Published in final edited form as:

Biochim Biophys Acta Gen Subj. 2018 April ; 1862(4): 846–854. doi:10.1016/j.bbagen.2017.12.011.

The 3'-end region of the human PDGFR- β core promoter nuclease hypersensitive element forms a mixture of two unique end-insertion G-quadruplexes

Buket Onel^a, Megan Carver^e, Prashansa Agrawal^d, Laurence H. Hurley^{e,f,g}, and Danzhou Yang^{a,b,c,*}

^aPurdue University, College of Pharmacy, Medicinal Chemistry and Molecular Pharmacology, 575 W Stadium Ave, West Lafayette, IN 47907, USA

^bPurdue University, Center for Cancer Research, 201 S University St, West Lafayette, IN 47906, USA

^cPurdue Institute for Drug Discovery, 720 Clinic Dr, West Lafayette, IN 47907, USA

^dCase Western Reserve University, Department of Chemistry, Cleveland, OH 44106, USA

^eUniversity of Arizona, College of Pharmacy, 1703 East Mabel Street, Tucson, AZ 85721, USA

^fBIO5 Institute, 1657 East Helen Street, Tucson, AZ 85721, USA

^gArizona Cancer Center, 1515 North Campbell Avenue, Tucson, AZ 85724, USA

Abstract

Background: While the most stable G-quadruplex formed in the human PDGFR- β promoter nuclease hypersensitive element (NHE) is the 5'-mid G-quadruplex, the 3'-end sequence that contains a 3'-GGA run forms a less stable G-quadruplex. Recently, the 3'-end G-quadruplex was found to be a transcriptional repressor and can be selectively targeted by a small molecule for PDGFR- β downregulation.

Method: We use 1D and 2D high-field NMR, in combination with Dimethylsulfate Footprinting, Circular Dichroism Spectroscopy, and Electrophoretic Mobility Shift Assay.

Results: We determine that the PDGFR- β extended 3'-end NHE sequence forms two novel end-insertion intramolecular G-quadruplexes that co-exist in equilibrium under physiological salt conditions. One G-quadruplex has a 3'-non-adjacent flanking guanine inserted into the 3'-external tetrad (3'-insertion-G4), and another has a 5'-non-adjacent flanking guanine inserted into the 5'-external tetrad (5'-insertion-G4). The two guanines in the GGA-run move up or down within the G-quadruplex to accommodate the inserted guanine. Each end-insertion G-quadruplex has a low

*Corresponding author at: Purdue University, College of Pharmacy, Medicinal Chemistry and Molecular Pharmacology, 575 W Stadium Ave, West Lafayette, IN 47907, USA. yangdz@purdue.edu (D. Yang).

Transparency document

The <http://dx.doi.org/10.1016/j.bbagen.2017.12.011> associated with this article can be found, in online version.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.bbagen.2017.12.011>.

thermal stability as compared to the 5'-mid G-quadruplex, but the selective stabilization of GSA1129 shifts the equilibrium toward the 3'-end G-quadruplex in the PDGFR- β NHE.

Conclusion: An equilibrium mixture of two unique end-insertion intramolecular G-quadruplexes forms in the PDGFR- β NHE 3'-end sequence that contains a GGA-run and non-adjacent guanines in both the 3'- and 5'-flanking segments; the novel end-insertion structures of the 3'-end G-quadruplex are selectively stabilized by GSA1129.

General significance: We show for the first time that an equilibrium mixture of two unusual end-insertion G-quadruplexes forms in a native promoter sequence and appears to be the molecular recognition for PDGFR- β downregulation.

Keywords

G-quadruplex; End-insertion G-quadruplexes; PDGFR- β gene downregulation; Drug target

1. Introduction

The platelet-derived growth factor receptor beta (PDGFR- β) gene encodes a cell surface tyrosine kinase receptor and is involved in a variety of crucial cellular events, such as proliferation, differentiation, survival, motility, and tissue repair in adults [1,2]. Mutation and abnormal expression of PDGFR- β contribute to several malignancies such as dermatofibrosarcoma [3], glioblastoma [4], endocrine pancreatic tumor [5], and medulloblastoma [6]. Therefore, the PDGFR- β signaling pathway is considered to be an attractive target for cancer therapeutics and has been tested in different animal models as well as in clinical trials [7–10]. Current PDGFR- β -targeted strategies inhibit the PDGFR- β signaling pathway by targeting the PDGF ligand or the PDGFR- β receptor protein [3]. Modulation of PDGFR- β at the transcriptional level provides an alternative approach to target the PDGFR- β signaling pathway for potential therapeutic development.

DNA G-quadruplexes have recently been found to form in the proximal location of promoters in a number of human proto-oncogenes under transcriptionally induced supercoiling and to act as transcriptional regulators [11–13], such as c-MYC [14,15], VEGF [16], HIF-1 α [17], BCL-2 [18,19], KRAS [20], Rb [21], c-KIT [22,23], RET [24], hTERT [25], as well as PDGFR- β [26]. A recent study from the Balasubramanian group showed that G4 formation in gene promoters is associated with regulatory, nucleosome-depleted chromatin and elevated transcription [27]. The core proximal promoter region of the human PDGFR- β contains a GC-rich nuclease hypersensitive element (NHE) (base pairs –166 to –132 with respect to the transcriptional start site) and has been shown to be an important regulatory element for PDGFR- β transcription activity [26]. The G-rich strand of the PDGFR- β NHE consists of seven poly-guanine tracts that can form a mixture of at least four G-quadruplexes from overlapping sequences, namely, the 3'-end (R1–R4), 3'-mid (R2–R6), 5'-mid (R3–R6), and 5'-end (R4–R7) (Fig. 1) [26]. The 5'-mid G-quadruplex is the most stable in the PDGFR- β NHE, forming a broken-strand structure in K⁺ solution [28]. Interestingly, the ligand-mediated stabilization of G-quadruplexes by Telomestatin, which binds more selectively to the 3'-end G-quadruplex, was shown to decrease PDGFR- β promoter activity [26].

Recently it was found through mutational analysis that the 3'-end G-quadruplex is a transcriptional repressor of PDGFR- β ; furthermore, a small molecule ellipticine analog, GSA1129, which selectively stabilizes the 3'-end G-quadruplex, was shown to downregulate PDGFR- β promoter activity, inhibit PDGFR- β -driven cell proliferation and migration in neuroblastoma cells, and attenuate endotoxin-mediated acute lung inflammation in an established preclinical murine model of acute lung injury associated with PDGFR- β overexpression [29]. The PDGFR- β NHE 3'-end minimal sequence consists of G-runs R1–R4; G-runs R2–R4 contain three, three, and four guanines, respectively, but the 3'-most purine-run R1 (GGA) contains only two guanines (Fig. 2A). Using the 3'-end minimal sequence Pu18 (Fig. 2A), the DMS footprinting showed clear protection of guanines in R2, R3, and R4, whereas R1 is only partially protected, suggesting that this end is partially frayed [29]. The 3'-most adenine A18 of R1 was thought to be utilized to form an imperfect 3'-external G:G:G:A tetrad (Fig. 2B) which leads to low stability of the 3'-end G-quadruplex [29].

