation of G-quadruplexes in mitochondrial DNA, a 247-nt human mtDNA fragment starting from the LSP was inserted into a pMD-19T-simple plasmid downstream of a T7 promoter and upstream of a T7 terminator (Figure 1A, scheme). G-quadruplex formation was first detected by a ligand-induced photocleavage technique (19). This technique uses a small ligand Zn-TTAPc that has photocleavage activity and high affinity/selectivity toward G-quadruplexes over other structural forms (20–24). After binding to G-quadruplexes, the ligand causes specific cleavage at the guanines in a G-quadruplex upon UV irradiation (19). Transcribed plasmid was incubated with Zn-TTAPc, followed by a UV irradiation, and covalently labeled with a fluorescent dye (Figure 1A, scheme). Photocleavage to the non-template DNA strand was specifically detected at

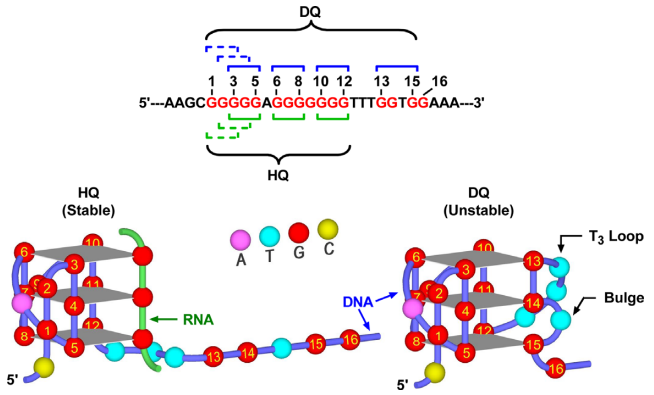


**Figure 3.** Illustration of possible assembly of G-tracts in the formation of HQ and DQ in the transcription at the CSB II. Formation of a G-quadruplex requires four G-tracts. Brackets above and below the sequence indicate the DNA G-tract that can contribute to a DQ or an HQ, respectively. The $G_5$ tract at the 5′ end has three options to supply a G-tract (solid or dashed bracket) for the assembly of the HQ and DQ.

the CSB II region in the plasmid transcribed in a $K^+$ solution, both in the absence and presence of polyethyleneglycol (PEG) (Figure 1A, lanes 3 and 6; filled bar), indicating the formation of G-quadruplexes in the plasmid under the two conditions. In addition, relatively weaker signal was seen downstream of the CSB II motif (lanes 2, 3 and 6; open bar), which could be attributed to possible G-quadruplex formation involving the GGTGG. The cleavages at both sites were G-quadruplex-specific, because they were greatly reduced when the transcription was conducted in a $Li^+$ solution (Figure 1A, lane 2), which does not stabilize G-quadruplex (12). The weak cleavages seen in the $Li^+$ solution might reflect the presence of weak G-quadruplexes stabilized by a low concentration of $Na^+$ (∼10 mM) and $Mg^{2+}$ (8 mM) introduced from the NTP solution and polymerase buffer. Previously, PEG has been used to stabilize G-quadruplex and facilitate its detection in linear dsDNA (25). Our current as well as previous results show that it is not required for plasmid in which G-quadruplex is sustained most likely by the superhelicity of plasmid (26).

We recently reported that the $G_5AG_7$ motif without the GGTGG tract forms an HQ of three G-quartet layers in transcription (16). To examine if HQ was formed at the CSB II, transcription was carried out with a substitution of GTP with 7-deaza-GTP (dzGTP). The 7-nitrogen (N7) in dzGTP is replaced with a carbon that is unable to form the required Hoogsteen hydrogen bond in a G-quadruplex; thus, dzGTP inhibits the RNA from participating in G-quadruplex formation (13). This substitution resulted in very faint cleavage signals at the $G_5AG_7$ motif (Figure 1A, lane 4), suggesting that the majority of the cleavages seen in the transcription with GTP (Figure 1A, lane 3) was induced by HQ that required participation of RNA.

We then used photo-crosslinking to verify the participation of RNA in the formation of HQ in the plasmid in the absence of PEG. A uridine analog 4-thio-uridine (4S-U) was incorporated into the RNA during transcription (27) and crosslinking to the non-template DNA strand catalyzed by light irradiation was detected by primer extension (Figure 1B, scheme at right). Formation of an HQ brings the participating G-tracts in the non-template DNA and those in the RNA transcript into close vicinity. This allows interstrand crosslinking between the two partners near an HQ and stalls the primer extension by DNA polymerase (14,16). It is known that mitochondria DNA forms an R-loop structure in transcription in which a nascent RNA remains hybridized with the template DNA strand (4,28). To differentiate the possible contribution of R-loop, crosslinking was performed before and after a post-transcription digestion with RNase H to cleave the R-loop (16). Extensive crosslinking was detected at three bases most likely through the three 4S-Us right downstream of the CSB II regardless of the RNase H treatment (Figure 1B, lanes 3 and 6; filled bar). The crosslinking was dramatically reduced when the plasmid was transcribed in $Li^+$ solution (lane 2, filled bar), suggesting that the majority of the crosslinking in $K^+$ solution was based on HQ formation. The crosslinking in $Li^+$ solution could be attributed to R-loop because it was greatly reduced after the removal of R-loop with an RNase H digestion (lane 5 versus lane 2). The residual crosslinking remaining (lane 5) could be due to the formation of weak G-quadruplexes promoted by the low concentration of $Na^+$ and $Mg^{2+}$. Taken together, these results demonstrated a presence of RNA in the HQ at the CSB II. In support of this, they all disappeared when the transcription was conducted using dzGTP (Figure 1B, lanes 4 and 7). Besides the major crosslinking site, a few other crosslinking bands were observed at positions away from the CSB II (Figure 1B, lanes 2 and 3; open bar). They possibly aroused from RNA in R-loops and were irrelevant to G-quadruplex because they were insensitive to cation species (lanes 2 versus 3, open bar), but disappeared with a prior RNase H digestion to remove R-loops (lanes 5–7).

Besides the chemical methods, we also detected the formation of G-quadruplex by an RNA polymerase arrest assay. G-quadruplex structure in a DNA strand impairs translocation of motor protein (29). In particular, it arrests RNA polymerase when it forms on the template strand (30). We modified the plasmid to have an SP6 promoter downstream of the T7 terminator and an SP6 terminator upstream of the T7 promoter (Figure 1C, scheme). Without a prior transcription by T7 polymerase, transcription with SP6 yielded full-length RNA transcripts (Figure 1C, lanes 1 and 2). Then the plasmid was first transcribed with T7 polymerase to produce HQ on the template strand for the SP6 transcription, followed by a transcription with SP6 RNA polymerase. In this case, the HQ formed with T7 transcription in $K^+$ solution caused robust termination to the SP6 in front of the CSB II (Figure 1C, lane 5; filled arrowhead). In agreement with the inability of $Li^+$ to stabilize G-quadruplex, only marginal termination was produced when the transcription was done in $Li^+$ solution (Figure 1C, lane 4; open arrowhead), which, again, could be the result of weak G-quadruplexes stabilized by the low concentration of $Na^+$ and $Mg^{2+}$.

