DNA Structures

DOI: 10.1002/anie.201407045

Formation of DNA:RNA Hybrid G-Quadruplexes of Two G-Quartet Layers in Transcription: Expansion of the Prevalence and Diversity of G-Quadruplexes in Genomes**

*Shan Xiao, Jia-yu Zhang, Jun Wu, Ren-yi Wu, Ye Xia, Ke-wei Zheng, Yu-hua Hao, Xiang Zhou, and Zheng Tan**

Abstract: *G-quadruplexes are implicated in important cellular processes. Previous studies mostly focused on intramolecular G-quadruplexes of three or more G-quartets. Those composed of two G-quartets were only shown to form in single-stranded oligonucleotides. On the basis of electrophoresis, DMS footprinting, fluorescence labeling, and photo-cross-linking, we detected the formation of DNA:RNA hybrid G-quadruplexes (HQs) of two G-quartets during the transcription of DNA duplexes. These HQs have a lifetime on the minute scale and are stabilized by a stabilizing ligand. They are far shorter-lived than the HQs of three G-quartets, which last for hours. The occurrence of putative formation motifs of such HQs shows a transcription-dependent strand-biased selection, thus supporting their formation and function in genomes. They are present in almost all human genes in large amounts. We speculate that the two-G-quartet HQs may be a distinct type of G-quadruplexes that may play a role in timely responsive processes and for purposes of fine-tuning.*

Guanine-rich (G-rich) nucleic acids bearing four guanine tracts (G-tracts) can form a four-stranded intramolecular G-quadruplex (GQ). Potential GQ-forming sequences are enriched in functional regions in genomes of a variety of organisms, thus suggesting their role in cellular processes and therapeutic applications.[1] Therefore, the study of G-quadruplex formation in cellular processes is important for understanding the functional role and pharmaceutical potential of G-quadruplexes. A G-quadruplex consists of four G-tracts in a stack of guanine quartets (G-quartets). Whereas most of the G-quadruplexes studied so far are stable structures containing three or more G-quartets, those with only two G-quartets were only shown to form in single-stranded oligonucleotides.[2] Such structures are significantly less stable than those composed of three or more G-quartets.[2b] It is not known whether they can form in double-stranded DNA (dsDNA) in cellular processes. In our recent studies, in which DNA:RNA hybrid G-quadruplexes (HQs) were found to form during transcription, we only detected those with at least three G-quartets.[3] Herein we describe the detection and characterization of two G-quartet HQs formed in transcription by the use of a T7 transcription model and modified experimental settings.

To test whether a two-G-quartet HQ could form between nearby DNA and RNA moieties, we annealed a partial DNA:RNA duplex bearing an overhanging G_3AG_2 motif on both the DNA and RNA (Figure 1, top). When incubated in a solution containing K^+ ions, which stabilize G-quadruplex formation, the duplex appeared as two distinct bands by native gel electrophoresis (Figure 1, lane 2). The faster band (open arrowhead) was recognized as a duplex without a HQ because it also appeared in the same duplex incubated in a solution containing Li^+ ions (lane 4), which do not stabilize G-quadruplexes.[4] The slower band (half-filled arrowhead) is interpreted as a duplex with a HQ because it was not observed

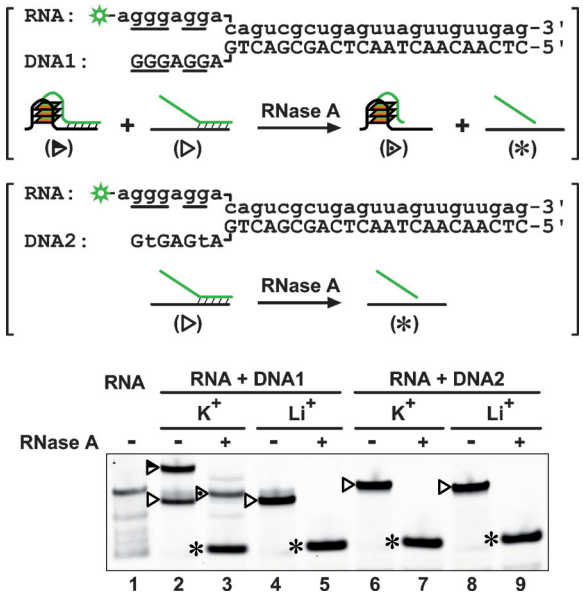

Figure 1. HQ formation between DNA and RNA G_3AG_2 in K^+ solution. A 5′-carboxyfluorescein (FAM)-labeled RNA was annealed with DNA1 or DNA2 to form a partial duplex, which was incubated at 37 °C overnight, digested with RNase A, and then resolved on a native gel at 4 °C. Half-filled and dotted arrowheads indicate the presence of a HQ, and open arrowheads indicate the absence of a HQ.