Herein we report that, instead of the proposed G-quadruplex with a 3'-G:G:G:A tetrad, the wild-type extended PDGFR- β NHE 3'-end sequence contains non-adjacent guanines in both the 3'- and 5'-flanking segments and can form two novel intramolecular end-insertion G-quadruplexes, one with a non-adjacent 3'-flanking guanine inserted into the 3'-external tetrad (3'-insertion-G4), and another with a non-adjacent 5'-flanking guanine inserted into the 5'-external tetrad (5'-insertion-G4). The two end-insertion G-quadruplexes co-exist in equilibrium under physiological salt conditions. Each end-insertion G-quadruplex has a low thermal stability as compared to the 5'-mid G-quadruplex; however, they are selectively stabilized by GSA1129 binding, which shifts the equilibrium toward the 3'-end G-quadruplex in the PDGFR- β NHE. The dynamic equilibrium of the two intramolecular G-quadruplexes may provide a potential means for PDGFR- β transcriptional regulation and the novel end-insertion structures may make it an effective target for GSA1129 binding. This study highlights the complex nature of the PDGFR- β NHE sequence and the importance of identifying the proper sequence for biologically important G-quadruplex.

2. Methods

2.1. DNA sample preparation

Oligonucleotides were synthesized and purified as previously described [30,31]. DNA NMR samples were prepared in either 50 mM or 25 mM K⁺ solution at pH 7, to a final concentration of 0.1–2.5 mM DNA in a solution that contains 90/10% H₂O/D₂O. The DNA oligo-nucleotides were heated to 95 °C for 5 min and then cooled slowly to room temperature to allow G-quadruplex formation. Quantification of the DNA oligonucleotides was performed by UV/Vis spectroscopy at 260 nm using their calculated extinction coefficients.

2.2. Nuclear Magnetic Resonance (NMR) spectroscopy experiments

NMR experiments were conducted using Bruker DRX-600 spectrometer or AV-800 spectrometer with cryoprobe. The NMR experiments for samples in water were performed using Watergate suppression technique. 1D ¹H NMR spectra were recorded using Bruker

DRX-600 and 512 scans were collected at 25 °C. Both H1 and H8 protons of the site-specific ^{15}N -labeled guanine were readily detected by ^{15}N -edited HMQC experiments. The guanine H1 imino protons, one-bond coupled to N1, and H8 protons, two-bond coupled to N7, can be unambiguously assigned by 1D ^{15}N -edited HMQC experiments [32,33]. DNA oligonucleotides with site-specific 6% incorporation of ^{15}N -labeled guanine at each guanine position were prepared [30,31]; the 1D GE-JRSE HMQC experiments were used for measuring ^{15}N -edited spectra to identify labeled-guanine imino and H8 protons [32]. Nuclear Overhauser Effect Spectroscopy (NOESY) experiments were collected at 25 °C for complete proton resonance assignment in water using AV-800 spectrometer with cryoprobe. The contribution from J -modulation and zero quantum coherence effect was suppressed by using z-gradient filter having gradient strength 20% and a duration of 1 ms. Relaxation delays were set to 2.5 s. The acquisition data points were set to 4096×400 (complex points). Peak assignments and integrations were achieved using the software Sparky (UCSF).

2.3. Dimethylsulfate (DMS) footprinting experiments

Gel-purified 3'-end PDGFR- β G-quadruplex oligonucleotides with additional T₇ flanking ends were 5'-end labeled using [γ - ^{32}P] ATP in the presence of T4 polynucleotide kinase [25,29,34]. The DNA oligonucleotide used for 3in-Pu20 is: 5'-TTTTTTTCGGGGCAGGGAGGGTGGACGTTTTTTT-3', and that used for 5in-Pu22 is: 5'-TTTTTTTCGGCGGG GCAGGGAGGGTGGACTTTTTTTT-3'. G-quadruplex formation was achieved by denaturing the DNA at ~95 °C for 5 min, then slowly cooling to room temperature in 10 mM Tris-HCl buffer (pH = 7.6) in the presence or absence of 140 mM KCl. Methylation of the DNA oligonucleotides was performed using 0.5% (final concentration) DMS, 1 μg calf thymus DNA for 7 min at room temperature. DMS reactions were stopped by the addition of β -mercaptoethanol and glycerol to final concentrations of 7.7% and 2.6%, respectively. Electrophoresis on an 8% native PAGE gel was used to separate the intramolecular G-quadruplex. DNAs were isolated from the gel slices by soaking in water and then recovered by ethanol precipitation. The samples were treated with 10% piperidine for 16 min at 90 °C. Piperidine was removed from the samples by complete drying on a Speedvac, followed by two successive water washes. The cleaved products were separated on a 16% sequencing denaturing PAGE gel, which was exposed to a Phosphor-Imager screen overnight. The signal was detected on a Storm PhosphorImager.

2.4. Circular Dichroism (CD) spectroscopy experiments

Circular Dichroism spectra were recorded using a Jasco-810 spec-tropolarimeter (Jasco Inc., Easton, MD) equipped with a thermodynamically controlled cell holder. Samples were prepared in either 25 or 50 mM K⁺-containing K-phosphate buffer. CD measurements were taken through a quartz cell with an optical path length of 1 mm. The CD spectra were attained using three averaged scans between 230 and 330 nm at 25 °C. The baseline was corrected by subtracting signal contributions from buffers. T_m values were determined using CD melting experiments at 264 nm with a heating rate of 1 °C/min and 1 s response time between 20 and 95 °C.

2.5. Native Gel Electrophoretic Mobility Shift Assay (EMSA)

Native Gel EMSA experiments were conducted using a 1.5 mm thick 10 × 7 cm native gel containing 16% acrylamide (Acrylamide:Bis-acrylamide = 29:1) in TBE buffer, pH = 8 supplemented with 12.5 mM KCl. 0.4 nmol DNA was loaded into each well. DNA bands were visualized using ethidium bromide staining.

3. Results

3.1. The PDGFR-β NHE 3'-end minimal sequence with G-runs R1 to R4 and a 3'-terminal adenine does not form a stable monomeric G-quadruplex but has a tendency to form a dimer

The PDGFR-β NHE 3'-end minimal sequence Pu18 consists of G-runs R1–R4, with its 3'-most purine-run R1 (GGA) containing only two guanines (Fig. 2A). It was proposed that the 3'-end G-quadruplex contains an imperfect 3'-external G:G:G:A tetrad utilizing the 3'-most adenine A18 of R1 (Fig. 2B), which would be responsible for low stability [29]. We prepared PDGFR-β NHE 3'-end minimal sequence Pu18 (Fig. 2A) for NMR analysis. Pu18 showed a nice 1D ¹H NMR spectrum indicative of G-quadruplex formation (Fig. 2C); however, the Pu18 G-quadruplex appeared to be of dimeric nature, as shown by the native EMSA experiments (Fig. 3A). As the Pu18 sequence starts with a guanine at its 5'-end, which could favor dimer formation, we tested the 19-mer sequence Pu19 with a 5'-flanking cytosine (Fig. 2A). The additional 5'-flanking cytosine clearly reduced the population of the dimeric species in Pu19 as shown by native EMSA (Fig. 3A). Moreover, the ¹H NMR spectrum of Pu19 is clearly different from that of Pu18 (Fig. 2C), indicating the formation of a different G-quadruplex. However, the ¹H NMR spectrum of Pu19 was not well resolved; we therefore tested additional 3'-end or 5'-end flanking residues for Pu19. Because both of the 3'- and 5'-end flanking segments beyond the minimal 3'-end sequence Pu18 (R1–R4) contain guanine residues (Fig. 2A), which could potentially influence the G-quadruplex formation, we first tested various non-guanine flanking residues (Table S1). However, none of these sequences were able to form a stable intramolecular G-quadruplex, as indicated by the absence of well-resolved imino proton signals at 10.5–12 ppm in the 1D ¹H NMR spectra (Fig. S1), a region characteristic of G-quadruplex structure.

