JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

ARTICLE

pubs.acs.org/JACS

Dissecting the Strand Folding Orientation and Formation of G-Quadruplexes in Single- and Double-Stranded Nucleic Acids by Ligand-Induced Photocleavage Footprinting

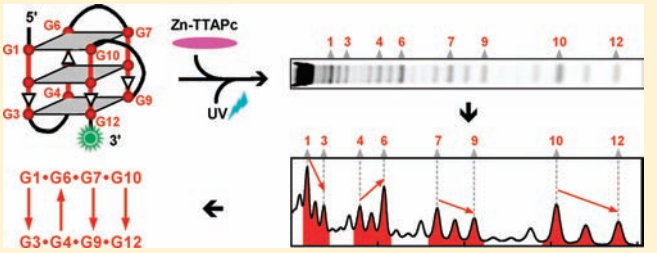
Ke-wei Zheng,[†] Dan Zhang,[‡] Li-xia Zhang,[‡] Yu-hua Hao,[†] Xiang Zhou,[‡] and Zheng Tan*[,†]

[†]State Key Laboratory of Biomembrane and Membrane Biotechnology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, People's Republic of China

[‡]College of Chemistry and Molecular Sciences, Wuhan University, Wuhan 430072, People's Republic of China

Supporting Information

ABSTRACT: The widespread of G-quadruplex-forming sequences in genomic DNA and their role in regulating gene expression has made G-quadruplex structures attractive therapeutic targets against a variety of diseases, such as cancer. Information on the structure of G-quadruplexes is crucial for understanding their physiological roles and designing effective drugs against them. Resolving the structures of G-quadruplexes, however, remains a challenge especially for those in double-stranded DNA. In this work, we developed a photocleavage footprinting technique to determine the folding orientation of each individual G-tract in intramolecular G-quadruplex formed in both single- and double-stranded nucleic acids. Based on the differential photocleavage induced by a ligand tetrakis(2-trimethylaminoethylethanol) phthalocyaninato zinc tetraiodine (Zn-TTAPc) to the guanines between the two terminal G-quartets in a G-quadruplex, this method identifies the guanines hosted in each terminal G-quartets to reveal G-tract orientation. The method is extremely intuitive, straightforward, and requires little expertise. Besides, it also detects G-quadruplex formation in long single- and double-stranded nucleic acids.



INTRODUCTION

Nucleic acids with four tracts of consecutive guanines can fold into a noncanonical four-stranded G-quadruplex structure in which the guanines in the G-tracts construct into a multilayered stack of planner G-quartets via Hoogsteen hydrogen bonding (Figure 1).[1] So far, a huge number of putative G-quadruplex sequences have been mapped in the genome of human and other species by bioinformatics analysis.[2−10] Such sequences are enriched in telomere and gene promoter regions and have been found to play important roles in the regulation of telomere extension and gene expression.[8−14] As a result, G-quadruplexes are becoming promising drug targets against a variety of diseases. For instance, the G-quadruplex formed by telomeric DNA is not a substrate for telomerase.[15] G-Quadruplex-stabilizing molecules have been shown to inhibit telomerase activity[16] and induce growth arrest, senescence, and apoptosis in cancer cells.[17,18] More details on G-quadruplexes in telomere region, genomic DNA, RNA, and their therapeutic applications can be found in many recent reviews.[19−25]

A G-quadruplex can assume a variety of topologies according to the relative folding orientation of each of the four G-tracts.[26,27] Knowledge on the structure of G-quadruplexes is crucial for developing effective drugs targeting G-quadruplexes and for understanding the physiological function of G-quadruplexes as well. There are a number of physical techniques readily available for studying different aspects of G-quadruplex structures.[24,28] Among them, circular dichroism (CD)[29] is a technique that has been widely employed to reveal whether the G-tracts are parallel or antiparallel with each other; the nuclear magnetic resonance spectroscopy (NMR)[30] and X-ray crystallography[31] can provide 3D information at atomic resolution. The radioprobing method analyzes strand folding via breaks produced by decay of ^{125}I incorporated into DNA.[32−34] These techniques have been an indispensable tool for studying G-quadruplex structures. Despite this, each of these techniques has certain limitations. For example, the determination of folding conformation by CD is empirical, and certain antiparallel G-quadruplexes exceptionally showed a CD profile that is generally assigned to parallel structure.[35−37] The NMR technique requires high concentration samples at millimolar level[30] that may lead to the formations of intermolecular structures that are unlikely to form in genomic DNA where a single copy of G-quadruplex strand is present. For the crystallography, the structure resolved in the solid state may

Received: October 5, 2010
Published: January 5, 2011

dx.doi.org/10.1021/ja108972e | J. Am. Chem. Soc. 2011, 133, 1475–1483

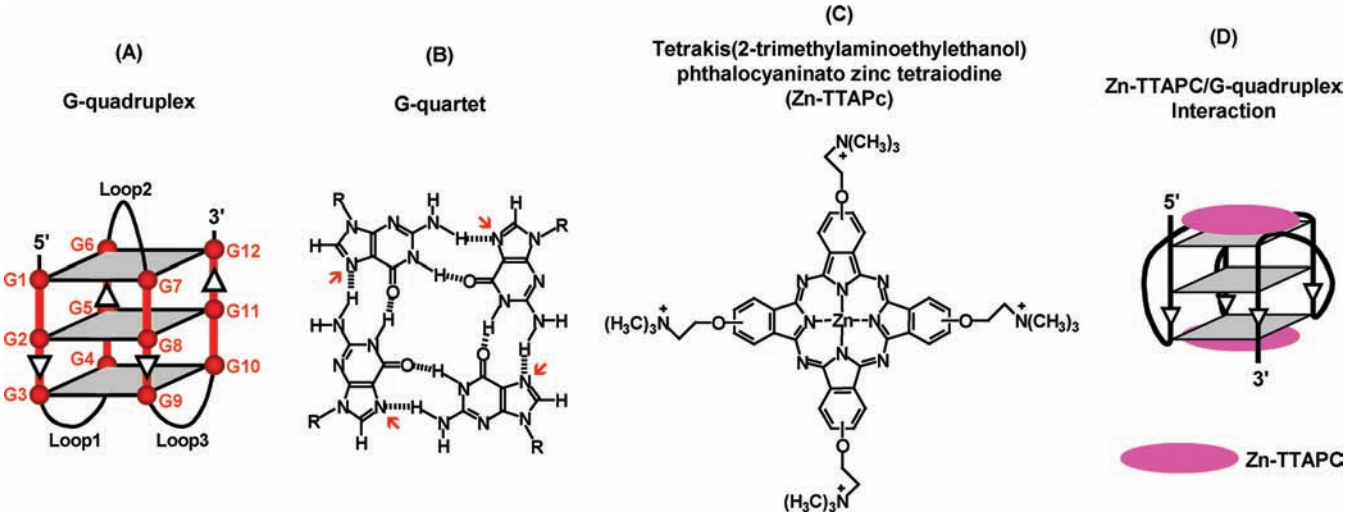


Figure 1. Structure of G-quadruplex, G-quartet, tetrakis(2-trimethylaminoethylethanol) phthalocyaninato zinc tetraiodine (Zn-TTAPc), and their interactions. (A) A human telomere G-quadruplex structure showing an antiparallel basket-type topology consisting of three stacked G-quartets with four three-guanine G-tracts (red rods) oriented consecutively in a ↓↑↓↑ direction starting from the 5′ end as indicated by the arrowheads. (B) Structure of G-quartet in which the N7 (red arrows) of the guanines is protected from methylation by DMS and subsequent chemical cleavage. (C) Structure of Zn-TTAPc. (D) Reported interaction of Zn-TTAPc with the parallel human telomere G-quadruplex.[40]

