The Major G-Quadruplex Formed in the Human Platelet-Derived Growth Factor Receptor β Promoter Adopts a Novel Broken-Strand Structure in K + Solution

ARTICLE *in* JOURNAL OF THE AMERICAN CHEMICAL SOCIETY · AUGUST 2012

Impact Factor: 12.11 · DOI: 10.1021/ja305764d · Source: PubMed

JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Communication
pubs.acs.org/JACS

The Major G-Quadruplex Formed in the Human Platelet-Derived Growth Factor Receptor β Promoter Adopts a Novel Broken-Strand Structure in K^+ Solution

Yuwei Chen,[†,‡] Prashansa Agrawal,[†] Robert V. Brown,[†] Emmanuel Hatzakis,[†] Laurence Hurley,[†,‡,§,∥] and Danzhou Yang*[,†,‡,§,∥]

[†]Department of Pharmacology and Toxicology, College of Pharmacy, [‡]Department of Chemistry and Biochemistry, [§]BIO5 Institute, and [∥]The Arizona Cancer Center, University of Arizona, 1703 East Mabel Street, Tucson, Arizona 85721, United States

Supporting Information

ABSTRACT: Overexpression of platelet-derived growth factor receptor β (PDGFR-β) has been associated with cancers and vascular and fibrotic disorders. PDGFR-β has become an attractive target for the treatment of cancers and fibrotic disorders. DNA G-quadruplexes formed in the GC-rich nuclease hypersensitivity element of the human PDGFR-β gene promoter have been found to inhibit PDGFR-β transcriptional activity. Here we determined the major G-quadruplex formed in the PDGFR-β promoter. Instead of using four continuous runs with three or more guanines, this G-quadruplex adopts a novel folding with a broken G-strand to form a primarily parallel-stranded intramolecular structure with three 1 nucleotide (nt) double-chain-reversal loops and one additional lateral loop. The novel folding of the PDGFR-β promoter G-quadruplex emphasizes the robustness of parallel-stranded structural motifs with a 1 nt loop. Considering recent progress on G-quadruplexes formed in gene-promoter sequences, we suggest the 1 nt looped G_iNG_j motif may have been evolutionarily selected to serve as a stable foundation upon which the promoter G-quadruplexes can build. The novel folding of the PDGFR-β promoter G-quadruplex may be attractive for small-molecule drugs that specifically target this secondary structure and modulate PDGFR-β gene expression.

Platelet-derived growth factor receptor β (PDGFR-β) is a cell-surface-receptor tyrosine kinase that plays an essential role in cellular growth, proliferation, differentiation, and development.[1] Overexpression of PDGFR-β has been associated with a number of diseases, including different types of cancer and vascular and fibrotic disorders.[2] Enhanced PDGFR-β signaling has been shown to be involved in tumor growth, angiogenesis, invasion, and metastasis.[2] Inhibiting PDGFR-β can increase the efficacy of chemotherapy through a tumor-selective increase in drug uptake.[3a] Therefore, inhibiting PDGFR-β signaling recently has become an attractive pursuit for anticancer therapeutics[3] and the treatment of fibrotic disorders.[4]

