Published in final edited form as:

J Am Chem Soc. 2012 August 15; 134(32): 13220–13223. doi:10.1021/ja305764d.

The Major G-Quadruplex Formed in the Human Platelet-Derived Growth Factor Receptor β (PDGFR- 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 University of Arizona, 1703 E. Mabel St, Tucson, AZ 85721

[†]Department of Chemistry University of Arizona, 1703 E. Mabel St, Tucson, AZ 85721

[§]BIO5 Institute University of Arizona, 1703 E. Mabel St, Tucson, AZ 85721

^{||}The Arizona Cancer Center, University of Arizona, 1703 E. Mabel St, Tucson, AZ 85721

Abstract

Overexpression of PDGFR- β has been associated with cancers, vascular and fibrotic disorders. PDGFR- β has become an attractive target for the treatment of cancers and fibrotic disorders. DNA G-quadruplexes formed in GC-rich nuclease hypersensitivity element (NHE) of the human PDGFR- β gene promoter has been found to inhibit PDGFR- β transcriptional activity. Here we determined the major G-quadruplex formed in the PDGFR- β promoter. Instead of using the four continuous runs with three or more guanines, this G-quadruplex adopt a novel folding with a broken G-strand, to form a primarily parallel-stranded intramolecular structure with three 1-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. Together with recent progress on G-quadruplexes formed in gene promoter sequences, the 1-nt-looped G_1NG_j motif may be evolutionarily selected to serve as a stable foundation for the promoter G-quadruplexes to build upon. The novel folding of the PDGFR- β promoter G-quadruplex may represent an attractive target for small molecule drugs that specifically target this secondary structure and modulate PDGFR- β gene expression.

The platelet-derived growth factor receptor β (PDGFR- β) is a cell surface receptor tyrosine kinase that plays an essential role in cell growth, cell proliferation, cellular differentiation, and development¹. Overexpression of PDGFR- β has been associated with a number of diseases, including different types of cancer, vascular disorders, and fibrotic disorders^{2, 3}. Enhanced PDGFR- β signaling has been shown to be involved in tumor growth, angiogenesis, invasion, and metastasis^{2, 3}. Inhibiting PDGFR- β can increase the efficacy of chemotherapy through a tumor-selective increase in drug uptake⁴. Therefore, inhibiting PDGFR- β signaling recently has become an attractive pursuit for anticancer therapeutics^{4–7} as well as for the treatment of fibrotic disorders⁸.

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

*Corresponding Author. yang@pharmacy.arizona.edu.

Supporting Information Available. Methods, DMS footprinting of the full-length wild-type PDGF-Pu38, and h8/h6-h1' region of the exchangeable 2D-NOESY of Pu22m1. This material is available free of charge via the Internet at <http://pubs.acs.org>.

PDGFR- β promoter activity⁹. This GC-rich promoter region is a nuclease hypersensitive element (NHE) and can assume non-B-form secondary structures in supercoiled plasmid⁹. The G-rich strand in this NHE can form G-quadruplexes and can be targeted by the G-quadruplex-interactive molecule telomestatin to inhibit the PDGFR- β transcriptional activity in both the plasmid transfection experiments and human cancer cells⁹. 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, with one run of seven guanines, one run of four guanines, three runs of three guanines, and two runs of two guanines (Figure 1A), and has the potential to form multiple G-quadruplexes by using different combinations of G-tracts. Four overlapping intramolecular G-quadruplex structures, namely, 5'End, Mid5', Mid3', and 3'End, have been shown to form in PDGF-Pu38 in K⁺-containing solution by polymerase stop assay⁹. The Mid5' G-quadruplex has been shown to be specifically stabilized by telomestatin which can inhibit the PDGFR- β transcriptional activity in human cancer cells⁹. 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 study can determine the formation of G-quadruplex 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 (Figure 1A), and is still capable of forming different G-quadruplex structures. Based on the previously known G-quadruplex folding rules¹¹, 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) (Figure 1A) would be involved in the formation of a three-tetrad G-quadruplex. However, while G6 was cleaved as expected in the DMS footprinting, surprisingly, together with the G17 of the 4-guanine run IV, G20 and G21 of the 3-guanine run V were clearly cleaved, while the G-to-A mutation at G20 or G21 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 deconvolution of the major G-quadruplex formation in a given sequence. The wild-type 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 (Figure 1A). 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 ¹H NMR in pH 7, 100 mM K⁺ solution. The imino protons at 10–12 ppm indicate the formation of G-quadruplex structures (Figure 1C)¹². The wild-type Pu23 (Figure 1A) showed the formation of a major G-quadruplex with sharp peaks, and the presence of a broad envelope (Figure 1C), likely due to the formation of multiple higher-order structures. The Pu23m1 sequence (Figure 1A) with the single G-to-A mutation at G6 showed the formation of the same major G-quadruplex as in Pu23 (Figure 1C), indicating 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 level of the broad envelope in NMR spectrum (Figure 1C). A significantly higher CD melting temperature of Pu23 than Pu23m1 also indicated the formation of higher-order structures in Pu23 (Figure 1A). This result indicates that the presence of seven continuous guanines of run II in Pu23 significantly increase the chance of the formation of higher-order G-quadruplex structures. Pu23m2, which contained an additional G-to-A mutation at G17, formed the same G-quadruplex as that in Pu23m1 (Figure 1C), indicating G17 is not involved in the G-tetrad formation. Similarly, G20 and G21 were shown to not to be involved in the G-tetrad formation by NMR spectra of Pu23m3 and Pu23m4 with G-to-A mutations at G20 and G21 (Figure 1C). Thus the NMR result was in complete agreement with the DMS footprinting data of the PDGFR- β Mid5' G-quadruplex (Figure 1B), that G6, G17, G20, and G21 are not involved in