not necessarily represent the native structure adopted in solution. The radioprobing requires dual labeling in which a ^{125}I has to be incorporated at a nucleotide near or within the G-quadruplex region. This extra labeling limits its application to only those targets that can be labeled with ^{125}I before G-quadruplex forms. Moreover, it takes a couple of weeks to accumulate strand break. While the application of these techniques has been successfully used to short single-stranded DNA, resolving G-quadruplex structure in long double-stranded DNA would be strongly desired because of the double-stranded nature of genomic DNA.

In the present work, we describe a ligand-induced photocleavage footprinting technique that can identify the orientation of each of the four G-tracts in an intramolecular G-quadruplex formed in both single- and double-stranded nucleic acids. This method is based on the phenomenon that a compound tetrakis-(2-trimethylaminoethylethanol) phthalocyaninato zinc tetraiodine (Zn-TTAPc) binds G-quadruplex and, under ultraviolet (UV) irradiation, induces differential cleavages to the guanines between the two terminal G-quartets in the G-quartet core of a G-quadruplex; thus the guanines accommodated in each terminal G-quartet can be identified by the relative cleavage of the two guanines at the two ends of each G-tract to deduce the orientation of the G-tracts. This method is extremely simple, straightforward, and requires little expertise. We expect it may provide a useful addition to the available techniques for the analysis of G-quadruplex structure, especially those in double-stranded DNA.

METHODS AND MATERIALS

Chemicals and Sample Preparation. Zn-TTAPc was synthesized as described.[38] Oligonucleotides were purchased from (Takara, China) and made in buffer of 10 mM Tris-HCl (pH 7.4) containing 1 mM EDTA and the indicated concentration of KCl, NaCl, or LiCl, respectively, with or without 40% (w/v) PEG 200. For single-stranded DNA (ssDNA), G-quadruplex was prepared by heating DNA to 95 °C and then slowly cooling to room temperature. Double-stranded DNA (dsDNA) was annealed at high concentration in 10 mM Tris-EDTA, pH 7.4, 150 mM LiCl by heating to 95 °C followed by slowly cooling. G-Quadruplex was generated by heat denaturation/renaturation or in vitro transcription in the presence of 40% (w/v) PEG 200 as recently described.[39] DMS footprinting was conducted as described.[39]

Purification of Transcribed DNA. Twenty microliters of Streptavidin−Sepharose 4B conjugate (Invitrogen, America) was mixed with 200 μL of 0.1 μM biotinylated dsDNA after heat denaturation/renaturation or transcription. After incubation on ice for 1 h with vortexing several times, the sample was centrifuged at 1000*g* for 5 min. The pellet was washed with buffer of 10 mM Tris-HCl (pH 7.4), 1 mM EDTA, and 40% (w/v) PEG 200 three times before photocleavage.

Photocleavage with Zn-TTAPc. 0.1 μM fluorescein (FAM) labeled ssDNA or dsDNA in 200 μL volume was put on ice for 5 min, and then mixed with 0.1 μM final concentration of Zn-TTAPc. After incubation on ice for 5 min, the sample was mixed with 100 μg of sperm DNA and left on ice for another 5 min. The sample was then transferred to a 24-well microtiter plate (Greiner Bio-One, Germany) and irradiated for 10 min with 365 nm UV light in a UVP CL-1000 Ultraviolet Crosslinker (UVP, America), followed by addition of β-mercaptoethanol to 0.72 M final concentration. After two phenol/chloroform extractions and one ethanol precipitation, the DNA was dissolved in 80% (v/v) deionized formamide in water, denatured at 95 °C for 5 min, and resolved on 19% (ssDNA) or 12% (dsDNA) denaturing polyacrylamide gel. The gel was scanned on a Typhoon 9400 (GE Healthcare, America) imager and processed using the software ImageQuant 5.2.

RESULTS

The Zn-TTAPc used is a water-soluble, positively charged phthalocyanine derivative carrying four ammonium group on the periphery.[38] With a large rigid planar configuration, phthalocyanines have been shown to have especially good shape complementarity with the G-quadruplex quartet plane and interact with the parallel human telomere G-quadruplex via stacking externally to the terminal G-quartets[40] in a 2/1 ligand/DNA stoichiometry (Figure 1).[38,40−42] In addition, phthalocyanines have been shown to stabilize and induce G-quadruplex formation and inhibit telomerase activity.[38,41] Under light irradiation, phthalocyanines cleave DNA via generation of singlet oxygen.[43−45] Because of these properties, we attempted to explore the phthalocyanine