[*] S. Xiao, J.-y. Zhang, R.-y. Wu, Y. Xia, K.-w. Zheng, Y.-h. Hao, Prof. Dr. Z. Tan
State Key Laboratory of Biomembrane and Membrane Biotechnology, Institute of Zoology, Chinese Academy of Sciences
Beijing 100101 (P. R. China)
E-mail: z.tan@ioz.ac.cn
Dr. J. Wu, Prof. Dr. X. Zhou
College of Chemistry and Molecular Sciences, Wuhan University
Wuhan 430072 (P.R. China)

[**] This research was supported by Grants 2012CB720601 and 2013CB530802 from MSTC and 21072189 from NSFC.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201407045.

in the Li^+ sample (Figure 1, lane 4). To verify the presence of a HQ in the slow band, we treated another set of samples with RNase A before electrophoresis. The RNase A digested the RNA at the hybridization region[3b] and thus released a large portion of the G_3AG_2 RNA motif from the duplex, as indicated by a major fast-migrating band at the bottom of the gel (lanes 3 and 5, asterisk). On the other hand, a retarded band was detected in the K^+ sample (lanes 3, dotted arrowhead), but not in the Li^+ sample (Figure 1, lane 5). This band was an indication of RNA:DNA association and its dependence on K^+ supported the formation of a HQ. The formation of the HQ required the G_3AG_2 motif from the DNA. When this motif was mutated (DNA2) to abolish its participation, no retarded band was seen in the corresponding lanes (Figure 1, lanes 6 and 7 versus 2 and 3).

These results showed that the formation of a HQ of two G-quartets was physically feasible. They prompted us to inspect whether such a HQ could also form in the transcription of dsDNA. In our previous studies, the G-core motifs were kept apart from a T7 promoter in the downstream region.[3a,b] Full-length transcription was conducted in the presence of four nucleoside triphosphates (NTPs), followed by digestion with RNase A or both A and H to cleave the RNA, except that in HQs, and further digestion with protease K to remove the T7 polymerase before structural analysis.[3] This approach might not detect short-lived HQs that do not survive the lengthy post-transcription processing. In the current study, we placed G-cores immediately downstream of a T7 promoter (Figure 2, top). Transcription was performed by supplying only ATP, GTP (and UTP if indicated), so that it could only proceed through the G-cores or beyond by a few nucleotides. Because the RNA transcripts were small, RNase treatment was omitted. By adding sodium dodecyl sulfate (SDS) to stop the reaction and destroy the interaction of the T7 polymerase with the DNA, we could immediately perform HQ detection.

The two methods were compared by the use of four different G-cores for HQ detection by native gel electrophoresis (Figure 2 A), in which DNA bearing a G-quadruplex migrates more slowly than the corresponding DNA without a G-quadruplex.[3,5] All G-cores had less than four G-tracts, so that they were unable to form the canonical intramolecular DNA G-quadruplex. In principle, the G_3AG_3 motif is capable of forming a HQ composed of three G-quartets, whereas the other motifs may form a HQ of only two G-quartets (Figure 2 A, illustration above gels). When transcription was followed by treatment with an RNase and protease digestion, HQ was only detected in the DNA containing the G_3AG_3 motif, as indicated by the appearance of an extra band behind that of the original DNA (Figure 2 A, top gel, lane 2 versus lane 1). Because Li^+ does not stabilize G-quadruplexes,[4] this extra band disappeared when the transcription was carried out in a solution containing Li^+ ions (top gel, lane 3). In contrast, no extra band was observed for the other three G-cores in the presence of K^+ ions (Figure 2 A, top gel, lanes 5, 8, and 11).