The promoter of the human PDGFR-β gene has previously been characterized, and a highly GC-rich proximal region (base pairs −165 to −132) was found to be critical for basal PDGFR-β promoter activity.[5] This GC-rich promoter region is a nuclease hypersensitivity element (NHE) and can assume non-B-form secondary structures in supercoiled plasmid.[5] The G-rich strand in this NHE can form G-quadruplexes and is targeted by the G-quadruplex-interactive molecule telomestatin to inhibit the PDGFR-β transcriptional activity in both plasmid transfection experiments and human cancer cells.[5] Thus, the G-quadruplex formed in the human PDGFR-β promoter could be a potential target for drug development in the treatment of cancer and other diseases. The 38-mer G-rich strand (PDGF-Pu38) of the PDGFR-β promoter NHE contains seven runs of guanines (one with seven, one with four, three with three, and two with two guanines; see Figure 1A) and has the potential to form multiple G-quadruplexes by using different combinations of G-tracts. Four overlapping intramolecular G-quadruplex structures, 5′End, Mid5′, Mid3′, and 3′End, form in PDGF-Pu38 in K^+-containing solution, as shown by polymerase stop assay.[5] The Mid5′ G-quadruplex has been shown to be specifically stabilized by telomestatin, which can inhibit the PDGFR-β transcriptional activity in human cancer cells.[5] The Mid5′ G-quadruplex is formed in the region containing the II−V G-runs of the PDGFR-β NHE (Figure 1A). Dimethyl sulfate (DMS) footprinting studies can determine the formation of G-quadruplexes under physiological conditions, as N7 of guanine in a G-tetrad is protected from DMS methylation and subsequent cleavage, in contrast to duplex or single-stranded DNA. The region of the PDGFR-β NHE that forms the Mid5′ G-quadruplex (Pu20, Figure 1A) contains one run of seven guanines, one run of two guanines, one run of four guanines, and one run of three guanines and is still capable of forming different G-quadruplex structures. Based on the previously known G-quadruplex folding rules,[6] one would predict that G3−G5 and G7−G9 of run II, run IV of four guanines (G14−G17), and run V of three guanines (G20−G22) would be involved in the formation of a three-tetrad G-quadruplex. G6 of run II and G17 of run IV were cleaved as expected in the DMS Footprinting (Figure 1B). However, it is surprising to find that G20 and G21 of the three-guanine run V were clearly cleaved, whereas G-to-A mutations at G11 and G12 destabilized the G-quadruplex (Figure 1B). Thus, it appears that a novel folding is adopted in the PDGFR-β Mid5′ G-quadruplex.

NMR mutational analysis is very useful in deconvoluting the major G-quadruplex formation in a given sequence. The wild-