dx.doi.org/10.1021/ja108972e |*J. Am. Chem. Soc.* 2011, 133, 1475–1483

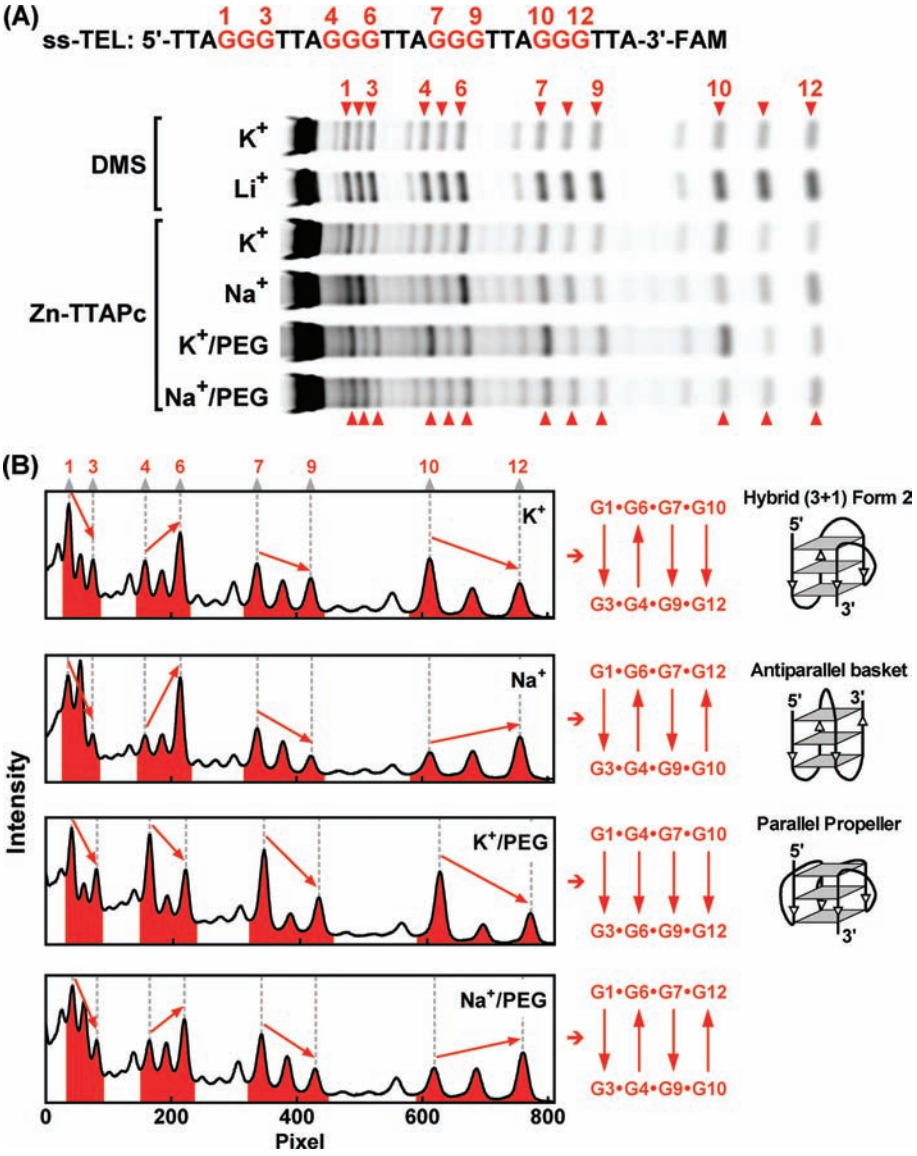


Figure 2. Photocleavage footprinting of human telomere G-quadruplexes in various solutions. (A) Gel electrophoresis of fragments produced by Zn-TTAPc-induced photocleavage to G-quadruplexes prepared in 150 mM K^+ or Na^+ solution in the absence or presence of 40% PEG 200, respectively. The DMS footprinting was included to indicate the guanine residuals by protection in Li^+ versus that in K^+ solution. (B) Intensity profiles of the photocleavage footprinting obtained by density scan of the gel. The bands of guanines indicated by red "▲" and peaks in red correspond to the four G-tracts participated in G-quartet formation. The red arrows in the graph were drawn according to the intensity of the first and last peaks of each G-tract and then combined to obtain the folding orientations of the G-tracts. Drawing at the right side of each profile shows the known structure of the corresponding G-quadruplex identified by other techniques.[46−48,50]

derivatives as probes for G-quadruplex structures. After testing several derivatives with the G-quadruplexes formed by the single-stranded human telomere sequence $T_2AG_3(T_2AG_3)_3T_2A$, we found that the Zn-TTAPc could produce differential photocleavage to the guanines between the two terminal G-quartets in the G-quadruplexes that faithfully reflects the strand folding orientation of each individual G-tract (Figure 2).

Analysis of Human Telomere G-Quadruplexes. We first show the results obtained in K^+ solution. Because K^+ is the most abundant cation in animal cells (150 mM), telomere G-quadruplex in K^+ solution is the most extensively studies structure. It is known that the human telomeric sequences $T_2AG_3(T_2AG_3)_3T_2A$,[46] $TAG_3(T_2AG_3)_3T_2$,[47] and $T_2AG_3(T_2AG_3)_3T_2$[46] form a major antiparallel/parallel hybrid (3 + 1) Form 2 structure in K^+ solution. For the $T_2AG_3(T_2AG_3)_3T_2A$ G-quadruplex in our study, the G1, G6, G7, and G10 had a higher cleavage than the G3, G4, G9, and G12 (Figure 2), which correlated with the actual arrangement of these guanines: the former four guanines participate in the 5′ terminal G-quartet and the latter four in the 3′ terminal G-quartet. This phenomenon seems to show that guanines at the same G-quartet tend to undergo similar photocleavage; therefore, the relative extent of cleavage to the two guanines at the two ends of a G-tract reflects how they are grouped into the two terminal G-quartets. Those with higher cleavage were in the same one G-quartet, and those with lower cleavage were in the G-quartet at the other terminal. Thus, from the differential cleavage we can easily determine the orientation of each of the G-tracts. On the basis of this, the sequential orientation of the four

dx.doi.org/10.1021/ja108972e |J. Am. Chem. Soc. 2011, 133, 1475–1483

G-tracts for this G-quadruplex is derived as ↓↑↓↓, starting from the 5′ end of the DNA strand, which is truly what is adopted by the hybrid (3 + 1) Form 2 structure.

We further analyzed the G-quadruplex prepared in 150 mM K^+ solution in the presence of 40% (w/v) PEG 200 that has often been used to mimic the intracellular molecular crowding condition. As judged from its CD spectra (Supporting Information Figure S1) and our previous studies,[48,49] the sequence formed a parallel G-quadruplex. The result for this parallel G-quadruplex (Figure 2) shows that the guanines G1, G4, G7, and G10 in the 5′ terminal G-quartet had a much higher cleavage than the G3, G6, G9, and G12 in the opposite 3′ terminal G-quartet. From this cleavage polarity, a parallel ↓↓↓↓ orientation can be derived that is fully in agreement with the parallel structure adopted by the DNA.

In Na^+ solution, the $T_2AG_3(T_2AG_3)_3T_2A$ showed a CD signature featuring a positive peak near 296 nm and negative peak near 250 nm (Supporting Information Figure S1) similar to that of the $T_2AG_3(T_2AG_3)_3T_2$ and $AG_3(T_2AG_3)_3$[50] that are known to form a basket-type G-quadruplex.[51] This implies that the sequence we used also assumed a basket-type structure. Its cleavage profile showed an antiparallel ↓↑↓↑ orientation that complies with this structure (Figure 2). For the G-quadruplex formed in 150 mM Na^+ and 40% (w/v) PEG 200 solution, no information is available yet from the literature, and its cleavage showed that it had a same G-tract orientation as the G-quadruplex made in Na^+-only solution (Figure 2). This does not necessarily imply that they have the same structure because the ↓↑↓↑ orientation is compatible to either the basket or the chair topology, both of which belong to the antiparallel type.

Analysis of C-KIT G-Quadruplexes. The success of telomere G-quadruplexes encouraged us to extend the analysis to single-stranded human genomic sequence whose G-quadruplex structure has already been resolved.[52] In Figure 3, the wild sequence from the C-KIT2 gene WT and its variant G21T with a G to T substitution at the last guanine was used with three flanking nucleotides at both ends. The flanking nucleotides were included to better precipitate and separate cleavage fragments on gel electrophoresis, but did not alter the structure of the corresponding G-quadruplex as examined by CD spectroscopy (Supporting Information Figure S2). These two sequences were shown to both form parallel G-quadruplex in K^+ solution.[52] While the C-KIT2-WT has only three guanines in its first three G-tracts, the last G-tract has four. In this case, the G-quadruplex formation may involve either the first or the last three guanines in the last G-tract. To identify the guanines participated in G-quartet formation, they were also subjected to DMS footprinting.[53] A much better protection to the first three guanines in the last G-tract in K^+ versus Li^+ solution indicates a dominant participation of the first three guanines (Figure 3A). In agreement with previous reports,[52] a parallel orientation ↓↓↓↓ is derived from their photocleavage profiles for both the C-KIT2-WT and C-KIT2-G21T (Figure 3A,B). In a recent study, the C-KIT2-G21T has been shown to form intramolecular G-quadruplex in 20 mM K^+ solution and dimeric intermolecular G-quadruplex in 100 mM K^+ solution.[54] At the concentration we used, 87% of the C-KIT2-WT and 93% of the C-KIT2-G21T formed intramolecular G-quadruplex as examined by native gel electrophoresis (Supporting Information Figure S3). The reported dimeric G-quadruplex might be related to the high DNA concentration used in the NMR analysis.