the G-tetrad formation. The CD spectra of the wild-type and mutant Pu23 sequences are almost identical (Figure 2A), supporting that the same G-quadruplex is formed in these Pu23 sequences.

The 22-mer Pu22m1 (Figure 1A) without the 3'-A forms the same major G-quadruplex as that in Pu23m1 with even reduced formation of higher-order structures in K^+ solution (Figure 1C), and thus was chosen for NMR structural analysis of the major G-quadruplex formed in the PDGFR- β promoter NRE. This G-quadruplex appears to be of unimolecular nature as indicated by the concentration-independent melting temperature measured by CD and NMR. The well-resolved imino proton peaks located between 10.5–12 ppm indicated 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 Pu22m1¹². The guanine imino H1 proton has one-bond coupling to N1, and the aromatic H8 proton has two-bond coupling to N7 (Figure 2B), hence the H1 and H8 protons of the site-specific labeled guanine are readily detected by 1D ^{15}N -edited experiments. Using such experiments, the imino H1 and aromatic H8 protons of guanine residues were unambiguously assigned, as shown in Figure 2C and 3D. Pu22m1 contains fifteen guanines; however, only twelve imino peaks are observed in the 10.5–12 ppm region (Figure 2C), indicating this intramolecular G-quadruplex formed in Pu22m1 contains three G-tetrads. In agreement with NMR mutational analysis and DMS footprinting, G17, G20, and G21 were not involved in G-tetrad formation, as their imino protons were not detected in the 10.5–12 ppm region.

The assignment of the imino H1 and base H8 protons of guanines led to the direct determination of the folding topology of this major G-quadruplex formed in the PDGFR- β promoter (Figure 2E). 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 H1s of the two adjacent guanines, and to the base H8 of one of the adjacent guanines (Figure 2B). The through-space 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 NOE interactions between G3H1/G7H1, G11H1/G14H1, G14H1/G3H1 (Figure 2F), and G7H8/G3H1, G11H8/G7H1, G14H8/G11H1, and G3H8/G14H1 (Figure 2G) defined a G-tetrad plane of G3-G7-G11-G14 (Figure 2E). Similarly, the NOE interactions between G8H8/G4H1, G12H8/G8H1, G15H8/G12H1, and G4H8/G15H1 (Figure 2G) defined a G-tetrad plane of G4-G8-G12-G15 (Figure 2E). Unexpectedly, the third G-tetrad plane was shown to be G5-G9-G22-G16 (Figure 2E), based on the NOE interactions between G5H1/G9H1, G9H1/G22H1 (Figure 2F), and G9H8/G5H1, G22H8/G9H1, G16H8/G22H1, and G5H8/G16H1 (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 previous two G-tetrads, but G22 was discontinued from the G11-G12 strand (Figure 2E). Therefore, whereas three G-strands are continuous, there is one broken-strand, namely, G11-G12---G22 in this PDGFR- β G-quadruplex. This novel G-quadruplex folding is further supported by inter-tetrad NOEs. Inter-tetrad NOEs, such as 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. The same kind of NOE connectivities was observed between the middle and bottom planes (Figure 2E), such as 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 (Figure 2F) observed for continuous G3-G4-G5 strand, the NOE of G22H1/G12H1 was clearly observed in addition to G11H1/G12H1 (Figure 2F) connecting the broken strand G11-G12---G22. Our NMR results thus showed that the major PDGFR- β G-quadruplex formed by Pu22m1 is a primarily parallel-stranded, intramolecular structure with a broken G-strand of G11-G12---G22 (Figure 2E). 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, i.e., 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 from that of G11 and G12. Indeed, as indicated by the much stronger intra-residue H8-H1' NOE intensity of G22 (Figure S2), G22 appears to be in *syn* conformation whereas all other tetrad-guanines are in *anti* conformation. The primarily parallel-stranded folding of this G-quadruplex is consistent with the CD spectra of Pu23 sequences, with a positive maximum around 265 nm and a negative minimum at 240 nm (Figure 2A), characteristic of the parallel-stranded G-quadruplex structures¹³. It is noted that these characteristic CD spectral features of the 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.