Received: June 13, 2012
Published: August 6, 2012

   dx.doi.org/10.1021/ja305764d | J. Am. Chem. Soc. 2012, 134, 13220−13223

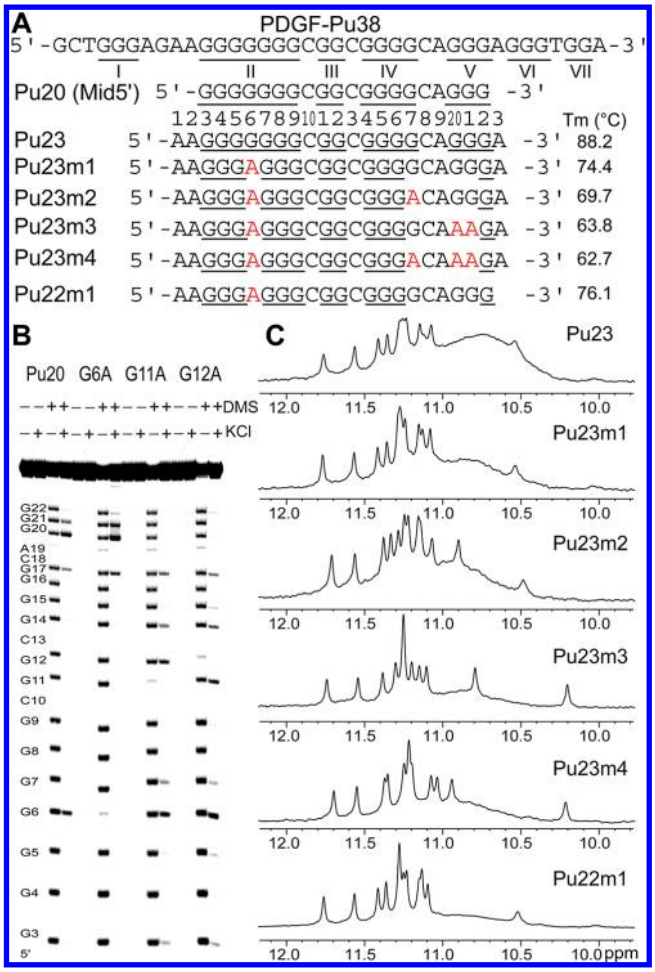

Figure 1. (A) Promoter sequence of the PDGFR-β gene and its modifications. Top: the 38-mer wt G-rich sequence. The seven G-runs with two or more guanines are underlined and numbered. The Mid5′ G-quadruplex is formed in the region containing G-runs II, III, IV, and V (Pu20). Pu23 is the 23-mer wt sequence. Pu23m1, Pu23m2, Pu23m3, and Pu23m4 are modified Pu23 sequences with G-to-A mutations shown in red. The CD melting temperatures of Pu23 sequences in 100 mM K^+ are shown. Pu22m1 is the same as Pu23m1 but without the 3′-A. (B) DMS footprinting of the wt Pu20 (sequence shown) and three modified Pu20 sequences containing a single G-to-A mutation (top). (C) Imino regions of the 1D ^{1}H NMR spectra of PDGFR-β Mid5′ promoter sequences at 25 °C in 25 mM potassium phosphate, 70 mM KCl, pH 7.0.

type (wt) Pu23 sequence was selected to determine the formation of the Mid5′ G-quadruplex, as Pu23 contains at least one flanking residue at each end of Pu20. Based on the DMS footprinting data, we prepared Pu23 sequences with specific G-to-A mutations at G6, G17, G20, and G21 (Figure 1A) and tested them by 1D ^{1}H NMR analysis in 100 mM, pH 7 K^+ solution. Signals for imino protons at 10−12 ppm indicate the formation of G-quadruplex structures (Figure 1C).[7] Wild-type Pu23 shows the formation of a major G-quadruplex with sharp peaks and the presence of a broad envelope, likely due to the formation of multiple higher-order structures. The Pu23m1 sequence with the single G-to-A mutation at G6 shows the formation of the same major G-quadruplex as in Pu23, indicating that G6 is not involved in the G-tetrad formation. The formation of higher-order structures was much reduced in Pu23m1, as shown by the markedly lower intensity of the broad envelope in the NMR spectrum. The significantly higher circular dichroism (CD) melting temperature of Pu23 relative to Pu23m1 also indicated the formation of higher-order structures in Pu23 (Figure 1A). This result indicates that the presence of seven continuous guanines in run II of Pu23 significantly increases the chance for the formation of higher-order G-quadruplex structures. Pu23m2 containing an additional G-to-A mutation at G17 forms the same G-quadruplex as Pu23m1, indicating that G17 is not involved in the G-tetrad formation. Similarly, the NMR spectra of Pu23m3 and Pu23m4 with G-to-A mutations at G20 and G21, respectively, show that G20 and G21 are not involved in the G-tetrad formation. Thus, our NMR results are in complete agreement with the DMS footprinting data for the PDGFR-β Mid5′ G-quadruplex, showing that G6, G17, G20, and G21 are not involved in the G-tetrad formation. The CD spectra of the wt and mutant Pu23 sequences are almost identical (Figure 2A), supporting the conclusion that the same G-quadruplex is formed in these Pu23 sequences.

The 22-mer Pu22m1 without the 3′-A forms the same major G-quadruplex as Pu23m1 with even further reduced formation of higher-order structures in K^+ solution, so it was chosen for NMR structural analysis of the major G-quadruplex formed in the PDGFR-β promoter NHE. This G-quadruplex appears to be of unimolecular nature, as indicated by the concentration-independent CD and NMR melting temperatures. The well-resolved imino proton peaks between 10.5 and 12 ppm indicate the formation of a stable G-quadruplex structure in Pu22m1. Site-specific low-concentration (6%) incorporation of ^{15}N-labeled guanine nucleoside was prepared for each guanine position of Pu22m1.[7] The guanine imino H1 proton has one-bond coupling to N1, and the aromatic H8 proton has two-bond coupling to N7 (Figure 2B); thus, the H1 and H8 protons of the site-specific labeled guanine can readily be detected by 1D ^{15}N-edited experiments. Using such experiments, the imino H1 and aromatic H8 protons of guanine residues are unambiguously assigned (Figure 2C,D). Pu22m1 contains 15 guanines; however, only 12 imino peaks are observed in the 10.5−12 ppm region, indicating this intramolecular G-quadruplex formed in Pu22m1 contains three G-tetrads. In agreement with NMR mutational analysis and DMS footprinting, it was found that G17, G20, and G21 are not involved in G-tetrad formation, as their imino protons are not detected in the 10.5−12 ppm region.