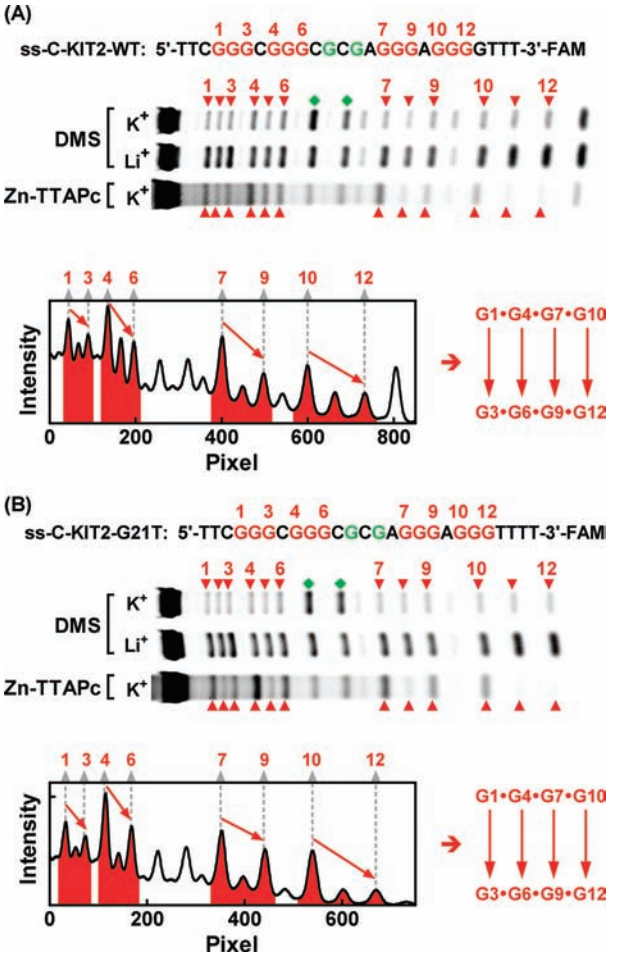


Figure 3. Photocleavage footprinting of C-KIT2 G-quadruplexes prepared in 150 mM K^+ solution. The guanines hosted in G-quartets were identified by DMS footprinting, and the Zn-TTAPc photocleavage profiles were generated as described as in Figure 2. Both sequences were identified to form the same parallel ↓↓↓↓ G-quadruplex as previously reported.[52,54] Bands of guanines participated in G-quartet formation are indicated by red ▲, and those not in G-quartet are indicated by green ◆.

Analysis of G-Quadruplexes in Double-Stranded DNA. Next, we examined whether the photocleavage could be performed on double-stranded DNA (dsDNA) because G-quadruplex formed in dsDNA is more representative of such structure in genome. G-Quadruplex was generated in two dsDNAs that carried the G-rich sequence from the ILPR or NRAS gene in 150 mM K^+ solution containing 40% (w/v) PEG according to the heat denaturation/renaturation method we recently described.[39] For the ILPR sequence, the DMS footprinting showed that the guanines 1−16 participated in G-quartet formation resulting in a G-quadruplex of four stacked G-quartets (Figure 4A). The much higher cleavage at G1, G5, G9, and G13 and a much lower cleavage at G4, G8, G12, and G16 suggested that the ILPR sequence formed parallel G-quadruplex (Figure 4A). For the NRAS sequence, the G-quartet formation might involve either G4-G6 or G5-G7 in the second G-tract from the 5′ end because it has four consecutive guanines, while the other three G-tracts had only three (Figure 4B). In this case, G5 and G6 were always in G-quartet and protected from DMS, but G4 and G7 would compete for G-quartet formation, and thus could only be partially