### HQ and DQ are both formed at CSB II with different stabilities

To determine the structural assembly of the G-quadruplexes, we carried out DMS footprinting analysis

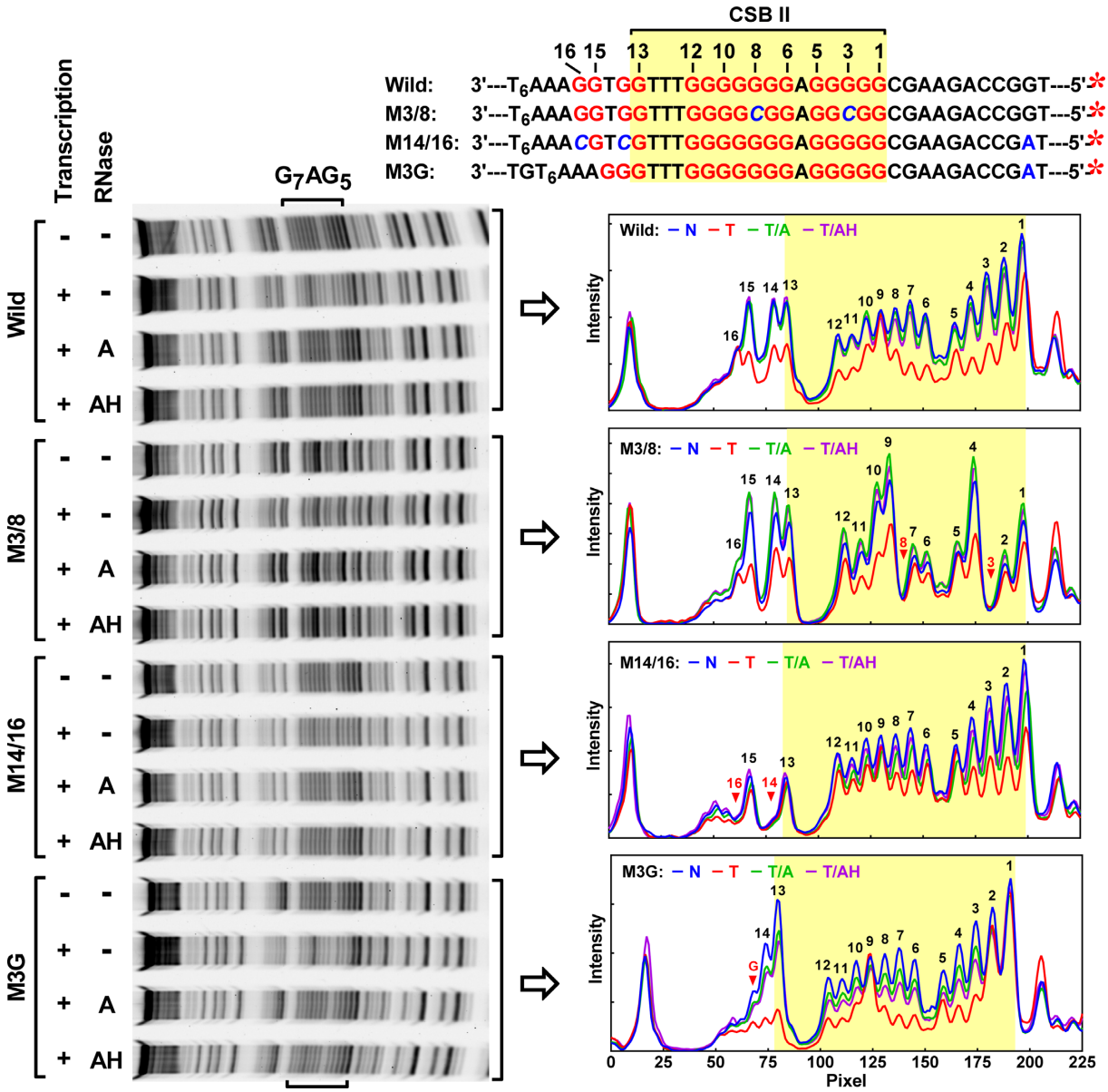


**Figure 4.** Identification of G-quadruplexes in linear dsDNA at CSB II and mutated motifs by DMS footprinting when the formation of HQ is inhibited. Linear dsDNA was prepared and processed as in Figure 2, except that the GTP was replaced by dzGTP in the transcriptions.

on linear dsDNA containing the wild or mutated CSB II before and after a post-transcription digestion with RNase A or both A and H (Figure 2). The N7 of guanine in a duplex is modified by DMS (31), which leads to subsequent cleavage at the guanine; but it is protected in a G-quadruplex by the Hoogsteen bonding (32). For the wild CSB II, a protection was found to the 1G-5G, 6G-8G, 10G-12G and 13G-15G, respectively (wild panel, red versus blue curve). The simplest interpretation of this protection pattern was the formation of DQ of three G-quartet layers using G-tracts from the four regions (Figure 3, blue brackets). Each G in the 1G-5G seemed to have an opportunity to participate because they were all protected. Since the 10G was not protected, we inferred that it was used as a loop, separating the 6G-12G into two G-tracts as we recently reported (16). The protection to the 13G-15G tract suggested that the T within it could be looped out as a bulge (Figure 3, DQ scheme), a feature of a discontinuous G-tract in G-quadruplex reported recently (33).

However, our RNase treatment revealed at least an additional G-quadruplex within the 1G-12G region. It can be noticed that an RNase A digestion totally removed the protection to the 13G-15G (Figure 2, wild panel, green curve). Because RNase A digests the R-loop outside of the 1G-12G region (16), this result suggested that the G-quadruplex involving the 13G-15G was unstable and collapsed without the support of an intact R-loop. In contrast, the protection to the 1G-12G remained although in a reduced degree (green versus red curve). This clearly indicated the presence

of HQ at this region because the $G_5AG_7$ is unable to form a DQ by itself (16). This HQ might be the same as the HQ formed by the $G_5AG_7$ motif without a flanking GGTGG tract we recently reported (16), using three $G_3$ tracts from the $G_5AG_7$ and a $G_3$ tract from the RNA (Figure 3, green brackets and HQ scheme). On the other hand, the reduction in the protection to the 1G-12G upon the RNase A digestion (green curve) was marginal, suggesting that the DQ involving 13G-15G was a minor form and the HQ was likely the major form of the G-quadruplexes at the CSB II. When the R-loop was completely removed by a digestion with both RNase A and H (16), the protection to the 1G-12G was further reduced, but still remained significant. This meant that the R-loop also helped stabilize the HQ.

To verify the presence of HQ at the 1G-12G, we mutated the 3G and 8G into C (Figure 2, M3/8 panel, red arrowheads). In this case, the protection pattern at the 1G-12G was completely disrupted (red curve) comparing to the wild sequence (wild panel, red curve). However, the 13G-15G was still protected as that in the wild sequence. This result suggested that the three Gs might have formed DQ/HQ with other G-tracts (e.g. 9G-12G) available from the DNA or/and RNA. Because this is a quite complicated situation, it is difficult to make any assumptions on the assembly of the possible structures. Anyway, such potential structures were apparently unstable that they largely disappeared after a treatment with RNase A or A and H to digest R-loop (green and purple curves). The capability of the 1G-12G to form HQ was further confirmed by the mutation at 14G and 16G (M14/16 panel, red arrowheads) that prevents the 13G-15G from forming G-quadruplex. In this situation, the 1G-12G was still well protected (red curve) and the RNase treatment only slightly reduced but did not remove the protection (green and purple curves). The presence of a bulge is expected to destabilize a G-quadruplex. When the T within the 13G-15G was removed to provide an intact $G_3$ tract, the G-quadruplex involving the 13G-15G was stabilized and became less susceptible to the RNases (M3G panel, green and purple versus red curves).