By the modified method, we observed additional slowly migrating bands not only in the DNA containing the G_3AG_3 motif (Figure 2 A, middle gel, lane 2) but also in that

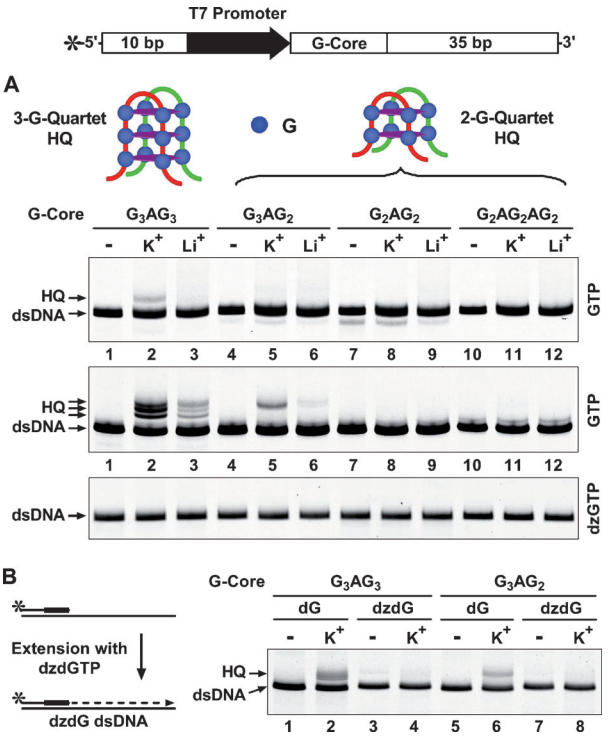

Figure 2. Detection of HQs (composed of two or three G-quartets) formed during transcription by native gel electrophoresis at 4 °C. DNA (top diagram) containing the indicated G-core was not transcribed (−) or transcribed in K^+ or Li^+ solution. DNA was visualized by the FAM dye (asterisk) present as a label at the 5′ end. The colored illustration above the gels show examples of possible HQ structures. A) Top gel: Transcription with four NTPs, followed by digestion with RNase A and H, then with protease K, before electrophoresis. Middle gel: Transcription with only GTP and ATP, followed by termination with SDS before electrophoresis. Bottom gel: Transcription and processing as for the middle gel, but with the substitution of dzGTP for normal GTP. B) Transcription and processing as for the middle gel in (A), but with a dzdG dsDNA sample, in which the dG residue on the nontemplate strand was replaced with dzdG after the promoter (thick bar).

containing the G_3AG_2 motif (middle gel, lane 5), thus suggesting HQ formation in both DNAs. The transcription of G-rich DNA can also produce a structure called an R-loop, in which a nascent RNA transcript remains annealed with the template DNA strand.[6] The slow bands might contain an R-loop because of the omission of RNase treatment. Although this possibility could not be excluded, the presence of a HQ was supported by the substantially reduced intensity of the additional bands (Figure 2 A, middle gel, lanes 3 and 6 versus 2 and 5) when the DNA was transcribed in the presence of Li^+ ions. The faint bands remaining in the Li^+ samples (middle gel, lanes 3 and 6) might be due to the presence of an R-loop. They could also indicate HQ formation promoted by Na^+ (ca. 10 mM) from the NTPs and Mg^{2+} (6 mM) supplied for the activity of T7 RNA polymerase. To confirm the formation of a HQ in the K^+ samples, we transcribed the DNAs by using 7-deaza-GTP (dzGTP) in place of normal GTP[3a,b] to prevent the RNA from forming a G-quadruplex.[7] In this case, no extra slowly migrating band was observed for any of the DNA samples (Figure 2 A, bottom gel), thus suggesting that an R-loop, even if it could form in these DNA molecules, could not

last without a HQ. Therefore, the slowly migrating band could serve as an indication of HQ formation. In the DNA containing the G_3AG_3 motif, the three slower bands (Figure 2A, middle gel, lanes 2 and 3) might be due to the presence of HQs with different conformations.

A HQ unites G-tracts from both DNA and RNA. To inspect whether the G-tracts of the DNA participated in HQ formation, we prepared a dsDNA sample in which the dG residue on the nontemplate strand was replaced with 7-deaza-dG (dzdG) after the T7 promoter by primer extension with 7-deaza-dGTP in place of normal dGTP (Figure 2B, left). This substitution prevented the formation of a HQ (Figure 2B, lanes 4 and 8), thus confirming that the slower bands in the normal DNA samples (lanes 2 and 6) contained a HQ.