3.2. The 3'-extended PDGFR-β NHE 3'-end sequence forms a novel intramolecular G-quadruplex with an inserted guanine from the 3'-end: the 3'-Insertion-G4

We decided to examine the wild-type extended PDGFR-β NHE 3'-end sequence containing the non-adjacent flanking guanines (Fig. 4A). Unexpectedly, we found that the wild-type 20-mer sequence 3in-Pu20 with the non-adjacent flanking guanine G20 (Fig. 4A) showed a ¹H NMR with well-resolved imino proton peaks between 10.5 and 12 ppm in K⁺ solution (Fig. 4B), indicating the formation of a stable G-quadruplex structure. Importantly, the native EMSA showed that the 3in-Pu20 G-quadruplex is monomeric in nature (Fig. 3B, second lane from the left). In contrast, the ¹H NMR spectrum of Pu19 was not well resolved (Fig. 2C), indicating that Pu19 doesn't form one specific, well-defined G-quadruplex structure. The observation of a well-defined intramolecular G-quadruplex formed in 3in-Pu20 (Fig. 4B) but not in Pu19 (Fig. 2C) suggested that the 3'-flanking guanine G20 is critical for the G-quadruplex formed in 3in-Pu20.

The guanine imino H1 proton is one-bond connected to N1, and the guanine H8 proton is two-bond connected to N7 (Fig. 2D); they are readily detectable by 1D ^{15}N -edited HMQC NMR experiments using guanine site-specific labeled DNA [32,33]. We prepared 3in-Pu20 oligonucleotides with site-specific 6% incorporation of ^{15}N -labeled guanine at each guanine position [30–32] for imino proton H1 assignments by ^{15}N -edited HMQC NMR experiments (Fig. 4C) [33]. To incorporate ^{15}N -labeled guanine at G20, an additional 3'-end residue, such as thy-mine, is needed for solid-phase DNA synthesis. Using this sequence, we were able to observe the imino proton of G20 in the 10.5–12 ppm region (Fig. 4C, second top spectrum), demonstrating that G20 was indeed involved in G-tetrad formation. The imino proton of G2 (Fig. 4A) was not detected in 3in-Pu20 (Fig. 4C), indicating that G2 was not involved in G-tetrad formation. The synthesis of 3in-Pu20 with ^{15}N -labeled G17 was unsuccessful, so the assignment of the G17H1 was made by elimination (Fig. 4C). We carried out DMS footprinting of the 3in-Pu20 in the pH 7 solution containing 0 or 140 mM K^+ (Fig. 4D). The DMS footprinting showed that in addition to all guanines in R1–R3 and G3–G5 in R4, G20 from the 3'-end was protected, whereas G2 was not protected, consistent with the NMR data.

Based on the NMR and DMS footprinting data that G2 is not involved in the G-tetrad formation, we made the G2-to-A mutation in 3in-Pu20a (Fig. 4A), which showed an almost identical, but improved, NMR spectrum to that of 3in-Pu20 (Fig. 4B). Additional 5'-flanking bases, as shown in 3in-Pu22 (Fig. 4A), appeared to stabilize the G-quadruplex and gave rise to a more improved NMR spectrum (Fig. 4B). The native EMSA showed that the 3in-Pu22 G-quadruplex is also monomeric in nature (Fig. 3C, second lane from the left). The clear similarity of the ^1H NMR spectra of 3in-Pu22 and the wild-type 3in-Pu20 (Fig. 4B) indicated the formation of the same G-quadruplex. We prepared site-specifically labeled 3in-Pu22 oligonucleotides with 9% ^{15}N -guanine at each guanine position for unambiguous assignment of guanine imino H1 and aromatic H8 protons (Fig. 4E; Fig. S2). The imino proton assignment of 3in-Pu22 confirmed all of the imino proton assignments of 3in-Pu20 (Fig. 4C). 3in-Pu22-T with an additional thymine at the 3'-end was used to incorporate ^{15}N -labeled-guanine at G20; the H1 and H8 protons of G20 were clearly detected by ^{15}N -edited HMQC NMR experiments (second top traces, Figs. 4E and S2), demonstrating the involvement of G20 in G-tetrad formation. In addition, G17, which was missing in the assignment of 3in-Pu20 (Fig. 4C), was detected and confirmed in 3in-Pu22 (Fig. 4E). We performed 2D NMR experiments of 3in-Pu22 to determine the folding structure of this G-quadruplex. The observation of twelve imino proton peaks indicates the formation of a three-tetrad G-quadruplex. A G-tetrad plane is connected with Hoogs-teen H-bonds; the imino H1 protons of adjacent guanines, as well as the imino H1 and one adjacent guanine H8 protons, are in the close spatial vicinity and detectable by NOE connections (Fig. 2D). The assignment of guanine H1 and H8 protons led to the direct determination of the folding structure of this G-quadruplex by 2D-NOESY (Fig. S3). The arrangement and topology of the three G-tetrad planes were determined to be G3–G8–G12–G16, G4–G9–G13–G17, and G5–G10–G14–G20 (Fig. S3). Therefore, the 3'-extended PDGFR- β NHE 3'-end sequence appears to form an intramolecular parallel-stranded G-quadruplex of three G-tetrads, with the 3' non-adjacent guanine G20 inserted to complete the formation of the 3' external G-tetrad. We name this the 3'-insertion-G4 (Fig. 4F). The CD spectrum of 3in-Pu20 showed a

positive peak at 265 nm and a negative peak at 245 nm (Fig. 4G), characteristic of a parallel-stranded G-quadruplex structure, consistent with the folding structure of the 3'-insertion-G4.