Our aforementioned DMS footprinting suggested the formation of a stable HQ structure at 1G-12G and that of an unstable DQ at 1G-15G. To further verify the hybrid and intramolecular nature of the structures, we repeated the transcriptions with a substitution of dzGTP for GTP to prevent the RNA from forming HQ (Figure 4). For the wild CSB II, the same protection pattern (Figure 4, wild panel, red curve) was obtained as the one in the transcription with GTP (Figure 2, wild panel, red curve), confirming the formation of DQ in the 1G-15G region. Without the formation of HQ, the protection was uniformly lost when the DNA was treated with either RNase A or A and H. This is in contrast to a partial loss of protection in the transcription with GTP (Figure 2, wild panel); therefore, it further confirmed that the DQ was unstable and depended on the presence of R-loop for its existence.

When a mutation was introduced to 3G and 8G (Figure 4, M3/8 panel; red arrowheads), a protection was found (red curve) that was roughly identical to that observed in the same DNA in the transcription with GTP (Figure 2, M3/8 panel; red curve). In both cases, the protection all disappeared when the DNAs were treated with the RNases. Because the protection was seen with both GTP and dzGTP, it could reflect either DQ structures or a single-stranded form of the non-template DNA strand which were unstable and collapsed when the R-loop was partially or totally removed (Figures 2 and 4, M3/8 panel; green and purple curves).

When the 14G and 16G were mutated (Figure 4, M14/16 panel; red arrowheads) to prevent the 13G-15G from DQ formation, the protection to the 13G-15G disappeared (red curve). However, the protection to the 1G-12G remained in a reduced degree (red curve). Since it largely disappeared when the R-loop was removed by RNase A or A and H (green and purple curves), this protection might reflect the formation of G-quadruplexes of two G-quartet layers or a single-stranded state of the non-template DNA strand. In the DNA in which the T in the 13G-15G was deleted to provide an intact $G_3$ tract, the formation of DQ was detected as judged from the protection to the Gs in the 1G-15G region (Figure 4, M3G panel; red curve). The protection was dramatically reduced after the RNase treatments (M3G panel; green and purple curves). This result showed that even the bulge was removed, the DQ was still less stable than the HQ at the 1G-12G that offered a protection insensitive to the RNase treatments (Figure 2, M3G panel; red, green and purple curves).

### HQ and DQ formation in oligonucleotides

Our DMS footprinting analysis uncovered two G-quadruplex structures, a stable HQ structure at 1G-12G and an unstable DQ at 1G-15G, formed in the transcription of the CSB II based on the degree of cleavage protection to the guanines before and after RNases digestion. Since a method to evaluate the physical stability of HQ and DQ in plasmid and dsDNA is unavailable, we examined the formation of HQ and DQ in several relevant oligonucleotides by CD spectroscopy and their stabilities by FRET melting and exonuclease hydrolysis.

To assess HQ formation at the $G_5AG_7$ motif, a trimeric structure was assembled by annealing an RNA and DNA oligonucleotide to the same cDNA (Figure 5A, scheme). Although the RNA and DNA are individually unable to form G-quadruplex because they carried less than four G-tracts, the formation of HQ was apparently indicated by a CD spectrum with a negative peak at 245 and a positive peak at 265 nm (Figure 5A), a feature of parallel G-quadruplex (34). The formation of HQ by the $G_5AG_7$ was further verified in a chimeric DNA:RNA oligonucleotide carrying a $G_3$ RNA tract that resulted in a similar CD spectrum (Figure 5B). The peak features of the two HQ structures are the same as the HQs in several recent reports (16,35).

In Figure 5C, DQ formation was analyzed for the wild CSB II motif and its mutants. In agreement with a DQ formation detected in DMS footprinting, the wild sequence displayed a negative peak at 245 nm and a positive peak at 265 nm. The mutation in the M3G provided an intact $G_3$ tract and apparently led to an enhanced formation of DQ as judged from the greater CD peaks. While the spectra of the M3/8 and M14/16 are less characteristic, G-quadruplex formation, either intramolecular or intermolecular, in a small fraction of the molecules might not be excluded.

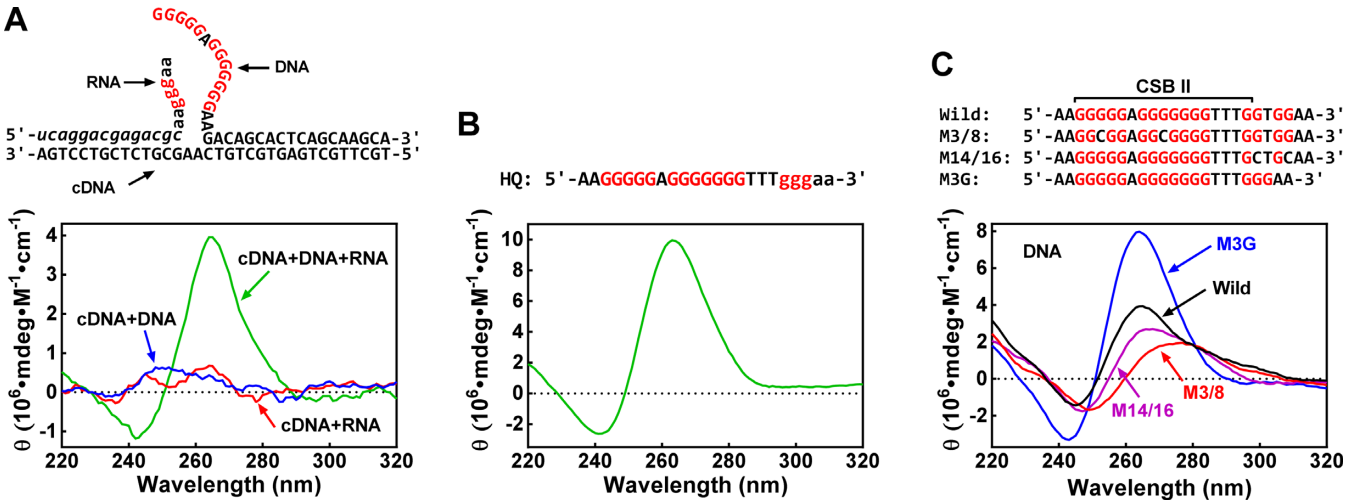


**Figure 5.** Formation of HQ and DQ in synthetic CSB II oligonucleotides in 50-mM $K^+$ by circular dichroism (CD) spectroscopy. (**A**) Intermolecular HQ between DNA and RNA oligonucleotides. Dimeric and trimeric oligomer structures were assembled as shown in the scheme and their CD spectra measured in 50-mM $K^+$ and $Li^+$ solution, respectively. Spectrum obtained in the $K^+$ solution was subtracted by the equivalent obtained in the $Li^+$ solution to remove the contribution of duplex. (**B**) Intramolecular HQ in a chimeric DNA:RNA oligonucleotide. (**C**) Intramolecular DQ in wild and mutant CSB II DNA oligomers. The DNA oligomer or moiety is shown in uppercase and that of RNA in lowercase in (A,B).