We employed DMS footprinting to verify the involvement of G-tracts from the nontemplate DNA strand. The N7 atom of a guanine residue in a G-quadruplex is protected from methylation and subsequent chemical cleavage.[8] The guanine residues in the G_3AG_3 motif were strongly protected (see Figure S1A in the Supporting Information), which is consistent with HQ formation in a large fraction of the DNA (Figure 2A, middle gel, lane 2). For the G_3AG_2 motif, protection was observed, but to a much smaller degree (see Figure S1B), in agreement with the much smaller amount of HQ formation in this DNA (Figure 2A, middle gel, lane 5 versus lane 2).

To detect the presence of RNA in the HQs, we used UTP labeled with fluorescein isothiocyanate (FITC) along with normal ATP and GTP for transcription. The G_3AG_3 and G_3AG_2 motifs were flanked at their 3′ side with ATC and AGTC, respectively. Transcription stopped before the C residue might incorporate a fluorescent FITC–UTP unit into the RNA, although it is not an effective substrate as the native nucleotide. The DNA carrying the G_3AG_3 motif showed a HQ band labeled with FITC (Figure 3A, lane 4). The labeling disappeared when the RNA was prevented from forming a G-quadruplex by the use of dzGTP instead of GTP (lane 5). It appeared, but to a significantly lower degree, in the transcription with Li^+ (lane 6). Similarly, the incorporation of FITC–UTP was also detected in the DNA containing the G_3AG_2 motif when transcribed in K^+ solution with normal GTP (lane 10), but not in that transcribed with dzGTP or in Li^+ solution (Figure 3A, lanes 11 and 12). The reduction in or disappearance of fluorescence labeling in the samples transcribed with dzGTP or in Li^+ solution was not caused by a possible reduction in transcription efficiency, because the different transcription conditions led to little difference in the RNA product (Figure 3B). Therefore, it supports the formation of a HQ in the presence of K^+ ions.

We further verified the presence of RNA in the HQs by cross-linking the RNA transcript with the nontemplate DNA strand.[3a,b] Transcriptions were carried out with ATP, GTP, and 4-thiouridine 5′-triphosphate (4S-UTP). We irradiated the incorporated 4S-UTP with UV light to cross-link it with the proximal nucleotides in the DNA, and the cross-linking sites were subsequently detected by primer extension. The stall signal seen at the TCT region (Figure 3C, lanes 2 and 6,