Based on the folding structure of 3'-insertion-G4 (Fig. 4F), it is not difficult to understand that the additional flanking sequence at the 5' end, such as in 3in-Pu22 (Fig. 4A), stabilizes the 3'-insertion-G4 structure, as shown by improved NMR (Fig. 4B), presumably by the formation of capping structures. On the other hand, the additional 3'-end flanking bases, such as the 3'-T in 3in-Pu23 (Fig. 4A), may induce a more dynamic conformation of the inserted G20 that is next to a bulged loop, as shown by a reduced NMR spectral quality (Fig. 4B). However, it was important to note that, in the 3in-Pu23 sequence with additional 3'-end flanking base, the imino proton peaks were observed at almost the same positions as those in 3in-Pu22 with no 3'-end flanking (Fig. 4B), indicating the additional 3'-end flanking did not affect the G-quadruplex structure formed. Because the inserted G20 is two bases away from G17, the 3'-insertion-G4 contains a bulge structure formed by A18 and C19 (Fig. 4F). The formation of this bulge structure suggests that the 3'-most adenine A18 is not involved in the tetrad formation, unlike the previously proposed structure with a 3'-G:G:G:A tetrad (Fig. 2B). To confirm this, we prepared the 3in-Pu20b sequence with an A18-to-T mutation (Fig. 4A). 3in-Pu20b exhibited a very similar ¹H NMR spectrum (Fig. 4B) and CD profile (Fig. 4G) to that of 3in-Pu20a, indicating that A18 is not specifically required for the G-quadruplex formation.

3.3. The 5'-extended PDGFR-β NHE 3'-end sequence forms a second novel intramolecular G-quadruplex with an inserted guanine from the 5'-end: the 5'-Insertion-G4

Additional guanine residues, i.e., G-run R5, are present 5' to the PDGFR-β NHE minimal 3'-end sequence consisting R1–R4 G-runs (3'-end sequence, Fig. 5A). The formation of the 3'-insertion-G4 (Fig. 4F) led us to investigate whether a 5'-end insertion G-quadruplex could also be possible for the extended PDGFR-β 3'-end sequence. Because the R1 G-run contains only two guanines, G16 and G17 (Fig. 4F), they could move down to the 3'-end of the G-quadruplex to possibly accommodate a 5'-end guanine insertion. We carried out DMS footprinting experiments of the wild-type 5'-extended 5in-Pu22 that contains the 5' G-run R5 (Fig. 5A), in the absence and presence of 140 mM K⁺ at pH 7 (Fig. 5B). The DMS footprinting on the wild-type 5'-extended 5in-Pu22 sequence showed that, in the presence of K⁺, all the guanines in R1–R3 and G3–G5 in R4 were protected from DMS (Fig. 5B), indicating that G3–G5 were involved in the tetrad formation. In contrast, G2 in R4 was clearly less protected from DMS; the partial protection of G2 is likely caused by its involvement in the capping structure (see discussion below). Intriguingly, G0 of R5 (Fig. 5A) appeared to be clearly protected from DMS methylation, indicating that this 5' flanking non-adjacent guanine is involved in the tetrad formation. On the other hand, G–1 of the G-run R5 (Fig. 5A), was not protected from DMS, indicating that G–1 was not involved in the G-tetrad formation.

Based on the DMS footprinting data that G–1 was shown not to be involved in the tetrad formation, we prepared the 5'-extended sequence 5in-Pu21 with the G–1-to-A mutation (Figure 5A). Indeed, 5in-Pu21 formed a major well-defined G-quadruplex in K⁺ solution, as shown by the well-resolved NMR spectrum with imino proton peaks located between 10.5

and 12 ppm (Fig. 5C). Flanking bases stabilized this G-quadruplex and improved the NMR spectral quality, as shown in 5in-Pu24 (Fig. 5A), which appeared to form the same G-quadruplex as in 5in-Pu21 (Fig. 5C). Importantly, the 5in-Pu25 sequence containing both guanines of the G-run R5 appeared to form the same G-quadruplex as that formed in 5in-Pu25a which contains a G-1-to-A mutation, as shown by ^1H NMR (Fig. 5C); this result confirmed that G-1 was not involved in the tetrad formation, according to the DMS footprinting data (Fig. 5B), and that the G-1-to-A mutation did not affect the G-quadruplex structure. The clear similarity of the ^1H NMR imino regions of all the 5in-sequences (Fig. 5C) indicated the formation of the same G-quadruplex. Importantly, native EMSA of 5in-Pu24 showed that this G-quadruplex is monomeric (intramolecular) in nature (Fig. 3D, second lane from the left).

We prepared site-specifically labeled 5in-Pu24 oligonucleotides with 9% ^{15}N -guanine at each guanine position for the unambiguous assignment of guanine imino H1 (Fig. 5D) and aromatic H8 protons (Fig. S4) by ^{15}N -edited HMQC NMR experiments [30–33]. Indeed, the imino proton of G0 in the 5'-flanking region (Fig. 5A) was detected at 11.4 ppm, confirming that this 5' non-adjacent guanine was involved in G-tetrad formation of this intramolecular G-quadruplex. Of the thirteen guanines in 5in-Pu24, the imino proton of G2 was the only one that was not detected (Fig. 5D), indicating that G2 is not involved in G-tetrad formation. The observation of twelve imino proton peaks indicates the formation of a 3-tetrad G-quadruplex.

We performed 2D NMR experiments of 5in-Pu24 to determine the folding structure of the 5'-insertion-G4. The assignment of guanine H1 and H8 protons led to the direct determination of the folding structure of the 5in-Pu24 G-quadruplex by 2D-NOESY (Fig. S5). The arrangement and topology of the three G-tetrad planes were determined to be G3–G8–G12–G0, G4–G9–G13–G16, and G5–G10–G14–G17 (Fig. S5). Therefore, the 5'-extended PDGFR- β NHE 3'-end sequence appears to form a second intramolecular parallel-stranded three-tetrad G-quadruplex, with the 5'-flanking guanine G0 inserted to complete the formation of the 5' external G-tetrad (Fig. 5E). We name this G-quadruplex the 5'-insertion-G4. The CD spectrum of 5in-Pu24 showed a positive peak at 265 nm and a negative peak at 245 nm (Fig. 4G), consistent with the parallel-stranded folding structure of the 5'-insertion-G4. As speculated, in the 5'-insertion-G4, G16 and G17 in the R1 G-run move down to the 3'-end to be involved in the middle and 3'-external G-tetrads, respectively (Fig. 5E). As shown by both NMR and DMS footprinting, G0, but not G-1, of the 5'-flanking G-run R5 (Fig. 5A), appears to be predominantly involved in the formation of the 5'-external G-tetrad in the 5'-insertion-G4 (Fig. 5E). Interestingly, the 5'-insertion-G4 appeared to be able to tolerate additional flanking residues at both the 5'- and 3'-ends (Fig. 5A and C).