### Stability of HQ and DQ formed in oligonucleotides

The formation of HQ and DQ in the oligonucleotides allowed us to examine the stability of the relevant structures by thermal melting using a FRET melting assay (18). In Figure 6A, the FRET melting curve of the structure of the wild oligonucleotide seemed to undergo an unfolding over a wide range of temperature starting from a low temperature until about 55°C. This result confirmed our conclusion in the DMS footprinting that the DQ formed by the 1G-15G was unstable. In contrast, the chimeric HQ oligomer and the M3G DNA mutant both showed a single melting peak at higher temperatures, suggesting more uniform and stable structures were formed.

The stability of the HQ and DQ was further analyzed by an exonuclease hydrolysis assay (36). The formation of HQ or DQ at the 3′ end of the DNA oligonucleotides will protect them from being digested and more stable structure is expected to protect more substrates. In agreement with the melting assay, the protection showed an order of HQ > DQ (M3G) > DQ (wild) which implies a stability at the same order (Figure 6B). The similar values obtained at 1 and 2 min implied that the assay condition was sufficient to fully digest the substrates without a G-quadruplex.

In both the FRET melting and exonuclease hydrolysis assays, the HQ showed a higher stability than the DQ of the M3G DNA that carried a mutation to provide four intact $G_3$ tracts to facilitate the formation of DQ, not to mention the DQs in the wild DNA. Therefore, the much greater stability of the HQ provided strong support to its dominant formation in the transcription of the CSB II DNA manifested in the DMS footprinting (Figures 2 and 4).

### HQ and DQ formation at wild and mutant CSB II in plasmid

Because the DMS footprinting revealed information leading to a predictable formation of HQ and DQ at the wild and mutant CSB II in the linear dsDNAs, we then analyzed the G-quadruplex formation in the correspondent motifs in plasmids by photocleavage and photo-crosslinking. Plasmids were transcribed, processed and analyzed as aforementioned. For the photocleavage (Figure 7A), specific fragments were detected corresponding to cleavages at 1G-12G (lane 2, filled bar) and 13G-15G (lane 2, open bar) in the wild plasmid without RNase treatment. The cleavages at the 13G-15G disappeared but those at the 1G-12G remained after a treatment with either A or A and H. The cleavages at the two distinct sites and their response to RNases treatment exactly correlated with the behavior of HQ and DQ predicted in the DMS footprinting (Figure 2, wild panel).

For the M3/8 mutant, weak cleavages at the 1G-12G (Figure 7A, lane 6; filled bar) and 13G-15G (lane 6; open bar) correlated with a protection at these two regions in DMS footprinting (Figure 2, M3/8 panel; red curve). As the RNase treatment removed the protection in both regions in DMS footprinting (Figure 2, M3/8 panel; green and purple curve), the cleavages were dramatically reduced or disappeared (Figure 7A, lanes 7 and 8). For the M14/16 mutant in which the 13G-15G was mutated to prevent formation of DQ, but allow that of HQ, the fragment became concentrated at the 1G-12G (lane 10, filled bar), but not obvious at the 13G-15G (lane 10, open bar). The resistance to the RNases (lane 11–12, filled bar) also correlated with the protection in the DMS footprinting (Figure 2, M14/16 panel; green and purple curve). For the M3G mutant in which an intact $G_3$ tract was available at the 13G-15G region, photocleavage was also observed in the 1G-12G and 13G-15G regions (Figure 7A, lanes 14–16; filled and open bars). Again this also correlated with the protection in the same regions in the DMS footprinting (Figure 2, M3G panel).

In the photo-crosslinking analysis, transcribed plasmids were subjected to a prior digestion with RNase H to remove R-loop. For the wild, M14/16 and M3G mutants, heavy crosslinking (filled bar) was detected downstream of

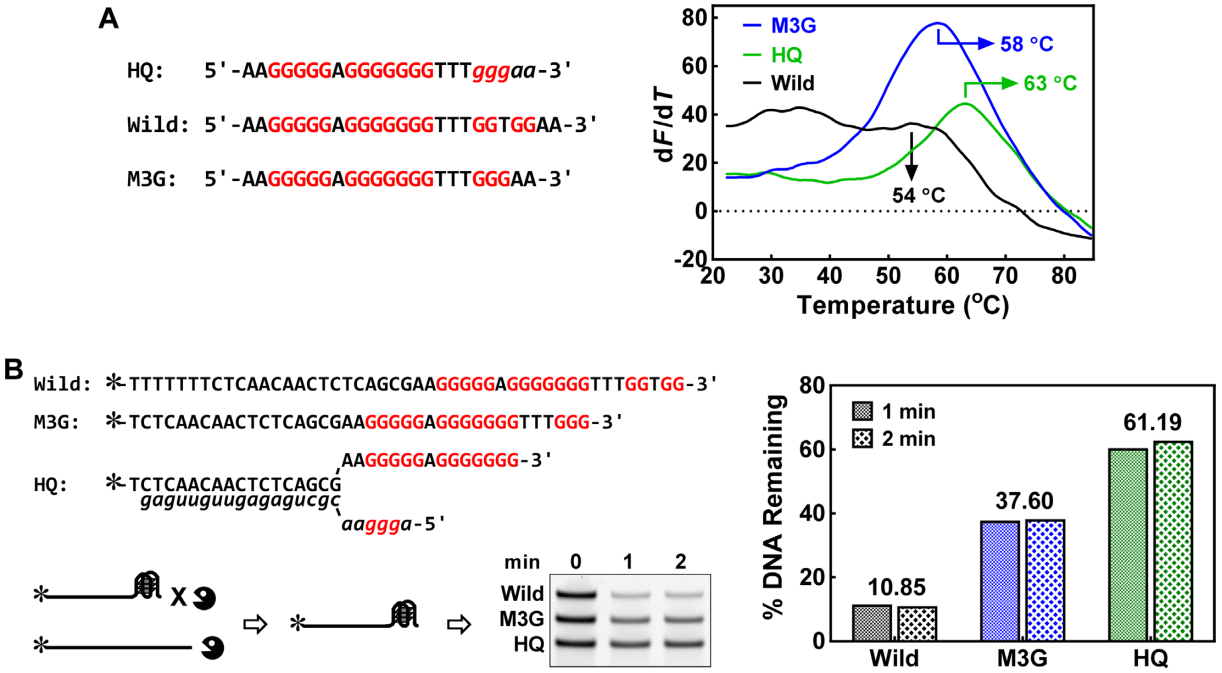


**Figure 6.** Stability of HQ and DQ formed in synthetic CSB II oligonucleotides in 50-mM $K^+$. (**A**) Melting profile of intramolecular DNA DQ and chimeric DNA:RNA HQ. Sequences used are shown on the left. Each of them carried a fluorescent donor FAM at the 5′ end and an accepter TAMRA at the 3′ end. The curves in the graph show the first derivative of FAM fluorescence over temperature as a function of temperature. (**B**) Protection of DNA by the formation of DQ or HQ. DQ or HQ formed in the single-stranded DNA or dimeric DNA:RNA partial duplex protected the DNA from being hydrolyzed from the 3′ end by Exo I exonuclease (scheme at left). Three substrates (Wild, M3G and HQ) were treated with Exo I in a single tube and those survived the hydrolysis were resolved on a denaturing gel. The RNA in the duplex region of the HQ substrate was hydrolyzed by RNase H prior to the exonuclease digestion. The DNAs were visualized by the FAM dye covalently labeled at their 5′ end, digitized, and the results are given on the right. The numbers above the bars indicate the average of the two time points. The DNA oligomer or moiety is shown in uppercase and that of RNA in lowercase in (A,B).