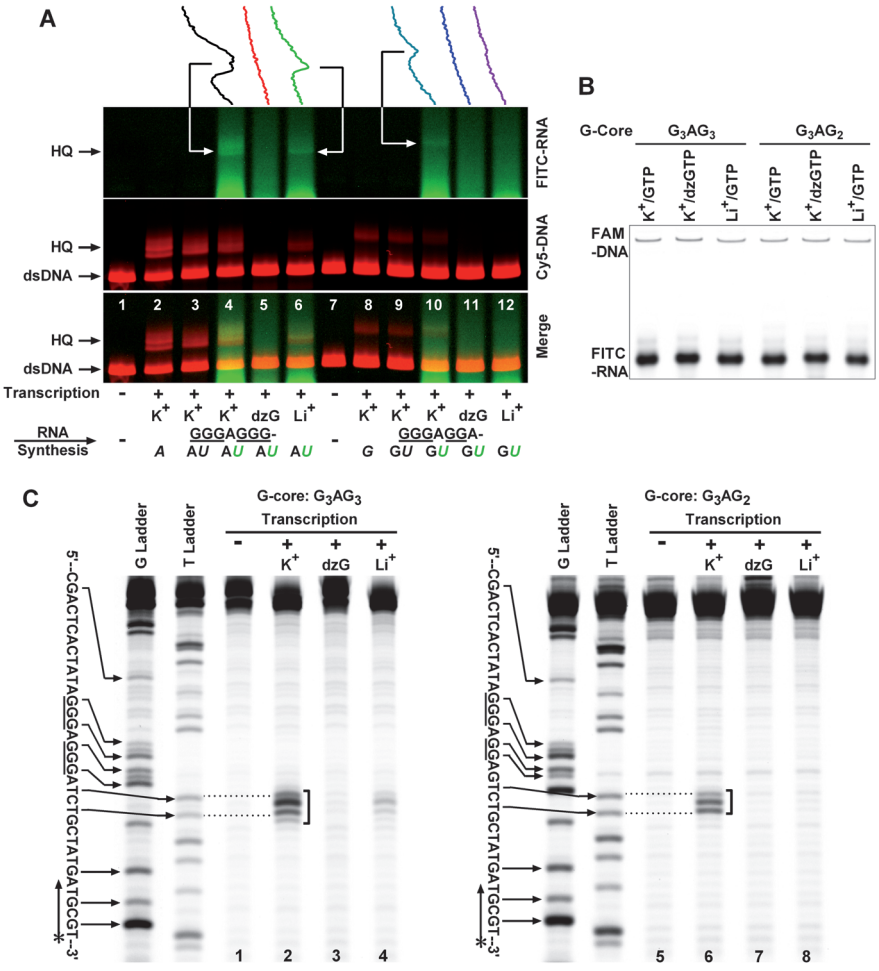

Figure 3. Detection of RNA participation in HQ formation in DNA containing a G_3AG_3 or G_3AG_2 motif. A) Comigration of RNA and DNA in a HQ. DNA was transcribed in K^+ or Li^+ solution with ATP, GTP, or dzGTP (dzG), and UTP (U in black) or FITC–UTP (U in green) and electrophoresed on a native gel as in Figure 2, middle gel. The terminating nucleotide in the RNA synthesis after the G-core is indicated in italic type. Incorporated FITC–UTP in RNA and the Cy5 label at the 5′ end of the nontemplate DNA strands were scanned separately. B) RNA transcripts obtained under the different transcription conditions corresponding to lanes 4, 5, 6, 10, 11, 12 in (A), respectively. The DNA was labeled with a FAM dye that had the same fluorescein fluorophore as the FITC. C) Cross-linking of RNA with DNA. DNA was transcribed in K^+ or Li^+ solution with ATP, GTP, or dzGTP (dzG), and 4S-UTP. The incorporated 4S-UTP was cross-linked to the nontemplate DNA strand by UV irradiation and detected by the stalling of primer extension in the direction indicated by the vertical arrow beside the sequence. The asterisk indicates a FAM fluorescent dye, and the bracket in lanes 2 and 6 indicates HQ-dependent cross-linking.

bracket) demonstrated the presence of RNA near the G-core in the DNA transcribed with K^+. In analogy with the results in Figure 3A, cross-linking was not detected when the participation of RNA in HQ formation was inhibited by dzGTP (lanes 3 and 7), and was greatly reduced (lane 4) or disappeared (lane 8) when the DNA was transcribed in Li^+ solution. The results in Figure 3 indicated that the RNA transcript was required for the formation of a HQ. Taken together, the results in Figures 2 and 3 (and Figure S1 in the Supporting Information) demonstrate a joint participation of the nontemplate DNA strand and the RNA transcript in the formation of the HQ structures.

T7 polymerase undergoes cycles of abortive transcription to produce many short transcripts of 8–12 nts before engaging in a elongation phase to produce a full-length transcript.[9] Because in our experiments the G-cores were placed next to the promoter and transcribed with two or three NTPs, the HQs were formed during the abortive transcription, which generated a large amount of abortive transcripts (Figure 3B). During this process, a nascent RNA molecule first formed a short DNA:RNA heteroduplex (R-loop) before it was released. We recently reported that a HQ of three G-quartets is produced in an R-loop→ssRNA→HQ cascade when a G-core was placed away from the promoter and transcribed in the elongation phase.[3b] HQ formation in the abortive phase also seemed to follow the same cascade because it was suppressed when the RNA transcripts were intercepted by a DNA oligomer complementary to the transcript (Figure 4). For these reasons, we assume that HQs composed of two G-quartets may also form when a G-core is transcribed in the elongation phase. This assumption is supported by the strand bias in the occurrence of the potential DNA:RNA hybrid-G-quadruplex-forming sequence (PHQS) of two G-quartets: This strand bias extends far downstream of the transcription start site (TSS; see Figure S2).