3.4. The extended PDGFR- β NHE 3'-end sequence with flanking segments at both ends forms a dynamic mixture of the 3'-Insertion-G4 and the 5'-Insertion-G4

As it appears that either the 3'-end or the 5'-end non-adjacent guanine in the flanking segments of the PDGFR- β NHE extended 3'-end sequence can be inserted to form a 3'-external or 5'-external G-tetrad, it is interesting to investigate whether these two

intramolecular end-insertion G-quadruplexes, i.e., the 3'-insertion-G4 and the 5'-insertion-G4, can co-exist in the wild-type PDGFR- β NHE extended 3'-end sequence. We prepared a 3'- and 5'-extended sequence, 3&5in-Pu23 (Fig. 6A), that contains 3in-Pu20 that forms the 3'-insertion-G4 and 5in-Pu24 that forms the 5'-insertion-G4 and hence has the potential to form both the end-insertion-G-quadruplexes. Indeed, the 1D ^1H NMR spectrum of 3& 5in-Pu23 indicated a mixture of more than one G-quadruplexes (Fig. 6B, top). We prepared site-specifically labeled 3&5in-Pu23 oligo-nucleotides with 6% ^{15}N -guanine at guanine positions G12, G13, G14, and G16. These four guanines were selected because their imino proton chemical shifts in the 3'-insertion-G4 (3in-Pu20) and the 5'-insertion-G4 (5in-Pu24) were distinctly different (Figs. 4C and 5D). We performed 1D ^{15}N -edited HMQC NMR experiments of each site-specifically labeled 3&5in-Pu23 for the assignment of guanine imino protons [30–33]. Significantly, each of these site-specifically labeled samples showed two imino proton peaks (Fig. 6B), indicating the co-presence of two G-quadruplexes in the extended sequence 3&5in-Pu23. Comparison of the imino proton chemical shifts in 3&5in-Pu23 to those in 3in-Pu20 and 5in-Pu24 (Fig. 6C) showed that both 3'-insertion-G4 and 5'-insertion-G4 co-exist in 3&5in-Pu23. Specifically, one set of imino protons of G12, G13, G14, and G16 for 3&5in-Pu23 was detected at nearly the same chemical shifts as those observed for 3in-Pu20 which forms 3'-insertion-G4 (labeled as “a”), while another set of imino protons was detected at nearly the same positions as those observed for 5in-Pu24 which forms 5'-insertion-G4 (labeled as “b”) (Fig. 6C). The CD spectrum of 3&5in-Pu23 markedly resembled those of 3in-Pu20 and 5in-Pu24 (Fig. 4G). Therefore, it appeared that the two end-insertion G-quadruplexes both form and co-exist in equilibrium in the PDGFR- β NHE extended 3'-end sequence, which contains non-adjacent guanines in both the 3'- and 5'-flanking segments (Fig. 6D). It is noted that the 12a imino peak in 3& 5in-Pu23 is shifted from that observed in 3in-Pu20 (Fig. 6C); this can be explained by the presence of additional 5'-flanking segment in 3&5in-Pu23, which could form a capping structure that specifically stacks over G12 of the 5'-external tetrad in the 3'-insertion-G4 to affect the imino chemical shift of G12.

3.5. GSA1129 significantly stabilizes both of the end-insertion G4s formed in the PDGFR- β NHE extended 3'-end sequence

We studied the effect of GSA1129 (Fig. S6 and S7) on the stability of the two end-insertion G-quadruplexes formed in the extended PDGFR- β NHE 3'-end sequence (Fig. 6D). We first tested the effect of GSA1129 on the thermal stability of the individual end-insertion G-quadruplex, i.e., 3'-insertion-G4 and 5'-insertion-G4 formed in 3in-Pu22 and 5in-Pu24, respectively (Fig. 6D), using CD melting experiments (Fig. S7). We found that GSA1129 markedly increased the T_m of the 3'-insertion-G4 and the 5'-insertion-G4, by about 30 °C and 26 °C, respectively, at a two-molecule drug equivalence in 50 mM K^+ (Table 1; Fig. S7). We then tested the stabilizing effect of GSA1129 on the both-end-extended sequence 3&5in-Pu23 (Fig. 6A), which forms a mixture of both the 3'-insertion-G4 and the 5'-insertion-G4 (Fig. 6D). GSA1129 was shown to increase the T_m of 3&5in-Pu23 by about 30 °C at a two-molecule equivalence in 50 mM K^+ (Table 1; Fig. S7). This increase of the T_m induced by GSA1129 of the two end-insertion G-quadruplexes formed in the extended 3'-end sequence was significantly greater as compared to other G-quadruplexes formed in the PDGFR- β NHE, including the most stable 5'-mid G4 [29]. We also carried out native EMSA

experiments of the 3'-insertion-G4 forming sequence 3in-Pu22 (Fig. 3C) and the 5'-insertion-G4 forming sequence 5in-Pu24 (Fig. 3D) in the presence of GSA1129 at 0.5, 1, and 2 equivalents. Importantly, the EMSA results showed that the complexes of GSA1129 with the 3'-insertion-G4 and 5'-insertion-G4 were both of monomeric nature (Fig. 3C and D). Collectively, the selective binding of GSA1129 significantly stabilizes the two end-insertion G-quadruplexes formed in the PDGFR- β NHE 3'-end sequence, which in the absence of drug are less stable than other G-quadruplexes formed in PDGFR- β NHE.

4. Discussion

The 3'-most guanine run (R1, GGA) of the PDGFR- β NHE 3'-end minimal sequence (R1–R4, Fig. 2A) contains only two guanines and an adenine. The wild-type PDGFR- β NHE sequence contains non-adjacent guanines in the flanking regions of the extended 3'-end sequence, i.e., two guanines in the 5' flanking G-run R5 (G0 and G–1) and a guanine (G20) in the 3' flanking (Fig. 2A). It is highly unexpected and intriguing that the extended 3'-end sequence with non-adjacent guanines can form two intramolecular end-insertion G-quadruplexes, with an inserted guanine from either the 3' or 5' flanking regions, and that the two end-insertion G-quadruplexes co-exist in equilibrium in potassium solution. In the 3'-insertion-G4 (Fig. 4F), the 3' non-adjacent guanine G20, which is two bases away 3' from the last guanine (G17) of R1, is involved in the formation of the 3'-external G-tetrad with G5, G10, and G14 from G-runs R4, R3, and R2, respectively, with the two guanines G16 and G17 from R1 involved in the formation of the 5'-external and middle G-tetrads, thereby inducing the formation of a bulge loop structure consisting of A18 and C19. In the 5'-insertion-G4 (Fig. 5E), the 5' non-adjacent guanine G0, which is two bases away 5' from the G3 of R4, is involved in the formation of the 5'-external G-tetrad with the G3, G8, and G12 from G-runs R4, R3, and R2, respectively. As a result, guanines G16 and G17 from R1 move down to be involved in the formation of the middle and 3'-external G-tetrads; C1 and G2 of the 5' flanking segment between G0 and G3 form a new lateral loop on top of the 5'-external tetrad (Fig. 5E). Therefore, it appears that the two guanines in R1 can move up or down the G-quadruplex, so that either a 5'-non-adjacent flanking guanine or a 3'-non-adjacent flanking guanine can be recruited to form the 5'-external G-tetrad in 5'-insertion-G4 or the 3'-external G-tetrad in 3'-insertion-G4, respectively (Fig. 6D); the formation of a specific end-insertion G-quadruplex appears to depend on the core G-rich sequence and the guanine-containing flanking sequences.

It was previously suggested that the PDGFR- β NHE 3'-end G-quadruplex contains an imperfect 3'-external G:G:G:A tetrad involving the 3'-terminal adenine of R1 (Fig. 2B), which leads to low stability [29]; however, our results show that the formation of a stable 3'-external G:G:G:A tetrad is unlikely. This is not surprising, as the adenine base contains a hydrogen instead of an amino group on C2 and an amino group instead of oxygen on C6, with the absence of the imino proton on N1; thus, the adenine base does not fit well with the Hoogsteen hydrogen-bonding network with neighboring guanines in a tetrad formation (Fig. 2D). It is noted that the PDGFR- β NHE 3'-end minimal sequence Pu18, while showing nice NMR and CD spectra, forms predominantly a dimer. This underscores the importance of identifying the proper sequence for the formation of the biologically relevant G-quadruplex structure.