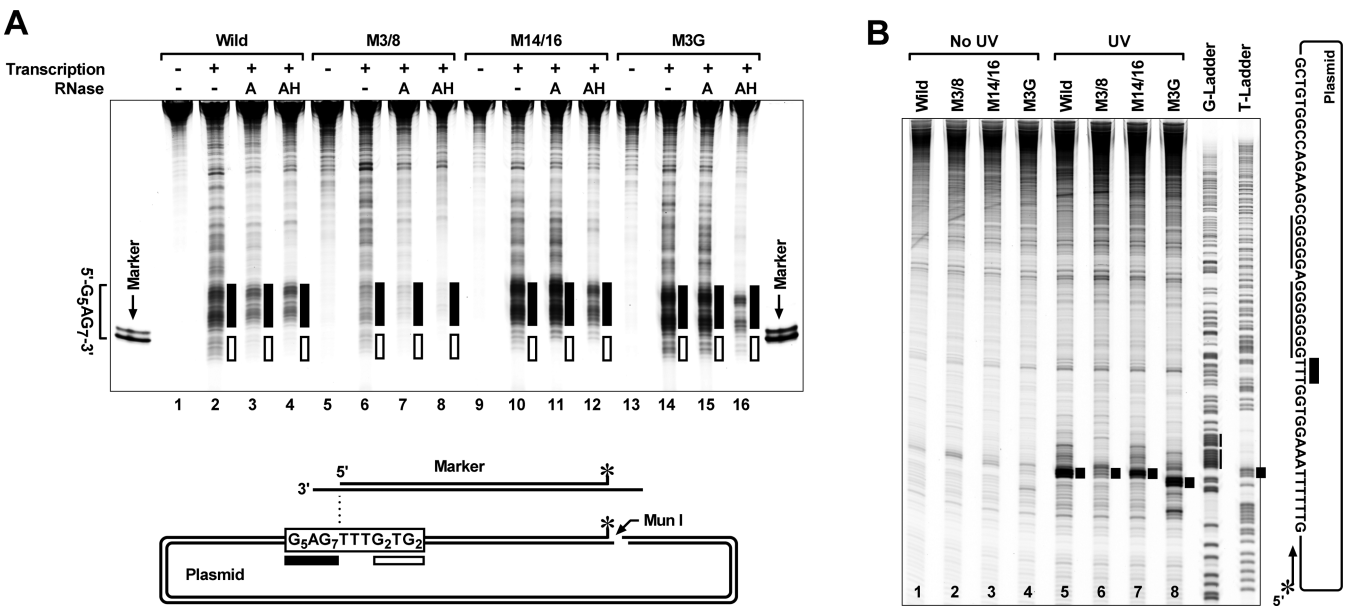


**Figure 7.** Identification of G-quadruplexes in plasmid at the wild and mutated CSB II by (**A**) ligand-induced photocleavage and (**B**) photo-crosslinking. (A) Plasmids transcribed in 50-mM $K^+$ solution were treated with RNase A (A) or A and H (AH), incubated with Zn-TTAPc, and irradiated with UV light. The plasmids were then cut with Mun I, labeled at the 3′ recessive end with an FAM dye by a fill-in reaction using fluorescein-dUTP. Marker was prepared in the same way using a synthetic dsDNA that has the same sequence as the plasmid at the correspondent region (scheme at bottom). The labeling may add one or two Ts, resulting in two bands. Cleavage fragments were resolved on a denaturing gel. (B) Plasmids were transcribed in 50-mM $K^+$ with 4-S-UTP and the other three NTPs, treated with RNase H, followed by UV irradiation. A 5′-FAM-labeled primer was extended on the non-template DNA strand that stalled at the crosslinking sites. Extension products were resolved on a denaturing gel.

the $G_5AG_7$ motif mediated by the 4S-$U_3$ tract (Figure 7B, lanes 5, 7 and 8; filled bar) as in Figure 1B. This also correlated with the formation of an HQ structure predicted at the 1G-12G in the correspondent DMS footprinting analysis (Figures 2 and 4). For the M3/8 mutant in which the $G_5AG_7$ was disrupted to prevent G-quadruplex formation, much faint cleavages were observed (Figure 7B, lane 6), indicating that the crosslinking in the rest three plasmids was specific to the HQ formed at the 1G-12G.

### HQ made a major contribution to the premature termination of transcription at the CSB II

The finding of the formation of a stable HQ in 1G-12G and an unstable DQ in 1G-15G in transcription prompted us to examine how they would affect premature termination of transcription at the CSB II. According to our aforementioned analysis, the three sets of mutations affected the structures differently. This allowed us to seek for connection between a specific structure and its effect on the termination. RNA transcripts produced by the plasmids were resolved by denaturing gel electrophoresis (Figure 8A). Three major premature termination bands were detected at the CSB II for the wild plasmid (lane 5). Disrupting the original DQ and HQ by mutations at 3G and 8G dramatically reduced the termination at the three positions, particularly at bands 2 and 3 (lane 6). The M14/16 mutation does not affect the formation of the HQ that involved only 1G-12G, but disrupts the DQ that required 13G-15G. This mutation showed little effect on bands 2 and 3, but only slightly reduced the intensity of band 1 (lane 7). This result strongly suggested that the DQ had a minor, if any, effect on the termination, or in other words, the HQ made the major contribution. In support of this, the M3G mutation that removes the bulge to stabilize the DQ just slightly enhanced the termination at band 1 and at the same time slightly reduced the termination at band 2 (lane 8). This effect might be indirect because the HQ could be affected via a competition between the HQ and DQ for the $G_5AG_7$ tract. When the transcriptions were performed in $Li^+$ solution, the overall pattern of termination remained among the four plasmids, except that the intensities of the termination bands were all reduced (lanes 1–4).

A most drastic effect was seen when the transcriptions were performed with dzGTP. Under this condition, the DQ was still able to form, but the formation of HQ was inhibited. In all the four plasmids, only a faint band could be detected before the CSB I (Figure 8A, lanes 9–12; bracket). It migrated slightly above band 1 seen in the other conditions, which might be attributed to the substitution of GMP with dzGMP in the RNA products. Among the DNAs, the DQ in the M3G is much more stable than the DQ in the wild DNA, but there was little difference between the two DNAs in the amount of terminations produced (lane 12 versus lane 9). This fact argued that the loss of transcription termination in the dzGTP transcription was not caused by the low stability of the DQ in the wild DNA, but rather by the lack of HQ formation (lane 9 versus lane 5). Therefore, the effect of GTP substitution further confirmed that HQ was the major determinant for the termination.