The failure to detect a HQ of two G-quartets by our previous method implies that the HQ was not stable enough to survive the time required to digest the RNA and T7 polymerase. To assess the stability of such HQs, we quantified the HQ bands remaining at different time points after transcription was terminated in the DNA containing a G-core of G_3AG_3, G_3AG_2, and $G_2AG_2AG_2$, respectively. We found that the three-G-quartet HQ formed by the G_3AG_3 motif lasted for hours (Figure 5A). In contrast, the two-G-quartet HQ formed by G_3AG_2 underwent fast decay before reaching a small equilibrium plateau (Figure 5B). It was stabilized by a G-quadruplex stabilizer, Zn-TTAPc,[10] which resulted in an increased lifetime and percentage of HQ at equilibrium. For $G_2AG_2AG_2$, no HQ was detected in the absence of a G-quadruplex stabilizer (Figure 2A, middle gel, lane 11). However, when the DNA was transcribed in the presence of Zn-TTAPc, a HQ was detected that displayed a lifetime of 5.4 min (Figure 5C).

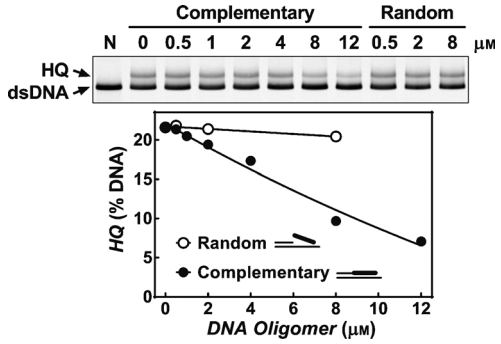

Figure 4. HQ formation in the DNA carrying a G_3AG_2 motif was inhibited when the RNA transcripts were captured by stacking-assisted hybridization with a complementary DNA molecule. Transcription was conducted as for Figure 2, middle gel in the presence of a partial DNA duplex bearing an overhang that was not or was complementary to the RNA transcript (thick bar in graph).

To investigate the biological relevance of the two-G-quartet HQs, we surveyed the occurrence of PHQSs of the two-G-quartet HQs in several eukaryotic and prokaryotic genomes. Gene sequences were searched for motifs satisfying $G_{\geq 2}(N_{1-7}G_{\geq 2})_{\geq 1}$, in which G denotes guanine and N any nucleotide, including G. Each motif found was then searched for $G_{\geq 3}(N_{1-7}G_{\geq 3})_{\geq 1}$ and $G_{\geq 2}(N_{1-7}G_{\geq 2})_{\geq 3}$, which represent motifs with the potential to form a G-quadruplex of three or more G-quartets and an intramolecular DNA G-quadruplex of two-G-quartets, respectively. If any match was found, the motif was discarded. All species showed a non-random, strand-biased distribution with more PHQS on the nontemplate than on the template DNA strand (see Figure S2), in analogy with that of the HQs composed of three or more G-quartets.[3c] The strand bias is transcription-dependent because it is only present downstream, but not upstream of the TSS. Since the formation of a HQ requires G-tracts on the nontemplate DNA strand, this result suggests a preferential selection of PHQSs on the nontemplate strand and supports the formation of two-G-quartet HQs and their physiological role in the genomes.

There are approximately 370000 putative G-quadruplex motifs in human genome.[11] Because a HQ can form with as few as two G-tracts instead of four on the nontemplate DNA strand, our recent survey revealed an additional approximately 1600000 PHQS motifs of three or more G-quartets that are present in $>97\%$ of human genes.[3a] PHQSs of two-G-quartet HQs are present in about 98% of human genes, and their number reached >1900000 (see Table S1 in the Supporting Information). Because we used G-cores of only 1 nt loops in our experiments and did not know if the HQ could form with longer loops, only those with 1 nt loops were counted. Thus, the formation of two-G-quartet HQs, if possible in vivo, will further expand the prevalence and diversity of G-quadruplexes in the genome far beyond what we previously anticipated.