The loop interactions and capping structures of the two end-insertion G-quadruplexes appear to be different. For the 3'-insertion-G4 (Fig. 4F), additional 5'-flanking bases appear to stabilize the structure (Fig. 4B), likely due to the capping interactions with the 5'-external G-tetrad, as indicated by the clear shift of the G12 imino proton (Fig. 6C). In contrast, additional residues at the 3'-end extension appear to reduce the stability of the 3'-insertion-G4; this is likely caused by the dynamic nature of the 3'-external-tetrad with the bulged A18–C19 structure. On the other hand, for the 5'-insertion-G4 (Fig. 5E), additional flanking residues at both the 5'- and 3'-ends appear to stabilize the structure (Fig. 5C), indicating favorable capping interactions at both ends of the G-quadruplex. Specifically, additional 5'-end flanking residues do not affect the stable formation of the 5'-insertion-G4, suggesting a favorable capping structure adopted by the lateral loop consisting of C1 and G2 (Fig. 5E). Interestingly, the run R5 contains two guanines (G0 and G-1) (Fig. 5A), but G0 is shown to be more preferred for the insertion into the 5'-external tetrad; this again suggests a specific capping structure formed by the lateral loop containing C1 and G2.

The PDGFR- β NHE 3'-end G-quadruplex has a significantly lower thermal stability, as compared to the most stable 5'-mid G-quadruplex [29]. GSA1129 appears to selectively bind to the 3'-end G-quadruplex in the PDGFR- β NHE and significantly increase the stabilities of the two end-insertion G-quadruplexes, by about 30 °C. The selective binding and stabilization of GSA1129 of the two end-insertion G-quadruplexes formed in the PDGFR- β NHE 3'-end sequence can shift the equilibrium to the formation of the less stable 3'-end G-quadruplex (minor conformation) in the PDGFR- β NHE for transcriptional repression.

5. Conclusions

G-quadruplexes have recently been found to form in the proximal promoters of human oncogenes. Unlike the human telomeric sequence, the G-quadruplex-forming regions in oncogene promoters are more diverse in sequence and length and can potentially form multiple G-quadruplexes [11,35]. The parallel-stranded structure, with the robust parallel-stranded structure motif G3NG3, is found to be prevalent in the promoter G-quadruplexes; however, the parallel-stranded G-quadruplexes can come in a variety of forms [35], including those with bulged, inserted, and vacant bases [28,36–41]. In this study, we show that the wild-type extended PDGFR- β NHE 3'-end sequence, which contains a GGA purine-run (R1) and 3'- and 5'- non-adjacent guanines in its flanking sequences, forms the two intramolecular end-insertion G-quadruplexes which co-exist in equilibrium in potassium solution (Fig. 6D). To our knowledge, this is the first example of the formation of two stable end-insertion G-quadruplexes that co-exist in equilibrium in a native promoter sequence. Significantly, the selective binding and stabilization of GSA1129 of the two end-insertion G-quadruplexes can shift the equilibrium to the formation of the less stable 3'-end G-quadruplex in the PDGFR- β NHE for transcriptional repression. The dynamic equilibrium of the two intramolecular G-quadruplexes may provide a potential means for PDGFR- β transcriptional regulation and the novel end-insertion structures may make an effective target for small molecule binding.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This research was supported by grants from the National Institutes of Health (R01CA177585 (DY), R01CA153821 (LHH), and P30CA023168 (DY, Purdue Center for Cancer Research). We thank Kaibo Wang for his help in preparing DNA oligonucleotides and CD measurements. We thank Drs. David Bishop, Clement Lin, and Pradeep Palakshan for proofreading the manuscript.

References

- [1]. Andrae J, Gallini R, Betsholtz C, Role of platelet-derived growth factors in physiology and medicine, *Genes Dev.* 22 (2008) 1276–1312. [PubMed: 18483217]
- [2]. Heldin CH, Westermark B, Mechanism of action and in vivo role of platelet-derived growth factor, *Physiol. Rev* 79 (1999) 1283–1316. [PubMed: 10508235]
- [3]. Heldin CH, Targeting the PDGF signaling pathway in tumor treatment, *Cell Commun. Signal* 11 (2013) 97. [PubMed: 24359404]
- [4]. Pietras K, Sjöblom T, Rubin K, Heldin CH, Ostman A, PDGF receptors as cancer drug targets, *Cancer Cell* 3 (2003) 439–443. [PubMed: 12781361]
- [5]. Fjallskog ML, Hessman O, Eriksson B, Janson ET, Upregulated expression of PDGF receptor beta in endocrine pancreatic tumors and metastases compared to normal endocrine pancreas, *Acta Oncol.* 46 (2007) 741–746. [PubMed: 17653895]
- [6]. Gilbertson RJ, Clifford SC, PDGFRB is overexpressed in metastatic medulloblastoma, *Nat. Genet* 35 (2003) 197–198. [PubMed: 14593398]
- [7]. Pietras K, Ostman A, Sjöquist M, Buchdunger E, Reed RK, Heldin CH, Rubin K, Inhibition of platelet-derived growth factor receptors reduces interstitial hypertension and increases transcapillary transport in tumors, *Cancer Res.* 61 (2001) 2929–2934. [PubMed: 11306470]
- [8]. Östman A, Heldin C-H, PDGF receptors as targets in tumor treatment, *Adv. Cancer Res* 97 (2007) 247–274. [PubMed: 17419949]
- [9]. Govindarajan B, Shah A, Cohen C, Arnold RS, Schechner J, Chung J, Mercurio AM, Alani R, Ryu B, Fan CY, Cuezva JM, Martinez M, Arbiser JL, Malignant transformation of human cells by constitutive expression of platelet-derived growth factor-BB, *J. Biol. Chem* 280 (2005) 13936–13943. [PubMed: 15695519]
- [10]. Heldin C-H, Rubin K, Pietras K, Östman A, High interstitial fluid pressure—an obstacle in cancer therapy, *Nat. Rev. Cancer* 4 (2004) 806–813. [PubMed: 15510161]
- [11]. Qin Y, Hurley LH, Structures, folding patterns, and functions of intramolecular DNA G-quadruplexes found in eukaryotic promoter regions, *Biochimie* 90 (2008) 1149–1171. [PubMed: 18355457]
- [12]. Brooks TA, Hurley LH, The role of supercoiling in transcriptional control of Myc and its importance in molecular therapeutics, *Nat. Rev. Cancer* 9 (2009) 849–861. [PubMed: 19907434]
- [13]. Yang D, Okamoto K, Structural insights into G-quadruplexes: towards new anticancer drugs, *Future Med. Chem* 2 (2010) 619–646. [PubMed: 20563318]
- [14]. Simonsson T, Pecinka P, Kubista M, DNA tetraplex formation in the control region of c-Myc, *Nucleic Acids Res.* 26 (1998) 1167–1172. [PubMed: 9469822]
- [15]. Siddiqui-Jain A, Grand CL, Bearss DJ, Hurley LH, Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-Myc transcription, *Proc. Natl. Acad. Sci. U. S. A* 99 (2002) 11593–11598. [PubMed: 12195017]
- [16]. Sun D, Guo K, Rusche JJ, Hurley LH, Facilitation of a structural transition in the polypurine/polypyrimidine tract within the proximal promoter region of the human VEGF gene by the presence of potassium and G-quadruplex-interactive agents, *Nucleic Acids Res.* 33 (2005) 6070–6080. [PubMed: 16239639]