Because RNA transcript has an identical sequence as the non-template DNA strand, all the mutations introduced into the DNA were also present in the RNA transcript. For this reason, the effect we observed might also be interpreted in the same way based on the correspondent changes in the RNA instead of those in the DNA. To clarify this, we prepared linear dsDNAs by PCR amplification of the four plasmids using 7-deaza-dGTP (dzdGTP) in place of normal dGTP. The substitution of 7-deaza-dGMP for dGMP inhibits the non-template DNA strand from forming either DQ or HQ, but it will not affect the RNA transcript. As is shown in Figure 8B, this modification led to a disappearance of almost all the terminations at the CSB II. In particular, the removal of the bulge in the M3G mutant is expected to result in a more stable RNA G-quadruplex; but this did not enhance the termination (Figure 8B, lane 8). These results clearly show that the RNA alone was unable to mediate the premature termination of transcription, even if a unimolecular G-quadruplex might form during transcription.

## DISCUSSION

In this study, we identified the formation of HQ at the CSB II upon transcription of mtDNA and attributed the HQ as the major G-quadruplex structure responsible to the transcription termination at the CSB II. A recent study reported HQ formation at the CSB II in transcribed mtDNA. In that work, the disappearance of an ~50-nt RNA fragment resistant to RNase A or H upon a substitution of 7-deaza analog of guanine was considered as an indication of HQ formation in the transcription (8). While HQ was indeed detected by CD spectroscopy and gel electrophoresis in a mixture of DNA and RNA oligonucleotides incubated overnight without transcription, this did not necessarily mean such structure would form in transcription. In addition, the size of the ~50-nt RNA fragments is suspiciously too large because the CSB II G-core is merely 14 nt. The formation of HQ involving only a few nucleotides from DNA and RNA is unlikely to protect such a large RNA fragment from the cleavage of the RNases. Because no structural analysis was carried out in the transcribed samples, it was possible that the RNA fragments might represent alternative higher-order structures that were also sensitive to the 7-deaza analog substitution. To seek for a definitive conclusion on the formation of the HQ in transcription, we carried out detailed structural analysis in the transcribed DNAs (Figures 1–4 and 6). The identification of HQ utilized ligand-induced photocleavage, photo-crosslinking, RNA polymerase arrest and DMS footprinting with normal GTP and dzGTP substitution. The participation of DNA in the formation of HQ was directly indicated in these assays and the dzGTP substitution and photo-crosslinking provided direct evidence for the participation of RNA. The contribution of the HQ to the transcription termination was demonstrated by the disappearance of the termination upon the substitution of normal guanine base with the correspondent 7-deaza analog in both the DNA and RNA (Figure 8). This conclusion on the causal factor of the transcription termination differs from that in the recent study in which the premature transcription termination at CSB II was not affected by the substitution

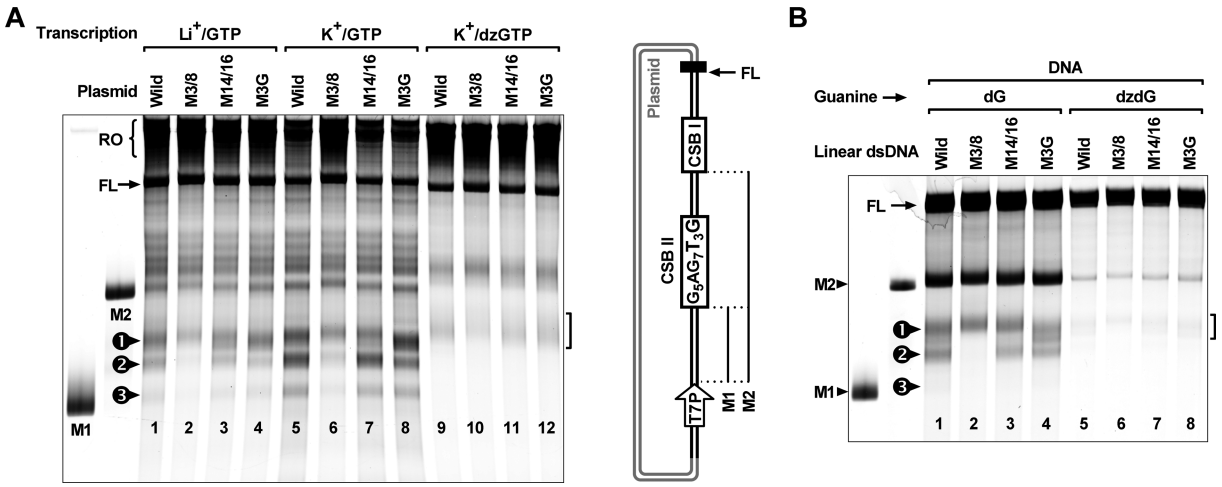


**Figure 8.** Premature termination of transcription at the wild and mutated CSB II in (**A**) plasmid and (**B**) linear dsDNA. (A) Transcription termination depended on the formation of G-quadruplex and participation of RNA. Plasmids were transcribed with GTP or dzGTP and the other three NTPs in 50 mM of $K^+$ or $Li^+$ solution. (B) Transcription termination depended on the formation of G-quadruplex and participation of DNA. Linear dsDNA (black lines in the scheme) was prepared by PCR using normal dGTP (dG) or 7-deaza-dGTP (dzdG) and transcribed with GTP and the other three NTPs in 50 mM of $K^+$ solution. RNA transcripts were resolved on a denaturing gel for both (A) and (B). RO and FL denote run-off and full-length transcripts ended at the T7 terminator, respectively.

of normal guanine base with 7-deaza analog in the DNA (8). While the effect of 7-deaza analog substitution is qualitatively obvious in our data (Figure 8), the confirmation of the irrelevance of the 7-deaza analog substitution to the transcription termination (8) may need a more quantitative assessment.

Besides confirming the formation of HQ, our results show that a DQ could also be formed in transcription at the CSB II (Figure 3, DQ scheme). Both structures used the $G_5AG_7$ DNA motif where the G in the middle of the $G_7$ served as a loop, suggesting a three G-quartet architecture for both the structures (Figure 2). The HQ might recruit three $G_3$ tracts from the DNA and a G3 tract from RNA. Preventing the RNA from forming G-quadruplex by a substitution of GTP with dzGTP abolished a large fraction of the crosslinking (Figure 1B), photo-cleavage (Figure 1A) and DMS protection (Figures 2 and 4) in the assays. This fact implies the formation of HQ and that the HQ was the dominant structure formed in the transcription. Instead of using RNA, the DQ recruited the GGTG from the downstream region of the $G_5AG_7$ DNA motif. The DQ was unstable and it collapsed when the R-loop was partially or totally removed by RNase digestion (Figures 2 and 4, wild panel). This instability is presumably caused by the bulky $T_3$ loop and the bulge that is known to destabilize G-quadruplexes (33,37). The different stabilities of the HQ and DQ were further confirmed by the melting and exonuclease hydrolysis assays of the correspondent G-quadruplexes formed in the oligonucleotides (Figure 6). Because the HQ and DQ both involve the $G_5AG_7$ tract from the CSB II, their formation is competitive of each other. For this reason, the instability of the DQ and the greater stability of the HQ (Figure 6) rendered a priority to the formation of the HQ.