Assignment of the imino H1 and base H8 protons of the guanines led to the direct determination of the folding topology of this major G-quadruplex formed in the PDGFR-β promoter. In a G-tetrad plane with the Hoogsteen H-bond network, the imino proton H1 of each guanine is in close spatial vicinity to the H1's of the two adjacent guanines and to the base H8 of one of the adjacent guanines (Figure 2B). The through-space nuclear Overhauser effect (NOE) connectivities of guanine H1−H1 and H1−H8 thus determine the arrangement and topology of a G-tetrad plane (Figure 2E). For example, the G3H1/G7H1, G11H1/G14H1, and G14H1/G3H1 NOE interactions (Figure 2F) and the G7H8/G3H1, G11H8/G7H1, G14H8/G11H1, and G3H8/G14H1 NOEs (Figure 2G) define a G-tetrad plane of G3−G7−G11−G14 (Figure 2E). Similarly, the G8H8/G4H1, G12H8/G8H1, G15H8/G12H1, and G4H8/G15H1 NOEs (Figure 2G) define a G-tetrad plane of G4−G8−G12−G15 (Figure 2E). Unexpectedly, the third G-tetrad plane is shown to be G5−G9−***G22***−G16 (Figure 2E) based on the G5H1/G9H1 and G9H1/G22H1 NOEs (Figure 2F) and the G9H8/G5H1, G22H8/G9H1, G16H8/G22H1, and G5H8/G16H1 NOEs (Figure 2G). It is interesting to note that, in this third G-tetrad plane, three guanines are continuous from the same G-runs of the