RNA G-quadruplexes of two G-quartets have been shown to modulate gene expression in vivo.[12] The thrombin-binding aptamer (TBA) interacts with its target by adopting a two-G-quartet intramolecular DNA G-quadruplex.[13] In principle, two-G-quartet HQs may be a unique type of G-quadruplexes that are unstable and more flexible. Their folding and unfolding are more responsive to transcription activity, which may be utilized to sense transcription in a timely manner or for interaction with other factors in a similar way to a TBA. On the other hand, a G-quadruplex imposes a physical obstacle to the translocation of a protein on a DNA

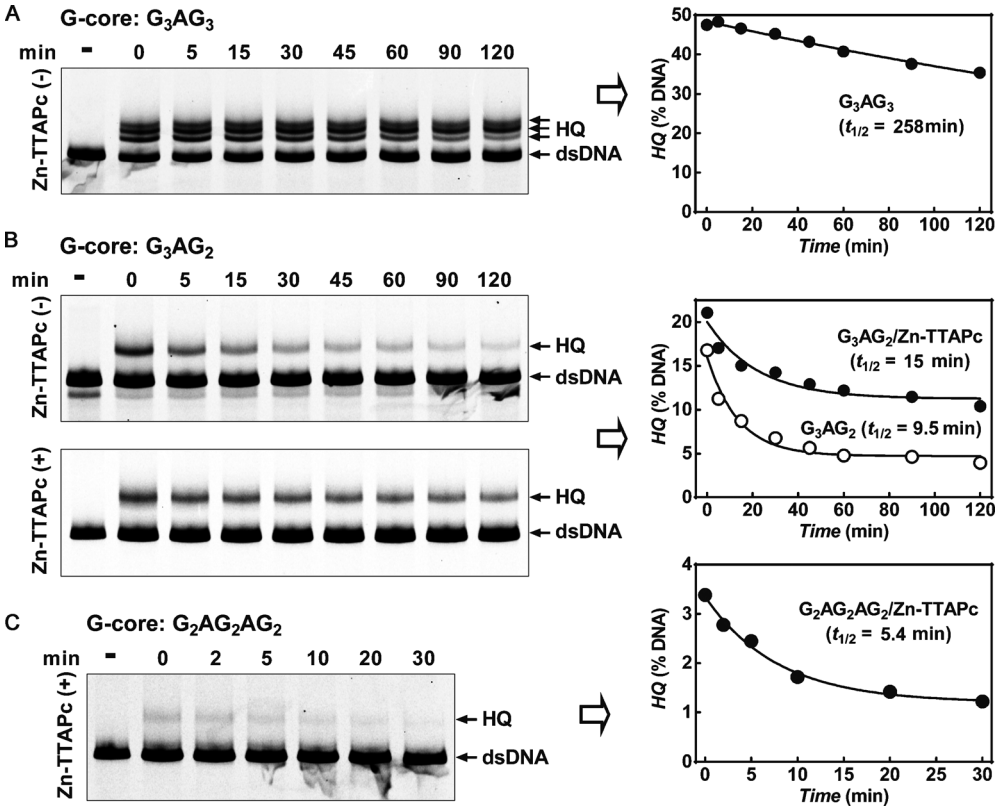

Figure 5. HQ remaining after transcription as a function of time. DNA containing the indicated G-core was transcribed as for Figure 2, middle gel in the absence or presence of the G-quadruplex stabilizer Zn-TTAPc. Transcription was stopped, and the samples were maintained for various periods at 37 °C before resolution by native gel electrophoresis. HQ-containing DNA was quantitated as a percentage of the total DNA in each lane.

tract.[14] In *Escherichia coli*, consecutive transcriptions can occur at intervals of about 10 min with large variation.[15] Thus, the lifetime of a two-G-quartet HQ could be longer than the interval between consecutive transcription events. As a result, transcriptions may be affected by such HQs. Whereas stable HQs composed of three or more G-quartets may exert a more profound influence on transcription, unstable two-G-quartet HQs may be more suited for fine-tuning purposes. G-quadruplexes are considered unique molecular targets in pharmaceutical applications.[1] Owing to their low stability, two-G-quartet HQs may be more susceptible or responsive to G-quadruplex-stabilizing ligands and may thus have an impact in such applications.

Received: July 9, 2014
Revised: August 17, 2014
Published online: September 29, 2014

Keywords: DNA · G-quadruplexes · genomes · RNA · transcription

[1] a) G. W. Collie, G. N. Parkinson, *Chem. Soc. Rev.* **2011**, *40*, 5867 – 5892; b) S. Balasubramanian, S. Neidle, *Curr. Opin. Chem. Biol.* **2009**, *13*, 345 – 353; c) J. Bidzinska, G. Cimino-Reale, N. Zaffaroni, M. Folini, *Molecules* **2013**, *18*, 12368 – 12395; d) P. Murat, S. Balasubramanian, *Curr. Opin. Genet. Dev.* **2014**, *25*, 22 – 29.