- [17]. De Armond R, Wood S, Sun D, Hurley LH, Ebbinghaus SW, Evidence for the presence of a guanine quadruplex forming region within a polypurine tract of the hypoxia inducible factor 1 α promoter, *Biochemistry* 44 (2005) 16341–16350. [PubMed: 16331995]
- [18]. Dai J, Dexheimer TS, Chen D, Carver M, Ambrus A, Jones RA, Yang DZ, An intramolecular G-quadruplex structure with mixed parallel/antiparallel G-strands formed in the human Bcl-2 promoter region in solution, *J. Am. Chem. Soc* 128 (2006) 1096–1098. [PubMed: 16433524]
- [19]. Dexheimer TS, Sun D, Hurley LH, Deconvoluting the structural and drug-recognition complexity of the G-quadruplex-forming region upstream of the Bcl-2 P1 promoter, *J. Am. Chem. Soc* 128 (2006) 5404–5415. [PubMed: 16620112]
- [20]. Cogoi S, Xodo LE, G-quadruplex formation within the promoter of the KRAS proto-oncogene and its effect on transcription, *Nucleic Acids Res.* 34 (2006) 2536–2549. [PubMed: 16687659]
- [21]. Xu Y, Sugiyama H, Formation of the G-quadruplex and i-motif structures in retinoblastoma susceptibility genes (Rb), *Nucleic Acids Res.* 34 (2006) 949–954. [PubMed: 16464825]
- [22]. Rankin S, Reszka AP, Huppert J, Zloh M, Parkinson GN, Todd AK, Ladame S, Balasubramanian S, Neidle S, Putative DNA quadruplex formation within the human c-kit oncogene, *J. Am. Chem. Soc* 127 (2005) 10584–10589. [PubMed: 16045346]
- [23]. Fernando H, Reszka AP, Huppert J, Ladame S, Rankin S, Venkitaraman AR, Neidle S, Balasubramanian S, A conserved quadruplex motif located in a transcription activation site of the human c-kit oncogene, *Biochemistry* 45 (2006) 7854–7860. [PubMed: 16784237]
- [24]. Guo K, Pourpak A, Beetz-Rogers K, Gokhale V, Sun D, Hurley LH, Formation of pseudosymmetrical G-quadruplex and i-motif structures in the proximal promoter region of the RET oncogene, *J. Am. Chem. Soc* 129 (2007) 10220–10228. [PubMed: 17672459]
- [25]. Palumbo SL, Ebbinghaus SW, Hurley LH, Formation of a unique end-to-end stacked pair of G-quadruplexes in the hTERT core promoter with implications for inhibition of telomerase by G-quadruplex-interactive ligands, *J. Am. Chem. Soc* 131 (2009) 10878–10891. [PubMed: 19601575]
- [26]. Qin Y, Fortin JS, Tye D, Gleason-Guzman M, Brooks TA, Hurley LH, Molecular cloning of the human platelet-derived growth factor receptor β (PDGFR- β) promoter and drug targeting of the G-quadruplex-forming region to repress PDGFR- β expression, *Biochemistry* 49 (2010) 4208–4219. [PubMed: 20377208]
- [27]. Hansel-Hertsch R, Beraldi D, Lensing SV, Marsico G, Zyner K, Parry A, Di Antonio M, Pike J, Kimura H, Narita M, Tannahill D, Balasubramanian S, G-quadruplex structures mark human regulatory chromatin, *Nat. Genet* 48 (2016) 1267–1272. [PubMed: 27618450]
- [28]. Chen Y, Agrawal P, Brown RV, Hatzakis E, Hurley L, Yang D, The major G-quadruplex formed in the human platelet-derived growth factor β receptor promoter adopts a novel broken-strand structure in K⁺ solution, *J. Am. Chem. Soc* 134 (2012) 13220–13223. [PubMed: 22866911]
- [29]. Brown RV, Wang T, Chappeta VR, Wu G, Onel B, Chawla R, Quijada H, Camp SM, Chiang ET, Lassiter QR, Lee C, Phanse S, Turnidge MA, Zhao P, Garcia JGN, Gokhale V, Yang D, Hurley LH, The consequences of overlapping G-quadruplexes and i-motifs in the platelet-derived growth factor receptor beta core promoter nuclease hypersensitive element can explain the unexpected effects of mutations and provide opportunities for selective targeting of both structures by small molecules to downregulate gene expression, *J. Am. Chem. Soc* 139 (2017) 7456–7475. [PubMed: 28471683]
- [30]. Dai J, Chen D, Jones RA, Hurley LH, Yang D, NMR solution structure of the major G-quadruplex structure formed in the human Bcl-2 promoter region, *Nucleic Acids Res.* 34 (2006) 5133–5144. [PubMed: 16998187]
- [31]. Agrawal P, Hatzakis E, Guo K, Carver M, Yang D, Solution structure of the major G-quadruplex formed in the human VEGF promoter in K⁺: insights into loop interactions of the parallel G-quadruplexes, *Nucleic Acids Res.* 41 (2013) 10584–10592. [PubMed: 24005038]
- [32]. Phan AT, Patel DJ, A site-specific low-enrichment (15)N,(13)C isotope-labeling approach to unambiguous NMR spectral assignments in nucleic acids, *J. Am. Chem. Soc* 124 (2002) 1160–1161. [PubMed: 11841271]