Our results also showed that the HQ is responsible to the majority, if not all, of the premature transcription termination at the CSB II. The irrelevance of the DQ in transcription termination is supported by several facts. On the one hand, disrupting the DQ by mutation (M14/16) had a marginal effect on the termination (Figure 8A, lane 7 versus lane 5; Figure 8B, lane 3 versus lane 1). On the other hand, enhancing the stability of the DQ by mutation (M3G) did not elevate the termination (Figure 8A, lane 8 versus lane 5; Figure 8B, lane 4 versus lane 1), even when the formation of HQ was prevented (Figure 8A, lane 12 versus lane 9). However, preventing either the DNA or RNA from participation in HQ formation by either dzdGTP or dzGTP substitution largely removed the termination (Figure 8A, lanes 9–12 versus 5–8; Figure 8B, lanes 5–8 versus 1–4). Therefore, these facts demonstrated that the transcription termination at the CSB II was determined by the type of the G-quadruplex. They also propose a possibility that the transcription termination may be regulated by manipulating the type of the G-quadruplex, because the HQ and DQ are exclusive to each other in their formation.

The T7 RNA polymerase resembles the structurally unrelated multi-subunit RNA polymerases in key mechanistic features (38). Although our study was carried out with T7 polymerase, previous investigations have shown that the premature termination at CSB II is independent of the type of polymerase and occurs in transcriptions with the mitochondrial RNA polymerase and cofactors or T7 RNA polymerase (6–8). This fact suggests that the transcription terminations share a similar mechanism for the two polymerases (1). The wide spreading of potential HQ-forming sequences in the genomes of animals suggests multiple functions that HQs may play. While their formation and physiological function are beginning to catch attention, the data presented in this work provide a first example of how HQ may play a role in the initiation of DNA replication. In addition, our results demonstrate that the types of G-quadruplex (i.e. HQ or DQ) can lead to different consequences. Therefore, the structural conversion between the HQ and DQ may provide a possibility to manipulate the

function of a G-quadruplex-forming sequence. We speculate that the formation of the HQ may promote the replication initiation by inducing a premature transcription termination and preserving the terminated transcript from being expelled off the template DNA strand. A terminated transcript has to stay annealed with the template DNA strand to prime an initiation of DNA synthesis in mtDNA (1,2). This status can be destroyed because the two DNA strands tend to hybridize back into a duplex. The formation of a HQ may prevent the hybridization of the two DNA strands at the local region, thus facilitating the priming process.

Our conclusion differs from the one in the earlier study (6) with regard to the major structure formed in transcription and assumed responsible to the premature termination of transcription at the CSB II of mtDNA. The results in that work showed participation of RNA transcript in G-quadruplex formation and transcription termination, which do not conflict at all with the formation of HQ and its role in transcription we observed. The mutations they made at the guanines within the $G_5AG_7$ tract dramatically reduced termination while those in the flanking region did not (6). This fact implies that only the $G_5AG_7$ core was involved in the formation of G-quadruplexes that are responsible for the transcription termination, which is also in agreement with the HQ formation and its effect on transcription we observed. However, their failure to analyze the participation of DNA and RNA individually led to a conclusion that attributed the termination to the intramolecular G-quadruplexes formed in the RNA transcript. In principle, the $G_5AG_7$ core alone can only form intramolecular G-quadruplex of two G-quartet layers in principle. G-quadruplexes of two G-quartets are much less stable than those of three G-quartets (39). Alternatively, the RNA could also form an equivalent RNA DQ of three G-quartet layers by using the 13G-15G as the non-template DNA strand did in this work. However, the mutation in the 13G-15G in our work is expected to disrupt (M14/16) or promote (M3G) the formation of RNA DQ, respectively, as to the DNA DQ. But these showed little effect on the termination (Figure 8A, lanes 7 and 8 versus lane 5; Figure 8B, lanes 3 and 4 versus lane 1). This suggested that the RNA DQ was largely irrelevant to the termination. On the other hand, because of the bulky $T_3$ loop and bulge, the RNA DQ might also be unstable as was the DNA DQ. With the participation of $G_3$ tracts from the DNA and RNA, the HQ of three G-quartets was much more stable and should be more competitive to form and to play a role in transcription termination.

## FUNDING

Ministry of Science and Technology of China [2013CB530802, 2012CB720601]; National Science Foundation of China [30970617]; China Postdoctoral Science Foundation [2013T60172].
*Conflict of interest statement.* None declared.