[2] a) E. A. Raiber, R. Kranaster, E. Lam, M. Nikan, S. Balasubramanian, *Nucleic Acids Res.* **2012**, *40*, 1499 – 1508; b) M. A. Mullen, S. M. Assmann, P. C. Bevilacqua, *J. Am. Chem. Soc.* **2012**, *134*, 812 – 815; c) Z. Zhang, J. Dai, E. Veliath, R. A. Jones, D. Yang, *Nucleic Acids Res.* **2010**, *38*, 1009 – 1021; d) K. W. Lim, S. Amrane, S. Bouaziz, W. Xu, Y. Mu, D. J. Patel, K. N. Luu, A. T. Phan, *J. Am. Chem. Soc.* **2009**, *131*, 4301 – 4309; e) S. L. Palumbo, R. M. Memmott, D. J. Uribe, Y. Krotova-Khan, L. H. Hurley, S. W. Ebbinghaus, *Nucleic Acids Res.* **2008**, *36*, 1755 – 1769.

[3] a) K. W. Zheng, S. Xiao, J. Q. Liu, J. Y. Zhang, Y. H. Hao, Z. Tan, *Nucleic Acids Res.* **2013**, *41*, 5533 – 5541; b) J. Y. Zhang, K. W. Zheng, S. Xiao, Y. H. Hao, Z. Tan, *J. Am. Chem. Soc.* **2014**, *136*, 1381 – 1390; c) S. Xiao, J. Y. Zhang, K. W. Zheng, Y. H. Hao, Z. Tan, *Nucleic Acids Res.* **2013**, *41*, 10379 – 10390.

[4] C. C. Hardin, A. G. Perry, K. White, *Biopolymers* **2000**, *56*, 147 – 194.

[5] a) C. Zhang, H. H. Liu, K. W. Zheng, Y. H. Hao, Z. Tan, *Nucleic Acids Res.* **2013**, *41*, 7144 – 7152; b) K. W. Zheng, Z. Chen, Y. H. Hao, Z. Tan, *Nucleic Acids Res.* **2010**, *38*, 327 – 338.

[6] A. Aguilera, T. García-Muse, *Mol. Cell* **2012**, *46*, 115 – 124.

[7] T. M. Fletcher, D. Sun, M. Salazar, L. H. Hurley, *Biochemistry* **1998**, *37*, 5536 – 5541.

[8] D. Sun, L. H. Hurley, *Methods Mol. Biol.* **2010**, *608*, 65 – 79.

[9] T. A. Steitz, *Curr. Opin. Struct. Biol.* **2009**, *19*, 683 – 690.

[10] Zn-TTAPc = tetrakis(2-trimethylaminoethylethanol) phthalocyaninato zinc tetraiodine; a) L. Zhang, J. Huang, L. Ren, M. Bai, L. Wu, B. Zhai, X. Zhou, *Bioorg. Med. Chem.* **2008**, *16*, 303 – 312; b) H. Yaku, T. Murashima, D. Miyoshi, N. Sugimoto, *Molecules* **2012**, *17*, 10586 – 10613.

[11] a) J. L. Huppert, S. Balasubramanian, *Nucleic Acids Res.* **2005**, *33*, 2908 – 2916; b) A. K. Todd, M. Johnston, S. Neidle, *Nucleic Acids Res.* **2005**, *33*, 2901 – 2907.

[12] M. Wieland, J. S. Hartig, *Chem. Biol.* **2007**, *14*, 757 – 763.

[13] K. Padmanabhan, K. P. Padmanabhan, J. D. Ferrara, J. E. Sadler, A. Tulinsky, *J. Biol. Chem.* **1993**, *268*, 17651 – 17654.

[14] a) C. Broxson, J. Beckett, S. Tornaletti, *Biochemistry* **2011**, *50*, 4162 – 4172; b) J. Q. Liu, C. Y. Chen, Y. Xue, Y. H. Hao, Z. Tan, *J. Am. Chem. Soc.* **2010**, *132*, 10521 – 10527.

[15] A. B. Muthukrishnan, M. Kandhavelu, J. Lloyd-Price, F. Kudasov, S. Chowdhury, O. Yli-Harja, A. S. Ribeiro, *Nucleic Acids Res.* **2012**, *40*, 8472 – 8483.