dx.doi.org/10.1021/ja305764d | *J. Am. Chem. Soc.* 2012, 134, 13220−13223

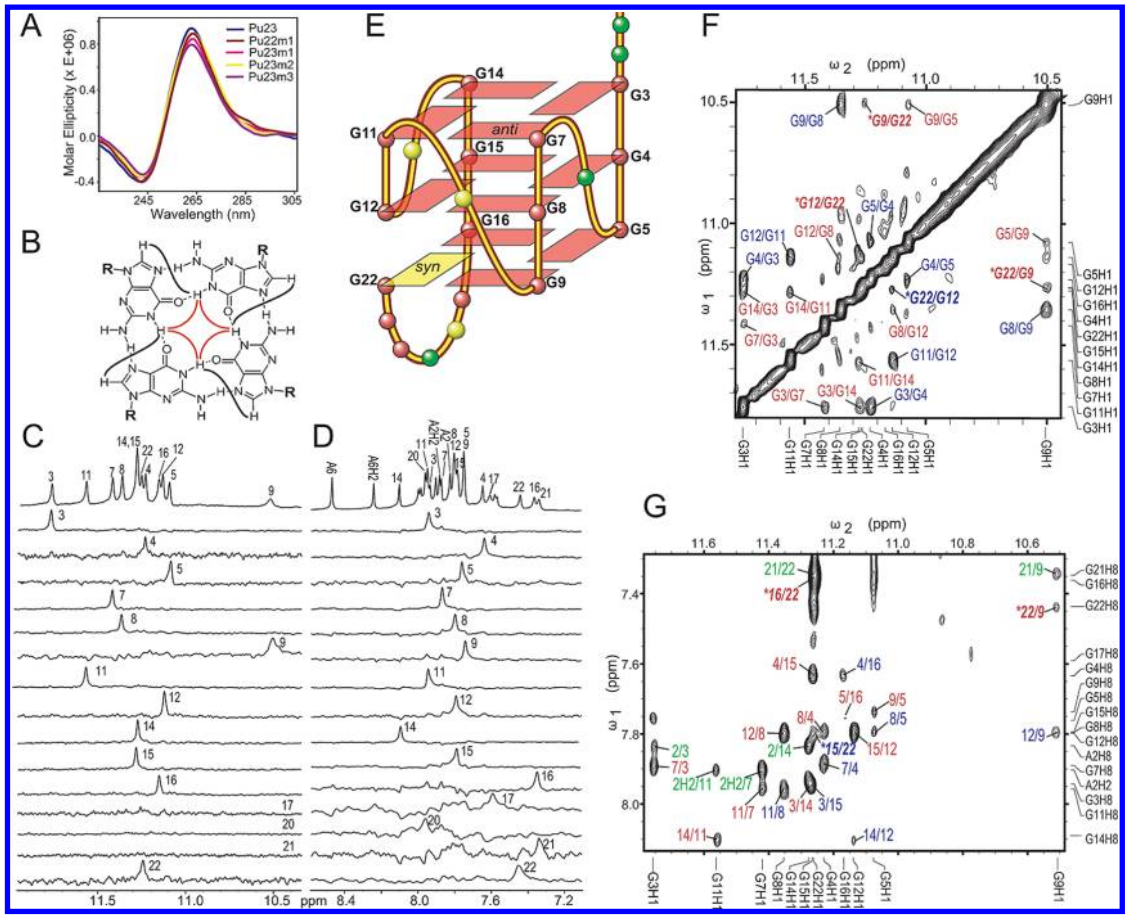

Figure 2. (A) CD spectra of Pu23 sequences in 100 mM K^+. (B) G-tetrad with H1−H1 and H1−H8 connectivity patterns detectable in NOESY experiments. (C) Imino and (D) aromatic H8 proton assignments of Pu22m1 from 1D ^{15}N-filtered experiments using site-specifically labeled oligonucleotides. (E) Schematic drawing of the folding topology of the G-quadruplex formed in Pu22m1 (G = red, A = green, C = yellow; red box = *anti*-G, yellow box = *syn*-G). (F) H1−H1 and (G) H1−H8 regions of the 2D NOESY spectrum of Pu22m1 in H_2O at 25 °C. Intra-tetrad NOEs are shown in red, inter-tetrad NOEs in blue, and NOEs with flanking bases in green.

previous two G-tetrads, but G22 is discontinued from the G11−G12 strand (Figure 2E). Therefore, whereas three G-strands are continuous, there is one broken strand, G11−G12···G22, in this PDGFR-β G-quadruplex. This novel G-quadruplex folding is further supported by inter-tetrad NOEs. G3H8/G15H1, G7H8/G4H1, G11H8/G8H1, and G14H8/G12H1 (Figure 2G) connect the top and middle G-tetrad planes (Figure 2E) and reflect the right-handed twist of the DNA backbone. Similar NOE connectivities are observed between the middle and bottom planes, including G4H8/G16H1, G8H8/G5H1, G12H8/G9H1, and specifically G15H8/G22H1 (Figure 2G), which stacks G22 to the G4−G8−G12−G15 plane. In addition, similar to the sequential H1/H1 connectivities observed along a G-strand, such as G3H1/G4H1 and G4H1/G5H1 observed for the continuous G3−G4−G5 strand (Figure 2F), the G22H1/G12H1 NOE is clearly observed in addition to G11H1/G12H1 (Figure 2F) connecting the broken strand G11−G12···G22. Our NMR results thus show that the major PDGFR-β G-quadruplex formed by Pu22m1 is a primarily parallel-stranded intramolecular structure with the broken G-strand G11−G12···G22. Because of this broken strand, in addition to three double-chain-reversal loops typically seen in a parallel-stranded structure (i.e., A6, C10, and C13), the Pu22m1 G-quadruplex contains a fourth loop, G17−C18−A19−G20−G21, which is a lateral loop (Figure 2E). Connected by a lateral loop, G22 is expected to have a different glycosidic conformation than G11 and G12. Indeed, as indicated by the much stronger intraresidue H8−H1′ NOE of G22 (Figure S2), G22 appears to be in a syn conformation, whereas all of the other G-tetrads are in anti conformations. The primarily parallel-stranded folding of this G-quadruplex is consistent with the CD spectra of Pu23 sequences, which exhibit a positive maximum around 265 nm and a negative minimum at 240 nm (Figure 2A), characteristic of parallel-stranded G-quadruplex structures.[8] It is noted that these characteristic CD spectral features of parallel-stranded G-quadruplexes are based on the guanine stacking conformations rather than the G-strand directionalities; in the case of the Mid5′ PDGFR-β G-quadruplex, although G22 has a different strand directionality, the guanine stacking is the same as in an all-parallel-stranded G-quadruplex.