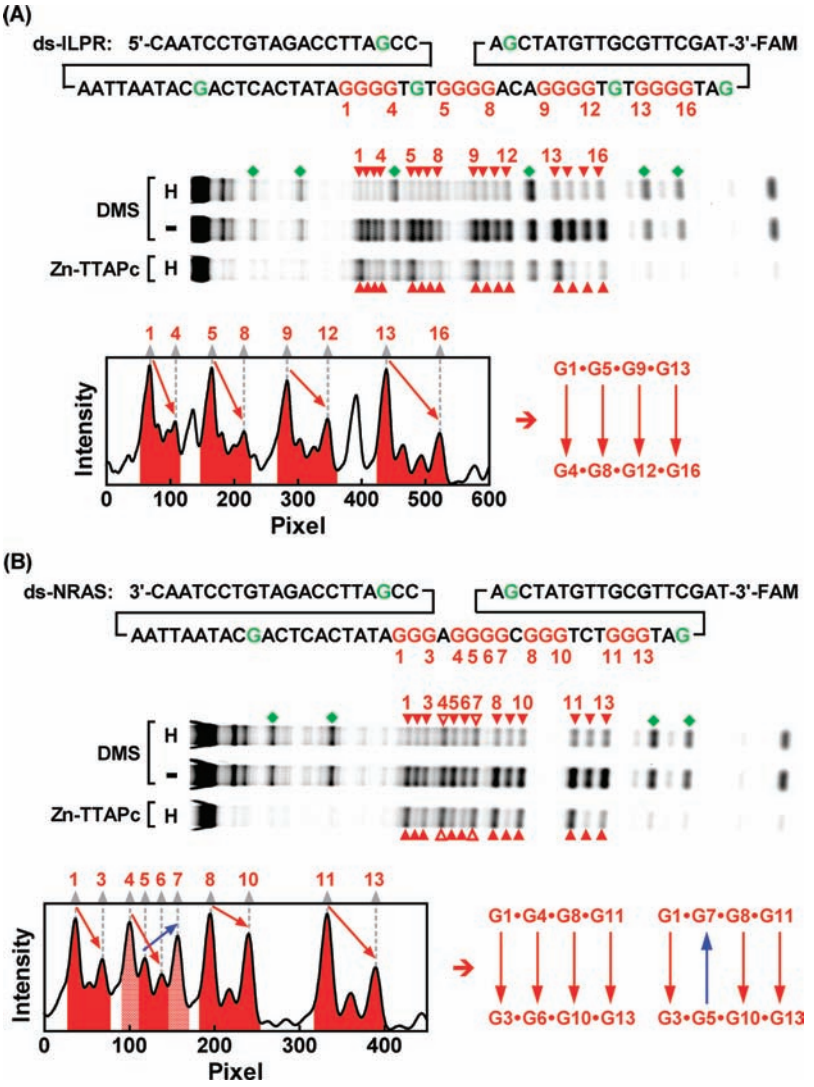


Figure 4. Photocleavage footprinting of (A) ILPR and (B) NRAS G-quadruplex in dsDNA. DNA in solution of 150 mM K^+, 40% (w/v) PEG 200 was subjected to heat denaturation/renaturation (H) to generate G-quadruplex. DNA without heating (−) did not form G-quadruplex. The guanines in G-quartet were identified by DMS footprinting (red ▲). Green "◆" indicate bands of cleavages at guanines not in G-quartet. The G-quadruplex in ILPR formed a parallel ⇓⇓⇓⇓ structure. The G-quadruplex in NRAS might adopt a parallel ⇓⇓⇓⇓ structure using the first three guanines or a (3 + 1) ⇓⇑⇓⇓ structure using the last three guanines in the second four-guanine G-tract.

protected. The weaker protection on G4 and G7 (△) as compared to that of G5 and G6 supports the presence of two structural isoforms in the NRAS: one involved the first three guanines (G4−G6), and the other involved the last three guanines (G5−G7) in the second G-tract (Figure 4B). The photocleavage resulted in a polarization corresponding to a parallel ⇓⇓⇓⇓ orientation for the former isoform and a (3 + 1) ⇓⇑⇓⇓ orientation for the latter.

We also carried out photocleavage analysis on G-quadruplex generated in RNA transcription in long dsDNA. Under this condition, a 78-bp dsDNA carrying a G-quadruplex sequence from the C-MYC gene was used. The G-rich core sequence was on the nontemplate strand and flanked by a promoter sequence for the T7 RNA polymerase at its 5′ side.[39] G-Quadruplex formation in the dsDNA was induced by transcription with T7 RNA polymerase in 150 mM K^+ and 40% PEG 200 solution.[39] Because the C-MYC sequence used here has four consecutive guanines in the second and fourth G-tracts from the 5′ end but the other two G-tracts had only three guanines, the G-quartet formation might involve either the first or the last three guanines in them. In Figure 5, the DMS footprinting suggested that the last guanine in the second G-tract did not participate in G-quadruplex formation induced by heat denaturation/renaturation because it was equally cleaved as the same guanine in the unheated DNA. However, transcription-induced G-quadruplex formation could involve either the first or the last three guanines in the second G-tract because the guanines in this G-tract were all protected (Figure 5B, top graph). For the fourth G-tract, it seems that G-quadruplex formation mainly involved the first three guanines that were similarly protected in the heated or transcribed samples. A weak protection could be seen on the last guanine. Because other guanines outside of the G-quadruplex region also showed slight variation in their cleavage, it is difficult to judge if this guanine participated in G-quadruplex formation or not. However, the chance was small if it did because its protection is minimal. To this end, the C-MYC sequence formed in RNA

dx.doi.org/10.1021/ja108972e |J. Am. Chem. Soc. 2011, 133, 1475–1483

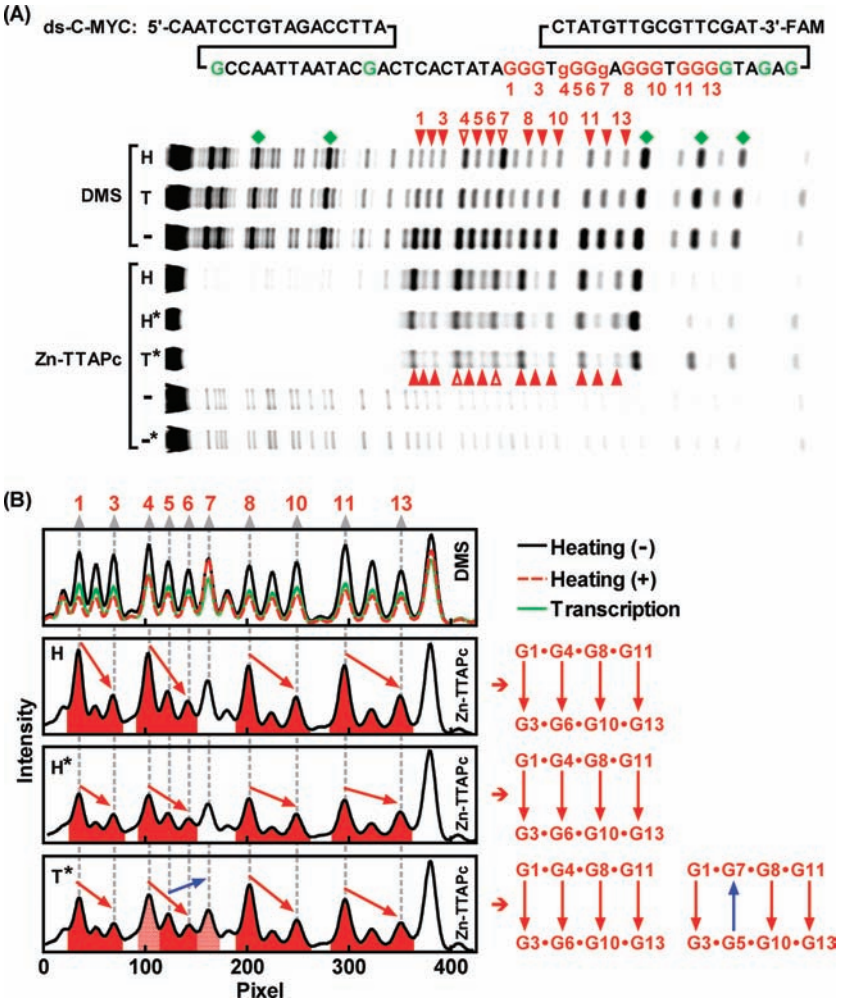


Figure 5. Photocleavage footprinting of C-MYC G-quadruplex in dsDNA prepared in 150 mM K^+, 40% (w/v) PEG 200 solution. DNA was subjected to heat denaturation/renaturation (H, H*) or transcription (T, T*), respectively, to generate G-quadruplex. DNA without heating or transcription (−) did not form G-quadruplex. The guanines in G-quartet were identified by DMS footprinting (red ▲). Green "◆" indicate bands of cleavages at guanines not in G-quartet. The DNA in one heated (H*) sample and the transcribed (T*) sample were purified by immobilization to streptavidin beads before photocleavage. The results reveal a parallel ⇓⇓⇓⇓ G-quadruplex that involved the first three guanines in the second four-guanine G-tract in both the heated and the transcribed samples. This structure is identical to the one reported of the single-stranded C-MYC sequence.[55,56] A (3 + 1) ⇓⇑⇓⇓ G-quadruplex might also be present that involved the last three guanines in the second four-guanine G-tract.

transcription at least two structural isoforms: one involved the first three guanines, and the other involved the last three guanines in the second G-tract. The cleavage pattern for the isoform involving the first three guanines in the second G-tract in both the heated (H) and the transcribed (T*) DNAs corresponds to a parallel ⇓⇓⇓⇓ orientation that is exactly what was reported for the single-stranded MYC22-G14T/G23T[55] and C-MYC-2345.[56] The cleavage profile of the isoform involving the last three guanines in the second G-tract in the transcribed (T*) DNA corresponds to a (3 + 1) ⇓⇑⇓⇓ orientation that has not been reported for the C-MYC.