Recently, DNA G-quadruplex secondary structures formed in the promoter regions of human proto-oncogenes were found to be transcriptional regulators and have emerged as a new class of cancer-specific molecular targets for anticancer drugs¹⁰. c-MYC promoter is the first and most extensively studied system for the promoter G-quadruplex formation. The major G-quadruplex formed in the c-MYC promoter is a parallel-stranded structure, which was shown to be very stable, containing two G₃NG₃ parallel-stranded motifs with a 1-nt loop (N) which appears to be highly favorable for the double-chain-reversal loop conformation¹⁴ (Figure 3). In addition, all the loop isomers formed in the c-MYC promoter also adopt parallel-stranded structures with G₃NG₃ motifs^{13, 15}. Since then, there has been a growing list of G-quadruplex structures formed in the promoter regions of human proto-oncogenes¹⁰. Intriguingly, parallel-stranded structures are found to be more common in the promoter G-quadruplexes, and there is a wide prevalence of the G₃NG₃ (N is a loop residue) sequence motif, which forms a robust parallel-stranded structural motif with a 1-nt loop (Figure 3)¹⁰. Interestingly, the same G₃NG₃ motif has also been found in the major G-quadruplex formed in the bcl-2 promoter (Figure 3), although the bcl-2 G-quadruplex adopts a hybrid-type mixed parallel/antiparallel structure which appears to be largely determined by the first 3-nt loop which adopts a lateral loop conformation¹⁶. It is quite significant that 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, G20-G22. The run II of seven guanines G3-G9 of the wild-type PDGF sequence (Figure 1A) could be considered as a G₃NG₃ motif with a 1-nt G-loop (N=G in this case) (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 resulted G-quadruplex would have three loops with sizes of 1, 4, 3, (or 1, 5, 2), which do not favor the parallel-stranded structure^{10, 16}. Instead, by having one broken G-strand and using G11-G12 and G14-G16, the Pu22m1 G-quadruplex now contains three 1-nt loops, i.e., G6, C10, and C13, which all favor the parallel-stranded structural motifs (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 wild-type 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 G₃NG₃). In the case of the major PDGFR- β G-quadruplex formed by Pu22m1, it can be expressed as G₃NG₃NG₂NG₃...G, a primarily parallel-stranded intramolecular structure with a broken G-strand (Figure 3). Another broken-strand G-quadruplex has been previously observed in the human c-KIT promoter, an all-parallel-stranded structure containing two 1-nt strand-reversal loops that can be expressed as (G₃NG₃NG...G₃...G₂)¹⁷. In fact, the 1-nt parallel-stranded

G_iNG_j motif is so common in the quadruplex-forming gene promoter sequences that this may suggest that the robust 1-nt-looped parallel-stranded motif may be evolutionarily selected to serve as a stable foundation which could be the first folding event for the promoter G-quadruplexes to build upon.

In summary, our results showed 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, emphasizing the robustness of parallel-stranded structural motifs with 1-nt loop, namely, the G_iNG_j motif. Together with recent progress on G-quadruplex structures formed in gene promoter sequences, the 1-nt-looped G_iNG_j motif may be evolutionarily selected to serve as a stable foundation for the promoter G-quadruplexes to build upon. The novel folding of the major G-quadruplex formed in the PDGFR- β promoter may represent an attractive target for small molecule drugs that specifically target this secondary structure and modulate PDGFR- β gene expression.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr. Megan Carver for proofreading the paper. This research is supported by the National Institutes of Health (S10RR16659 to DY, CA153821 to LH, DY).