DNA G-quadruplex secondary structures formed in the promoter regions of human proto-oncogenes were recently found to be transcriptional regulators and have emerged as a new class of cancer-specific molecular targets for anticancer drugs.[9] The c-MYC promoter is the first and most extensively studied system for promoter G-quadruplex formation. The major G-quadruplex Myc2345 formed in the c-MYC promoter is a parallel-stranded structure that has been shown to be very stable, containing two G_3NG_3 parallel-stranded motifs with a 1 nucleotide (nt) loop (N) that appears to be highly favorable for the double-chain-reversal loop conformation[10] (Figure 3). In addition, all of the loop isomers formed in the c-MYC promoter

dx.doi.org/10.1021/ja305764d | J. Am. Chem. Soc. 2012, 134, 13220−13223

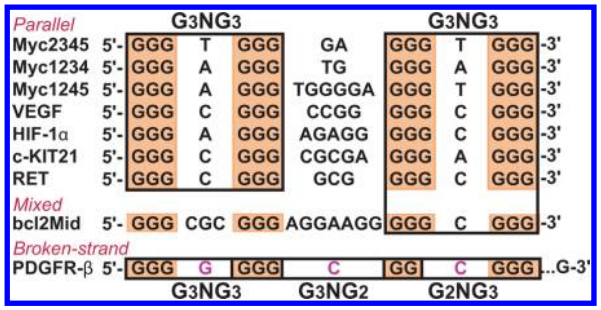

Figure 3. G-quadruplex-forming promoter sequences.

also adopt parallel-stranded structures with G_3NG_3 motifs.[8,11] The list of G-quadruplex structures formed in the promoter regions of human proto-oncogenes has been growing.[9] Parallel-stranded structures are found to be more common in the promoter G-quadruplexes; and, intriguingly, there is a wide prevalence of the G_3NG_3 (N is a loop residue) sequence motif that forms a robust parallel-stranded structural motif with a 1 nt loop (Figure 3).[9] Interestingly, the same G_3NG_3 motif has also been found in the major G-quadruplex formed in the bcl-2 promoter (Figure 3), although it adopts a hybrid-type mixed parallel/antiparallel structure that appears to be largely determined by the first 3 nt loop, which adopts a lateral loop conformation.[12] It is quite significant that the PDGFR-β Pu22m1 adopts a parallel-stranded G-quadruplex with a broken G-strand, despite the presence of four three-guanine runs in this sequence (i.e., G3−G5, G7−G9, G14−G17, and G20−G22). Run II containing the seven guanines G3−G9 in the wt PDGF sequence (Figure 1A) could be considered as a G_3NG_3 motif with a 1 nt G-loop (i.e., N = G) (Figure 3). However, if Pu22m1 had used the four continuous G-runs with three guanines [i.e., G3−G5, G7−G9, G14−G16 (or G15−G17), and G20−G22], the resulting G-quadruplex would have three loops with sizes of 1, 4, and 3 (or 1, 5, and 2), which do not favor the parallel-stranded structure.[9,12] Instead, by having one broken G-strand and using G11−G12 and G14−G16, the Pu22m1 G-quadruplex contains three 1 nt loops, G6, C10, and C13, which all favor the parallel-stranded structural motif (Figure 2E). As a consequence, G20 and G21 of the G20−G22 G-run are not involved in the G-tetrad formation (Figure 2E). Importantly, in the DMS footprinting of the full-length wt PDGF promoter sequence Pu38 (Figure S1), G20 and G21 were clearly cleaved by DMS, suggesting that the novel broken-strand Mid5′ G-quadruplex appears to be the major structure formed in the PDGFR-β NHE promoter sequence. The formation of the novel broken-strand G-quadruplex by Pu22m1 again emphasizes the robust parallel structural motif with a 1 nt loop connecting the two adjacent parallel G-strands, which can now be extended to G_iNG_j ($i = j = 3$ in G_3NG_3). The major PDGFR-β G-quadruplex formed by Pu22m1 can be expressed as $G_3NG_3NG_2NG_3\cdots G$, a primarily parallel-stranded intramolecular structure with a broken G-strand (Figure 3). Another broken-strand G-quadruplex was observed previously in the human c-KIT promoter, an all-parallel-stranded structure containing two 1 nt strand-reversal loops that can be expressed as $(G_3NG_3NG\cdots G_3\cdots G_2)$.[13] The fact that the 1 nt parallel-stranded G_iNG_j motif is so common in quadruplex-forming gene-promoter sequences may suggest that the robust 1 nt loop parallel-stranded motif may have been evolutionarily selected to serve as a stable foundation that could be the first folding event upon which the promoter G-quadruplexes can then build.