When the G-quadruplex was prepared via RNA transcription, a clear footprinting was difficult to obtain probably because of the interference from the polymerase and other reagents (e.g., the reducing agent DTT) present in the reaction. Therefore, the DNA was immobilized onto Sepharose beads via biotin−streptavidin interaction to remove the interference by washing before photocleavage. The comparison between immobilized (H*) and nonimmobilized (H) samples prepared by heat denaturation/renaturation indicates that the immobilization did not alter the result (Figure 5).

Detection of G-Quadruplex Formation in Long Single- and Double-Stranded DNA. In Figure 5, it can be noticed that the guanines experienced much higher cleavage in the G-quadruplex region than those in the other regions or in the dsDNA where no G-quadruplex was generated. This fact is a clear manifestation of the Zn-TTAPc's high selectivity to G-quadruplex over other structural forms.[40,41,45] This property was then explored to detect G-quadruplex formation in both single-stranded DNA (ssDNA) and double-stranded DNA (dsDNA). C-MYC G-quadruplex was generated in a ssDNA in the absence and presence of PEG 200, respectively, or in a dsDNA in the presence of PEG 200 by heat denaturation/renaturation in K^+ solution. The dsDNA without heat treatment was used as G-quadruplex negative control. In Figure 6, it can be seen that the guanines in the G-quadruplexes in both ssDNA and dsDNA showed much higher cleavage as compared to those either in the rest parts of the DNA or in the control dsDNA. This sharp difference in

dx.doi.org/10.1021/ja108972e |J. Am. Chem. Soc. 2011, 133, 1475–1483

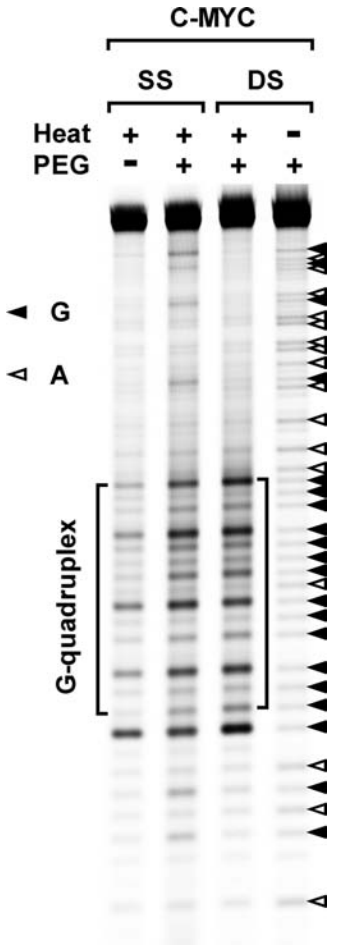


Figure 6. Detection of G-quadruplex formation in long single-stranded (SS) and double-stranded (DS) DNA by photocleavage. C-MYC DNA (see Figure 5 for sequence) was prepared in 150 mM K^+ solution in the presence or absence of 40% (w/v) PEG 200, then subjected to heat denaturation/renaturation. The unheated C-MYC dsDNA served as a G-quadruplex-negative control. The square bracket indicates the region of G-quadruplex formation; "▲" and "△" indicate the band of cleavage at guanine and adenine, respectively.

cleavage, therefore, can be used to detect G-quadruplex formation in long ssDNA and dsDNA.

DISCUSSION

We have shown that, in the several G-quadruplexes whose structure has so far been resolved by either X-ray crystallography or NMR, the Zn-TTAPc can induce differential photocleavage to the guanines between the two terminal G-quartets in the G-quartet core, in agreement with the strand folding orientations of the G-quadruplexes. The cleavage polarity allows us to determine how the guanines at the two ends of each of the four G-tracts in a G-quadruplex are grouped into the two terminal G-quartets. As a result, the folding orientation of an intramolecular G-quadruplex can be easily derived by connecting them in a 5′ to 3′ direction (Figure 7). The Zn-TTAPc has overall affinity similar to that of the different G-quadruplexes adopted by the human telomere DNA (Supporting Information Figure S4). This implies that it is unlikely to selectively detect a specific type of G-quadruplex, but rather the major form of structure if more than one structure is present.

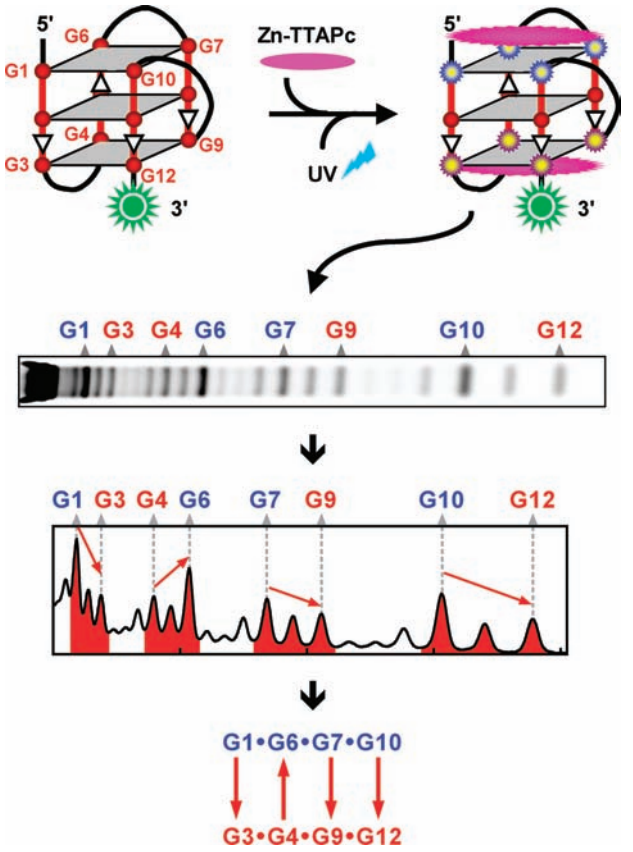


Figure 7. Scheme of detection of strand folding orientation in G-quadruplex by Zn-TTAPc (exemplified by telomeric DNA in K^+ solution). Zn-TTAPC binds to G-quadruplex presumably by stacking externally to the two terminal G-quartets of a G-quadruplex and cleaves the guanines in the G-quartets under UV irradiation. Because of the difference in local structure and interaction, the guanines in the two G-quartets are differentially cleaved. The cleavage fragments are separated by gel electrophoresis and visualized by a proper label (e.g., fluorescence or radioisotope) at the 3′ strand end. The guanines within each terminal G-quartet can be identified by the relative cleavage of the two guanines at the two ends of each G-tract. The strand folding orientation is obtained by drawing sequential connections through the two guanines in each G-tract in the 5′ to 3′ direction.