- [33]. Szewczak AA, Kellogg GW, Moore PB, Assignment of NH resonances in nucleic acids using natural abundance ^{15}N - ^1H correlation spectroscopy with spin-echo and gradient pulses, *FEBS Lett.* 327 (1993) 261–264. [PubMed: 7688695]
- [34]. Sun D, Hurley LH, Biochemical techniques for the characterization of G-quad-ruplex structures: EMSA, DMS footprinting, and DNA polymerase stop assay, *Methods Mol. Biol* 608 (2010) 65–79. [PubMed: 20012416]
- [35]. Onel B, Lin C, Yang D, DNA G-quadruplex and its potential as anticancer drug target, *SCIENCE CHINA Chem.* 57 (2014) 1605–1614.
- [36]. Phan AT, Kuryavyi V, Burge S, Neidle S, Patel DJ, Structure of an unprecedented G-quadruplex scaffold in the human c-kit promoter, *J. Am. Chem. Soc* 129 (2007) 4386–4392. [PubMed: 17362008]
- [37]. Phan AT, Kuryavyi V, Gaw HY, Patel DJ, Small-molecule interaction with a five-guanine-tract G-quadruplex structure from the human Myc promoter, *Nat. Chem. Biol* 1 (2005) 167–173. [PubMed: 16408022]
- [38]. Mukundan VT, Phan AT, Bulges in G-quadruplexes: broadening the definition of G-quadruplex-forming sequences, *J. Am. Chem. Soc* 135 (2013) 5017–5028. [PubMed: 23521617]
- [39]. Li XM, Zheng KW, Zhang JY, Liu HH, He YD, Yuan BF, Hao YH, Tan Z, Guanine-vacancy-bearing G-quadruplexes responsive to guanine derivatives, *Proc. Natl. Acad. Sci. U. S. A* 112 (2015) 14581–14586. [PubMed: 26553979]
- [40]. Heddi B, Martin-Pintado N, Serimbetov Z, Kari TM, Phan AT, G-quadruplexes with $(4n - 1)$ guanines in the G-tetrad core: formation of a G-triad-water complex and implication for small-molecule binding, *Nucleic Acids Res.* 44 (2016) 910–916. [PubMed: 26673723]
- [41]. Onel B, Carver M, Wu G, Timonina D, Kalarn S, Larriva M, Yang D, A new G-quadruplex with hairpin loop immediately upstream of the human Bcl-2 P1 promoter modulates transcription, *J. Am. Chem. Soc* 138 (2016) 2563–2570. [PubMed: 26841249]

Fig. 1. The PDGFR-β core promoter NHE. This PDGFR-β promoter region was previously shown to form four predominant G-quadruplexes from overlapping sequences, i.e., 3'-end, 3'-mid, 5'-mid, and 5'-end, with seven guanine runs underlined and numbered R1–R7 [26].

Fig. 2.

(A) The extended PDGFR- β NHE 3'-end sequence and the minimal 3'-end sequences Pu18 and Pu19 used in the previous studies [26,29]. The G-runs are underlined and numbered. (B) Schematic drawing of the proposed parallel-stranded G-quadruplex formed in Pu18 and Pu19 with a G:G:G:A 3'-external tetrad (G = red, A = green, C = yellow, T = blue). (C) Imino regions of 1D NMR spectra of Pu18 (top) and Pu19 (bottom) at 25 °C in 50 mM K⁺ solution at pH 7. (D) Schematic drawing of a G-tetrad.

Fig. 3. Native EMSA experiments in 50 mM K^+ -containing solution at pH 7. (A) Pu19 and Pu18. (B) 3in-Pu20. (C) 3in-Pu22 in the absence and presence of various equivalents of GSA1129. (D) 5in-Pu24 in the absence and presence of various equivalents of GSA1129.

Fig. 4.

(A) The 3'-extended wild-type PDGFR-β NHE 3'-end sequence 3in-Pu20 and its variants. The full imino H1 proton assignments were made for 3in-Pu20 and 3in-Pu22, which are shown in bold. The modified bases are underlined. The PDGFR-β NHE extended 3'-end sequence, with the numbering system and G-runs underlined, is shown for reference. (B) Imino regions of 1D NMR spectra of various 3in-sequences at 25 °C in 50 mM K⁺ solution at pH 7.0. The imino H1 proton assignments of 3in-Pu20 and 3in-Pu22 are shown. (C) Imino H1 proton assignment of 3in-Pu20 by 1D ¹⁵N-edited HMQC experiments using site-specifically labeled oligonucleotides at 25 °C in 50 mM K⁺ solution at pH 7. (D) DMS chemical footprinting of 3in-Pu20 in the presence of 140 mM K⁺ (lane 1) and 0 mM K⁺ (lane 2) with densitometric scans (right). Lanes 3 and 4 serve as control experiments in the absence of DMS. (E) Imino H1 proton assignment of 3in-Pu22 by 1D ¹⁵N-edited HMQC experiments using site-specifically labeled oligonucleotides at 25 °C in 50 mM K⁺ solution at pH 7. (F) Schematic drawing of the 3'-insertion-G4 with the 3' non-adjacent G20 inserted into the 3'-external G-tetrad (G = red, A = green, C = yellow, T = blue). G16 and G17 of R1 are involved in the 5'-external and middle G-tetrads. (G) CD spectra of Pu19, various 3in-sequences, 5in-Pu24, and 3&5in-Pu23 in 50 mM K⁺ solution at pH 7.

Fig. 5.

(A) The 5'-extended wild-type PDGFR- β NHE 3'-end sequence 5in-Pu22 and its variants. The full imino H1 proton assignments were made for 5in-Pu24, which is shown in bold. The PDGFR- β NHE extended 3'-end sequence, with the numbering system and G-runs underlined, is shown for reference. (B) DMS chemical footprinting of 5in-Pu22 in the presence of 140 mM K⁺ (lane 1) and 0 mM K⁺ (lane 2) with densitometric scans (right). Lanes 3 and 4 serve as control experiments in the absence of DMS. (C) Imino regions of 1D NMR spectra of various 5in-sequences at 25 °C in 50 mM K⁺ solution at pH 7.0. (D) Imino H1 proton assignment of 5in-Pu24 by 1D ¹⁵N-edited HMQC experiments using site-specifically labeled oligonucleotides at 25 °C, in 50 mM K⁺ solution at pH 7. (E) Schematic drawing of the 5'-insertion-G4 with the 5' non-adjacent G0 inserted into the 5'-external G-tetrad (G = red, A = green, C = yellow, T = blue). G16 and G17 of R1 are involved in the middle and 3'-external G-tetrads.

Fig. 6.

(A) Extended PDGFR- β NHE 3'-end sequences containing the 3' flanking guanine (3in-Pu20), the 5' flanking guanine (5in-Pu24), or both (3&5in-Pu23). The numbering system is shown. (B) Selected imino H1 proton assignments of 3&5in-Pu23 by 1D ^{15}N -edited HMQC experiments using site-specifically labeled oligonucleotides at 25 °C in 50 mM K^+ solution at pH 7. (C) Comparison of selected imino H1 proton assignments of 3 & 5in-Pu23 with 3in-Pu20 (3'-insertion-G4) and 5in-Pu24 (5'-insertion-G4) at 25 °C in 50 mM K^+ solution at pH 7. The imino proton peaks of 3&5in-Pu23 belonging to the 3'-insertion-G4 and the corresponding imino protons from 3in-Pu20 are labeled with "a", while the imino proton peaks of 3&5in-Pu23 belonging to the 5'-insertion-G4 and the corresponding imino protons from 5in-Pu24 are labeled with "b". (D) Schematic drawing of the two end-insertion G-quadruplexes that co-exist in equilibrium in the extended PDGFR- β NHE 3'-end sequence:

the 3'-insertion-G4 (left) and the 5'-insertion-G4 (right) (G = red, A = green, C = yellow, T = blue).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

The CD melting temperatures of 3in-Pu22, 5in-Pu24, and 3&5in-Pu23 in 50 mM K⁺ pH 7 solution, and the changes of T_m in the presence of GSA1129 at two equivalents.

DNA	T _m (°C)	T _m (°C)
3in-Pu22	50	30
5in-Pu24	44	26
3&5in-Pu23	46	30