## REFERENCES

1. Pham,X.H., Farge,G., Shi,Y., Gaspari,M., Gustafsson,C.M. and Falkenberg,M. (2006) Conserved sequence box II directs transcription termination and primer formation in mitochondria. *J. Biol. Chem.*, **281**, 24647–24652.
2. Xu,B. and Clayton,D.A. (1995) A persistent RNA-DNA hybrid is formed during transcription at a phylogenetically conserved mitochondrial DNA sequence. *Mol. Cell. Biol.*, **15**, 580–589.
3. Lee,D.Y. and Clayton,D.A. (1997) RNase mitochondrial RNA processing correctly cleaves a novel R loop at the mitochondrial DNA leading-strand origin of replication. *Genes Dev.*, **11**, 582–592.
4. Lee,D.Y. and Clayton,D.A. (1998) Initiation of mitochondrial DNA replication by transcription and R-loop processing. *J. Biol. Chem.*, **273**, 30614–30621.
5. Kiss,T. and Filipowicz,W. (1992) Evidence against a mitochondrial location of the 7–2/MRP RNA in mammalian cells. *Cell*, **70**, 11–16.
6. Wanrooij,P.H., Uhler,J.P., Simonsson,T., Falkenberg,M. and Gustafsson,C.M. (2010) G-quadruplex structures in RNA stimulate mitochondrial transcription termination and primer formation. *Proc. Natl Acad. Sci. U.S.A.*, **107**, 16072–16077.
7. Asari,M., Tan,Y., Watanabe,S., Shimizu,K. and Shiono,H. (2007) Effect of length variations at nucleotide positions 303–315 in human mitochondrial DNA on transcription termination. *Biochem. Biophys. Res. Commun.*, **361**, 641–644.
8. Wanrooij,P.H., Uhler,J.P., Shi,Y., Westerlund,F., Falkenberg,M. and Gustafsson,C.M. (2012) A hybrid G-quadruplex structure formed between RNA and DNA explains the extraordinary stability of the mitochondrial R-loop. *Nucleic Acids Res.*, **40**, 10334–10344.
9. Burge,S., Parkinson,G.N., Hazel,P., Todd,A.K. and Neidle,S. (2006) Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res.*, **34**, 5402–5415.
10. Balasubramanian,S., Hurley,L.H. and Neidle,S. (2011) Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat. Rev. Drug Discov.*, **10**, 261–275.
11. Patel,D.J., Phan,A.T. and Kuryavyi,V. (2007) Human telomere, oncogenic promoter and 5′-UTR G-quadruplexes: diverse higher order DNA and RNA targets for cancer therapeutics. *Nucleic Acids Res.*, **35**, 7429–7455.
12. Hardin,C.C., Perry,A.G. and White,K. (2000) Thermodynamic and kinetic characterization of the dissociation and assembly of quadruplex nucleic acids. *Biopolymers*, **56**, 147–194.
13. Fletcher,T.M., Sun,D., Salazar,M. and Hurley,L.H. (1998) Effect of DNA secondary structure on human telomerase activity. *Biochemistry*, **37**, 5536–5541.
14. Zheng,K.W., Xiao,S., Liu,J.Q., Zhang,J.Y., Hao,Y.H. and Tan,Z. (2013) Co-transcriptional formation of DNA:RNA hybrid G-quadruplex and potential function as constitutional cis element for transcription control. *Nucleic Acids Res.*, **41**, 5533–5541.
15. Xiao,S., Zhang,J.Y., Zheng,K.W., Hao,Y.H. and Tan,Z. (2013) Bioinformatic analysis reveals an evolutional selection for DNA:RNA hybrid G-quadruplex structures as putative transcription regulatory elements in warm-blooded animals. *Nucleic Acids Res.*, **41**, 10379–10390.
16. Zhang,J.Y., Zheng,K.W., Xiao,S., Hao,Y.H. and Tan,Z. (2014) Mechanism and manipulation of DNA:RNA hybrid G-quadruplex formation in transcription of G-rich DNA. *J. Am. Chem. Soc.*, **136**, 1381–1390.
17. Ujvari,A. and Martin,C.T. (1997) Identification of a minimal binding element within the T7 RNA polymerase promoter. *J. Mol. Biol.*, **273**, 775–781.
18. De Cian,A., Guittat,L., Kaiser,M., Sacca,B., Amrane,S., Bourdoncle,A., Alberti,P., Teulade-Fichou,M.P., Lacroix,L. and Mergny,J.L. (2007) Fluorescence-based melting assays for studying quadruplex ligands. *Methods*, **42**, 183–195.
19. Zheng,K.W., Zhang,D., Zhang,L.X., Hao,Y.H., Zhou,X. and Tan,Z. (2011) Dissecting the strand folding orientation and formation of G-quadruplexes in single- and double-stranded nucleic acids by ligand-induced photocleavage footprinting. *J. Am. Chem. Soc.*, **133**, 1475–1483.
20. Ren,L., Zhang,A., Huang,J., Wang,P., Weng,X., Zhang,L., Liang,F., Tan,Z. and Zhou,X. (2007) Quaternary ammonium zinc phthalocyanine: inhibiting telomerase by stabilizing G quadruplexes and inducing G-quadruplex structure transition and formation. *Chembiochem*, **8**, 775–780.
21. Zhang,L., Huang,J., Ren,L., Bai,M., Wu,L., Zhai,B. and Zhou,X. (2008) Synthesis and evaluation of cationic phthalocyanine

derivatives as potential inhibitors of telomerase. *Bioorg. Med. Chem.*, **16**, 303–312.
22. Yaku,H., Murashima,T., Miyoshi,D. and Sugimoto,N. (2012) Specific binding of anionic porphyrin and phthalocyanine to the G-quadruplex with a variety of in vitro and in vivo applications. *Molecules*, **17**, 10586–10613.
23. Yaku,H., Fujimoto,T., Murashima,T., Miyoshi,D. and Sugimoto,N. (2012) Phthalocyanines: a new class of G-quadruplex-ligands with many potential applications. *Chem. Commun. (Camb.)*, **48**, 6203–6216.
24. Yaku,H., Murashima,T., Miyoshi,D. and Sugimoto,N. (2010) Anionic phthalocyanines targeting G-quadruplexes and inhibiting telomerase activity in the presence of excessive DNA duplexes. *Chem. Commun. (Camb.)*, **46**, 5740–5742.
25. Zheng,K.W., Chen,Z., Hao,Y.H. and Tan,Z. (2010) Molecular crowding creates an essential environment for the formation of stable G-quadruplexes in long double-stranded DNA. *Nucleic Acids Res.*, **38**, 327–338.
26. Sun,D. and Hurley,L.H. (2009) The importance of negative superhelicity in inducing the formation of G-quadruplex and i-motif structures in the c-Myc promoter: implications for drug targeting and control of gene expression. *J. Med. Chem.*, **52**, 2863–2874.
27. Harris,M.E. and Christian,E.L. (2009) RNA crosslinking methods. *Methods Enzymol.*, **468**, 127–146.
28. Ohsato,T., Muta,T., Fukuoh,A., Shinagawa,H., Hamasaki,N. and Kang,D. (1999) R-Loop in the replication origin of human mitochondrial DNA is resolved by RecG, a Holliday junction-specific helicase. *Biochem. Biophys. Res. Commun.*, **255**, 1–5.
29. Liu,J.Q., Chen,C.Y., Xue,Y., Hao,Y.H. and Tan,Z. (2010) G-quadruplex hinders translocation of BLM helicase on DNA: a real-time fluorescence spectroscopic unwinding study and comparison with duplex substrates. *J. Am. Chem. Soc.*, **132**, 10521–10527.
30. Broxson,C., Beckett,J. and Tornaletti,S. (2011) Transcription arrest by a G quadruplex forming-trinucleotide repeat sequence from the human c-myb gene. *Biochemistry*, **50**, 4162–4172.
31. Maxam,A.M. and Gilbert,W. (1977) A new method for sequencing DNA. *Proc. Natl Acad. Sci. U.S.A.*, **74**, 560–564.
32. Sun,D. and Hurley,L.H. (2010) Biochemical techniques for the characterization of G-quadruplex structures: EMSA, DMS footprinting, and DNA polymerase stop assay. *Methods Mol. Biol.*, **608**, 65–79.
33. Mukundan,V.T. and Phan,A.T. (2013) Bulges in G-quadruplexes: broadening the definition of G-quadruplex-forming sequences. *J. Am. Chem. Soc.*, **135**, 5017–5028.
34. Randazzo,A., Spada,G.P. and Silva,M.W. (2013) Circular dichroism of quadruplex structures. *Top. Curr. Chem.*, **330**, 67–86.
35. Xu,Y., Ishizuka,T., Yang,J., Ito,K., Katada,H., Komiyama,M. and Hayashi,T. (2012) Oligonucleotide models of telomeric DNA and RNA form a Hybrid G-quadruplex structure as a potential component of telomeres. *J. Biol. Chem.*, **287**, 41787–41796.
36. Yao,Y., Wang,Q., Hao,Y.H. and Tan,Z. (2007) An exonuclease I hydrolysis assay for evaluating G-quadruplex stabilization by small molecules. *Nucleic Acids Res.*, **35**, e68.
37. Pandey,S., Agarwala,P. and Maiti,S. (2013) Effect of loops and G-quartets on the stability of RNA G-quadruplexes. *J. Phys. Chem. B*, **117**, 6896–6905.
38. McAllister,W.T. (1993) Structure and function of the bacteriophage T7 RNA polymerase (or, the virtues of simplicity). *Cell. Mol. Biol. Res.*, **39**, 385–391.
39. Mullen,M.A., Assmann,S.M. and Bevilacqua,P.C. (2012) Toward a digital gene response: RNA G-quadruplexes with fewer quartets fold with higher cooperativity. *J. Am. Chem. Soc.*, **134**, 812–815.