REFERENCES

- Heldin CH, Westermark B. Physiological Reviews. 1999; 79:1283. [PubMed: 10508235]
- Andrae J, Gallini R, Betsholtz C. Genes & Development. 2008; 22:1276. [PubMed: 18483217]
- Alvarez RH, Kantarjian HM, Cortes JE. Mayo Clinic Proceedings. 2006; 81:1241. [PubMed: 16970222]
- Ostman A. Cytokine & Growth Factor Reviews. 2004; 15:275. [PubMed: 15207817]
- Shen J, Vil MD, Prewett M, Damoci C, Zhang H, Li H, Jimenez X, Deevi DS, Iacolina M, Kayas A, Bassi R, Persaud K, Rohoza-Asandi A, Balderes P, Loizos N, Ludwig DL, Tonra J, Witte L, Zhu Z. Neoplasia. 2009; 11:594. [PubMed: 19484148]
- Coluccia AML, Cirulli T, Neri P, Mangieri D, Colanardi MC, Gnani A, Di Renzo N, Dammacco F, Tassone P, Ribatti D, Gambacorti-Passerini C, Vacca A. Blood. 2008; 112:1346. [PubMed: 18524994]
- Cao RH, Bjorndahl MA, Religa P, Clasper S, Garvin S, Galter D, Meister B, Ikomi F, Tritsarlis K, Dissing S, Ohhashi T, Jackson DG, Cao YH. Cancer Cell. 2004; 6:333. [PubMed: 15488757]
- Borkham-Kamphorst E, Herrmann J, Stoll D, Treptau J, Gressner AM, Weiskirchen R. Laboratory Investigation. 2004; 84:766. [PubMed: 15077122]
- Qin Y, Fortin JS, Tye D, Gleason-Guzman M, Brooks TA, Hurley LH. Biochemistry. 2010; 49:4208. [PubMed: 20377208]
- Yang D, Okamoto K. Future Med. Chem. 2010; 2:619. [PubMed: 20563318]
- Huppert JL, Balasubramanian S. Nuc. Acids Res. 2005; 33:2908.
- Zhang Z, Dai J, Veliath E, Jones RA, Yang DZ. Nuc. Acids Res. 2010; 38:1009.
- Hatzakis E, Okamoto K, Yang DZ. Biochemistry. 2010; 49:9152. [PubMed: 20849082]
- Ambrus A, Chen D, Dai J, Jones RA, Yang DZ. Biochemistry. 2005; 44:2048. [PubMed: 15697230]
- Mathad RI, Hatzakis E, Dai J, Yang D. Nuc. Acids Res. 2011; 39:9023.

16. Dai J, Dexheimer TS, Chen D, Carver M, Ambrus A, Jones RA, Yang DZ. *J. Am. Chem. Soc.* 2006; 128:1096. [PubMed: 16433524]
17. Phan AT, Kuryavyi V, Burge S, Neidle S, Patel DJ. *J. Am. Chem. Soc.* 2007; 129:4386. [PubMed: 17362008]

Figure 1.

(A) The promoter sequence of the PDGFR- β gene and its modifications. The top sequence is the 38mer wild-type 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 the II, III, IV, V G-runs (Pu20). Pu23 is the 23mer wild-type 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 wild-type Pu20 (sequence shown), and three modified Pu20 sequences containing a single G-to-A mutation shown at the top. (C) Imino regions of 1D ¹H NMR spectra of PDGFR- β Mid5' promoter sequences, at 25 °C, 25 mM K-phosphate, 70 mM KCl, pH 7.0.

Figure 2.

(A) CD spectra of Pu23 sequences in 100 mM K⁺. (B) A G-tetrad with H1-H1 and H1-H8 connectivity pattern detectable in NOESY experiments. Imino (C) and aromatic H8 (D) proton assignments of Pu22m1 using 1D ¹⁵N-filtered experiments using site-specific 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). H1-H1 region (F) and H1-H8 region (G) of 2D-NOESY spectrum of Pu22m1 in H₂O at 25 °C. Intra-tetrad NOEs are shown in red, inter-tetrad NOEs are shown in blue, and NOEs with flanking bases are shown in green.

Figure 3.
G-quadruplex-forming promoter sequences.