For the parallel human telomere G-quadruplex, the guanines in the two terminal G-quartets were more intensively cleaved than those in the middle G-quartet (Figure 2), strongly supporting the proposed specific stacking interaction of the ligand to the external face of the two terminal G-quartets (Figure 1D).[40] This cleavage polarity, with higher cleavage at the 5′ terminal G-quartet and lower at the 3′, should be related to the differences in the local structure near the two terminal G-quartets, which may result in different binding affinity or/and cleavage activity for the Zn-TTAPc. In the other types of G-quadruplexes studied, the two terminal G-quartets may still serve as two major binding sites for the Zn-TTAPc, because the cleavage to the guanines in the two terminal G-quartets dominated over the guanines in the middle G-quartet. This may establish the structural basis for the analysis. The only exception is the G2 in the basket-type human telomere G-quadruplex, which had higher cleavage than its two neighboring G1 and G3 residuals (Figure 2B). We are unable to think of a reason for this phenomenon. Certain factors, such as the presence of lateral loops or possible asymmetric binding of

Zn-TTAPc to the G-quadruplex, may produce various interferences to the cleavage of individual guanines in the terminal G-quartets. Therefore, the success of the photocleavage technique relies on the difference in cleavage between the two terminal G-quartets produced by the stacking binding. Larger difference will be more tolerable to such interferences. It has been reported that phthalocyanines bind to G-quadruplex DNA with at least 50- and 500-fold higher affinities than to tRNA and calf thymus DNA, respectively.[40] This high selectivity should greatly disfavor the binding to nonspecific sites, thus reducing interferences from these sources.

The Zn-TTAPc-induced photocleavage is expected to be a useful tool for the analysis of G-quadruplex structures. We have successfully used it to determine the G-quadruplex conformation that human telomere DNA initially assumes when liberated from DNA duplex in a kinetic process.[57] A unique advantage of the photocleavage method is that it can work, in addition to single-stranded nucleic acids, with long double-stranded nucleic acids. This is especially useful because all G-quadruplex sequences in genome, except the telomere overhang, are accommodated in double-stranded DNA. Unlike the CD spectroscopy that only tells whether the G-tracts are parallel or antiparallel, the photocleavage reveals the orientation of each individual G-tract. When working with single-stranded nucleic acids, its requirement for submicromolar sample concentration dramatically reduces the risk of intermolecular G-quadruplex formation that might occur in other physical methods, such as NMR. Because of its high selectivity on G-quadruplexes over non-G-quadruplex structures,[40,41,45] the Zn-TTAPc produces much stronger cleavage to the guanines in the G-quadruplex than in the other regions (Figure 6). Thus, the photocleavage can also be used to detect G-quadruplex formation in both long single- and double-stranded DNA.

Some technical issues should be considered for the photocleavage. The fluorescent dye had to be labeled at the 3′ end of the nucleic acids to obtain clear bands in gel electrophoresis. When 5′ end labeling was used, the bands became blurred. We suspect that certain chemical modification occurred at the 5′ side of the cleaving point. A few nucleotides should be present at the 5′ side of the first guanine to separate the fragment from the intact input nucleic acid. To obtain efficient precipitation of fragments produced by cleavages near the 3′ end, 3 or 4 nucleotides should be added to the 3′ end of the G-quadruplex core sequence.

To maximize cleavage polarity, low ligand concentration should be used. We observed that the polarity was best seen at low Zn-TTAPc concentration equivalent to that of the G-quadruplex being analyzed. At high ligand concentration, ligand may saturate binding sites even though their affinities may be different. In certain cases, for instance, in the K^+/PEG solution, addition of competitive DNA a few minutes before UV irradiation could improve the cleavage polarity. We anticipate that the weaker binding could be more rapidly dissociated by the competition; therefore, the differential binding between the two G-quartets might be enlarged.

Because the phthalocyanines may induce G-quadruplex structure transition and formation,[41] their incubation with nucleic acids and the UV irradiation that followed should be carried out at low temperature to ensure the original G-quadruplex structure is not altered. According to our experiments, the original structure of the telomere G-quadruplexes was well maintained at 4 °C as examined by CD spectroscopy (Supporting Information Figure S1).

For sequences whose guanine participation in G-quartet formation is not known for sure, it is recommended to use the photocleavage in combination with the DMS footprinting to identify the guanines participated in G-quartet formation.[53] The orientation of each individual G-tract in the G-quadruplex can then be revealed by photocleavage.

■ ASSOCIATED CONTENT

Supporting Information. Additional figures. This material is available free of charge via the Internet at http://pubs.acs.org

■ AUTHOR INFORMATION

Corresponding Author
z.tan@ioz.ac.cn

■ ACKNOWLEDGMENT

This work was supported by grants 2010CB945300 and 2007CB507402 from MSTC and grants 90813031, 30970617, and 20921062 from NSFC.

■ REFERENCES

(1) Lipps, H. J.; Rhodes, D. *Trends Cell Biol.* **2009**, *19*, 414.
(2) Catasti, P.; Chen, X.; Mariappan, S. V.; Bradbury, E. M.; Gupta, G. *Genetica* **1999**, *106*, 15.
(3) Rawal, P.; Kummarasetti, V. B.; Ravindran, J.; Kumar, N.; Halder, K.; Sharma, R.; Mukerji, M.; Das, S. K.; Chowdhury, S. *Genome Res.* **2006**, *16*, 644.
(4) Du, Z.; Zhao, Y.; Li, N. *Nucleic Acids Res.* **2009**, *37*, 6784.
(5) Hershman, S. G.; Chen, Q.; Lee, J. Y.; Kozak, M. L.; Yue, P.; Wang, L. S.; Johnson, F. B. *Nucleic Acids Res.* **2008**, *36*, 144.
(6) Verma, A.; Halder, K.; Halder, R.; Yadav, V. K.; Rawal, P.; Thakur, R. K.; Mohd, F.; Sharma, A.; Chowdhury, S. *J. Med. Chem.* **2008**, *51*, 5641.
(7) Yadav, V. K.; Abraham, J. K.; Mani, P.; Kulshrestha, R.; Chowdhury, S. *Nucleic Acids Res.* **2008**, *36*, D381.
(8) Zhang, R.; Lin, Y.; Zhang, C. T. *Nucleic Acids Res.* **2008**, *36*, D372.
(9) Huppert, J. L.; Balasubramanian, S. *Nucleic Acids Res.* **2007**, *35*, 406.
(10) Du, Z.; Zhao, Y.; Li, N. *Genome Res.* **2008**, *18*, 233.
(11) Zhao, Y.; Du, Z.; Li, N. *FEBS Lett.* **2007**, *581*, 1951.
(12) Du, Z.; Kong, P.; Gao, Y.; Li, N. *Biochem. Biophys. Res. Commun.* **2007**, *354*, 1067.
(13) Maizels, N. *Nat. Struct. Mol. Biol.* **2006**, *13*, 1055.
(14) Eddy, J.; Maizels, N. *Nucleic Acids Res.* **2006**, *34*, 3887.
(15) Zahler, A. M.; Williamson, J. R.; Cech, T. R.; Prescott, D. M. *Nature* **1991**, *350*, 718.
(16) Sun, D.; Thompson, B.; Cathers, B. E.; Salazar, M.; Kerwin, S. M.; Trent, J. O.; Jenkins, T. C.; Neidle, S.; Hurley, L. H. *J. Med. Chem.* **1997**, *40*, 2113.
(17) Shammas, M. A.; Shmookler Reis, R. J.; Akiyama, M.; Koley, H.; Chauhan, D.; Hideshima, T.; Goyal, R. K.; Hurley, L. H.; Anderson, K. C.; Munshi, N. C. *Mol. Cancer Ther.* **2003**, *2*, 825.
(18) Riou, J. F.; Guittat, L.; Mailliet, P.; Laoui, A.; Renou, E.; Petitgenet, O.; Megnin-Chanet, F.; Helene, C.; Mergny, J. L. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2672.
(19) Ou, T. M.; Lu, Y. J.; Tan, J. H.; Huang, Z. S.; Wong, K. Y.; Gu, L. Q. *ChemMedChem* **2008**, *3*, 690.
(20) Neidle, S. *FEBS J.* **2010**, *277*, 1118.
(21) Luedtke, N. W. *Chimia* **2009**, *63*, 134.
(22) Balasubramanian, S.; Neidle, S. *Curr. Opin. Chem. Biol.* **2009**, *13*, 345.