In summary, our results show that the major intramolecular G-quadruplex formed in the promoter region of the human PDGFR-β gene adopts a novel folding structure. By having a broken G-strand, this PDGFR-β promoter G-quadruplex is a primarily parallel-stranded G-quadruplex with three 1 nt double-chain-reversal loops and one additional lateral loop. This emphasizes the robustness of the parallel-stranded structural motif with 1 nt loop, the G_iNG_j motif. Together with recent progress on G-quadruplex structures formed in gene-promoter sequences, these results suggest that the 1 nt loop G_iNG_j motif may have been evolutionarily selected to serve as a stable foundation upon which the promoter G-quadruplexes can build. The novel folding of the major G-quadruplex formed in the PDGFR-β promoter may be attractive in the design of small-molecule drugs that specifically target this secondary structure and modulate PDGFR-β gene expression.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Methods, DMS footprinting, NOESY spectrum, and complete refs [3b−d]. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

yang@pharmacy.arizona.edu

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Dr. Megan Carver for proofreading the paper. This research was supported by the National Institutes of Health (S10RR16659 to D.Y., CA153821 to L.H. and D.Y.).

■ REFERENCES

(1) Heldin, C. H.; Westermark, B. *Physiol. Rev.* **1999**, *79*, 1283.

(2) (a) Andrae, J.; Gallini, R.; Betsholtz, C. *Genes Dev.* **2008**, *22*, 1276. (b) Alvarez, R. H.; Kantarjian, H. M.; Cortes, J. E. *Mayo Clin. Proc.* **2006**, *81*, 1241.

(3) (a) Ostman, A. *Cytokine Growth Factor Rev.* **2004**, *15*, 275. (b) Shen, J.; et al. *Neoplasia* **2009**, *11*, 594. (c) Coluccia, A. M. L.; et al. *Blood* **2008**, *112*, 1346. (d) Cao, R. H.; et al. *Cancer Cell* **2004**, *6*, 333.

(4) Borkham-Kamphorst, E.; Herrmann, J.; Stoll, D.; Treptau, J.; Gressner, A. M.; Weiskirchen, R. *Lab. Invest.* **2004**, *84*, 766.

(5) Qin, Y.; Fortin, J. S.; Tye, D.; Gleason-Guzman, M.; Brooks, T. A.; Hurley, L. H. *Biochemistry* **2010**, *49*, 4208.

(6) Huppert, J. L.; Balasubramanian, S. *Nucleic Acids Res.* **2005**, *33*, 2908.

(7) Zhang, Z.; Dai, J.; Veliath, E.; Jones, R. A.; Yang, D. Z. *Nucleic Acids Res.* **2010**, *38*, 1009.

(8) Hatzakis, E.; Okamoto, K.; Yang, D. Z. *Biochemistry* **2010**, *49*, 9152.

(9) Yang, D.; Okamoto, K. *Future Med. Chem.* **2010**, *2*, 619.

(10) Ambrus, A.; Chen, D.; Dai, J.; Jones, R. A.; Yang, D. Z. *Biochemistry* **2005**, *44*, 2048.

(11) Mathad, R. I.; Hatzakis, E.; Dai, J.; Yang, D. *Nucleic Acids Res.* **2011**, *39*, 9023.

(12) Dai, J.; Dexheimer, T. S.; Chen, D.; Carver, M.; Ambrus, A.; Jones, R. A.; Yang, D. Z. *J. Am. Chem. Soc.* **2006**, *128*, 1096.

(13) Phan, A. T.; Kuryavyi, V.; Burge, S.; Neidle, S.; Patel, D. J. *J. Am. Chem. Soc.* **2007**, *129*, 4386.