(23) Wong, H. M.; Payet, L.; Huppert, J. L. *Curr. Opin. Mol. Ther.* **2009**, *11*, 146.
(24) Huppert, J. L. *Chem. Soc. Rev.* **2008**, *37*, 1375.
(25) Neidle, S. *Curr. Opin. Struct. Biol.* **2009**, *19*, 239.
(26) Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S. *Nucleic Acids Res.* **2006**, *34*, 5402.
(27) Simonsson, T. *Biol. Chem.* **2001**, *382*, 621.
(28) Baumann, P. *G-Quadruplex DNA: Methods and Protocols*; Humana Press: Totowa, NJ, 2009.
(29) Paramasivan, S.; Rujan, I.; Bolton, P. H. *Methods* **2007**, *43*, 324.
(30) Webba da Silva, M. *Methods* **2007**, *43*, 264.
(31) Campbell, N. H.; Parkinson, G. N. *Methods* **2007**, *43*, 252.
(32) Onyshchenko, M. I.; Gaynutdinov, T. I.; Englund, E. A.; Appella, D. H.; Neumann, R. D.; Panyutin, I. G. *Nucleic Acids Res.* **2009**, *37*, 7570.
(33) Gaynutdinov, T. I.; Brown, P.; Neumann, R. D.; Panyutin, I. G. *Biochemistry* **2009**, *48*, 11169.
(34) He, Y.; Neumann, R. D.; Panyutin, I. G. *Nucleic Acids Res.* **2004**, *32*, 5359.
(35) Jing, N.; Rando, R. F.; Pommier, Y.; Hogan, M. E. *Biochemistry* **1997**, *36*, 12498.
(36) Dapic, V.; Bates, P. J.; Trent, J. O.; Rodger, A.; Thomas, S. D.; Miller, D. M. *Biochemistry* **2002**, *41*, 3676.
(37) Dapic, V.; Abdomerovic, V.; Marrington, R.; Peberdy, J.; Rodger, A.; Trent, J. O.; Bates, P. J. *Nucleic Acids Res.* **2003**, *31*, 2097.
(38) Zhang, L.; Huang, J.; Ren, L.; Bai, M.; Wu, L.; Zhai, B.; Zhou, X. *Bioorg. Med. Chem.* **2008**, *16*, 303.
(39) Zheng, K. W.; Chen, Z.; Hao, Y. H.; Tan, Z. *Nucleic Acids Res.* **2010**, *38*, 327.
(40) Alzeer, J.; Luedtke, N. W. *Biochemistry* **2010**, *49*, 4339.
(41) Ren, L.; Zhang, A.; Huang, J.; Wang, P.; Weng, X.; Zhang, L.; Liang, F.; Tan, Z.; Zhou, X. *ChemBioChem* **2007**, *8*, 775.
(42) Alzeer, J.; Vummidi, B. R.; Roth, P. J.; Luedtke, N. W. *Angew. Chem., Int. Ed.* **2009**, *48*, 9362.
(43) Zhang, A. M.; Huang, J.; Weng, X. C.; Li, J. X.; Ren, L. G.; Song, Z. B.; Xiong, X. Q.; Zhou, X.; Cao, X. P.; Zhou, Y. *Chem. Biodiversity* **2007**, *4*, 215.
(44) Gantchev, T. G.; Gowans, B. J.; Hunting, D. J.; Wagner, J. R.; van Lier, J. E. *Int. J. Radiat. Biol.* **1994**, *66*, 705.
(45) Wang, P.; Ren, L.; He, H.; Liang, F.; Zhou, X.; Tan, Z. *ChemBioChem* **2006**, *7*, 1155.
(46) Dai, J.; Carver, M.; Punchihewa, C.; Jones, R. A.; Yang, D. *Nucleic Acids Res.* **2007**, *35*, 4927.
(47) Phan, A. T.; Luu, K. N.; Patel, D. J. *Nucleic Acids Res.* **2006**, *34*, 5715.
(48) Xue, Y.; Kan, Z. Y.; Wang, Q.; Yao, Y.; Liu, J.; Hao, Y. H.; Tan, Z. *J. Am. Chem. Soc.* **2007**, *129*, 11185.
(49) Kan, Z. Y.; Lin, Y.; Wang, F.; Zhuang, X. Y.; Zhao, Y.; Pang, D. W.; Hao, Y. H.; Tan, Z. *Nucleic Acids Res.* **2007**, *35*, 3646.
(50) Ambrus, A.; Chen, D.; Dai, J.; Bialis, T.; Jones, R. A.; Yang, D. *Nucleic Acids Res.* **2006**, *34*, 2723.
(51) Wang, Y.; Patel, D. J. *Structure* **1993**, *1*, 263.
(52) Hsu, S. T.; Varnai, P.; Bugaut, A.; Reszka, A. P.; Neidle, S.; Balasubramanian, S. *J. Am. Chem. Soc.* **2009**, *131*, 13399.
(53) Sun, D.; Hurley, L. H. *Methods Mol. Biol.* **2010**, *608*, 65.
(54) Kuryavyi, V.; Phan, A. T.; Patel, D. J. *Nucleic Acids Res.* **2010**, *38*, 6757.
(55) Ambrus, A.; Chen, D.; Dai, J.; Jones, R. A.; Yang, D. *Biochemistry* **2005**, *44*, 2048.
(56) Phan, A. T.; Modi, Y. S.; Patel, D. J. *J. Am. Chem. Soc.* **2004**, *126*, 8710.
(57) Xue, Y.; Liu, J.-q.; Zheng, K.-w.; Kan, Z.-y.; Hao, Y.-h.; Tan, Z., under revision.