Published in final edited form as:

J Am Chem Soc. 2017 June 07; 139(22): 7456–7475. doi:10.1021/jacs.6b10028.

The Consequences of Overlapping G-Quadruplexes and i-Motifs in the Platelet-Derived Growth Factor Receptor β 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

Robert V. Brown[†], Ting Wang[‡], Venkateshwar Reddy Chappeta[†], Guanhui Wu[†], Buket Onel[†], Reena Chawla[§], Hector Quijada[‡], Sara M. Camp[‡], Eddie T. Chiang[‡], Quinea R. Lassiter[⊥], Carmen Lee^{⊥,||}, Shivani Phanse^{||}, Megan A. Turnidge^{||}, Ping Zhao[▽], Joe G. N. Garcia[‡], Vijay Gokhale^{†,§}, Danzhou Yang^{†,§,||}, and Laurence H. Hurley^{†,§,||,*}

[†]College of Pharmacy, University of Arizona, 1703 East Mabel Street, Tucson, Arizona 85721, United States

[‡]College of Medicine, University of Arizona, 1501 North Campbell Avenue, Tucson, Arizona 85724, United States

[§]BIO5 Institute, 1657 East Helen Street, Tucson, Arizona 85721, United States

[⊥]College of Agriculture & Life Sciences, University of Arizona, 1117 East Lowell Street, Tucson, Arizona 85721, United States

^{||}College of Science, University of Arizona, 1040 East Fourth Street, Tucson, Arizona 85721, United States

*Corresponding Author: hurley@pharmacy.arizona.edu.

R.B.: National Institute of Environmental Health Sciences, 111 T.W. Alexander Drive, Building 101, Research Triangle Park, North Carolina 27709, USA.

V.R.C.: VNR Biosciences Pvt. Ltd., Plot No. 22-362, IDA, Jeedimetla, Hyderabad, 500 055 Telangana, India.

B.O., G.W., and D.Y.: College of Pharmacy, Medicinal Chemistry and Molecular Pharmacology, Purdue University, 575 Wes Stadium Avenue, West Lafayette, Indiana 47907, USA.

D.Y.: Purdue University Center for Cancer Research, 201 South University Street, West Lafayette, Indiana 47906, USA.

D.Y.: Purdue Institute for Drug Discovery, 720 Clinic Drive, West Lafayette, Indiana 47907, USA.

Q.R.L.: Ventana Medical Systems, Inc., 1910 Innovation Park Drive, Tucson, Arizona 85755, USA.

C.L.: Laboratory Corporation of America, 5005 South 40th Street, Phoenix, Arizona 85040, USA.

S.P.: California School of Podiatric Medicine, 450 30th Street, Suite 2860, Oakland, California 94609, USA.

M.A.T.: The University of Arizona, 475 North 5th Street, Phoenix, Arizona 85004, USA.

ORCID

Laurence H. Hurley: 0000-0002-8522-450X

Notes

The authors declare the following competing financial interest(s): Laurence Hurley and Vijay Gokhale have a financial interest in Reglagene, a G-quadruplex-targeting company.

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.6b10028.

CD data of NSC338258 with WT 41-mer, control experiments with GSA1125 showing no effect on acute lung injury, detailed synthetic methods, and chemical characterization of compounds (PDF)

[‡]School of Chemistry and Chemical Engineering, Guangdong Pharmaceutical University, No. 280 Waihuandong Road, Education Mega Centre, Guanzhou 510006, Peoples Republic of China

[¶]University of Arizona Cancer Center, 1515 North Campbell Avenue, Tucson, Arizona 85724, United States

Abstract

The platelet-derived growth factor receptor β (PDGFR- β) signaling pathway is a validated and important target for the treatment of certain malignant and nonmalignant pathologies. We previously identified a G-quadruplex-forming nuclease hypersensitive element (NHE) in the human PDGFR- β promoter that putatively forms four overlapping G-quadruplexes. Therefore, we further investigated the structures and biological roles of the G-quadruplexes and i-motifs in the PDGFR- β NHE with the ultimate goal of demonstrating an alternate and effective strategy for molecularly targeting the PDGFR- β pathway. Significantly, we show that the primary G-quadruplex receptor for repression of PDGFR- β is the 3'-end G-quadruplex, which has a GGA sequence at the 3'-end. Mutation studies using luciferase reporter plasmids highlight a novel set of G-quadruplex point mutations, some of which seem to provide conflicting results on effects on gene expression, prompting further investigation into the effect of these mutations on the i-motif-forming strand. Herein we characterize the formation of an equilibrium between at least two different i-motifs from the cytosine-rich (C-rich) sequence of the PDGFR- β NHE. The apparently conflicting mutation results can be rationalized if we take into account the single base point mutation made in a critical cytosine run in the PDGFR- β NHE that dramatically affects the equilibrium of i-motifs formed from this sequence. We identified a group of ellipticines that targets the G-quadruplexes in the PDGFR- β promoter, and from this series of compounds, we selected the ellipticine analog GSA1129, which selectively targets the 3'-end G-quadruplex, to shift the dynamic equilibrium in the full-length sequence to favor this structure. We also identified a benzothiophene-2-carboxamide (NSC309874) as a PDGFR- β i-motif-interactive compound. In vitro, GSA1129 and NSC309874 downregulate PDGFR- β promoter activity and transcript in the neuroblastoma cell line SK-N-SH at subcytotoxic cell concentrations. GSA1129 also inhibits PDGFR- β -driven cell proliferation and migration. With an established preclinical murine model of acute lung injury, we demonstrate that GSA1129 attenuates endotoxin-mediated acute lung inflammation. Our studies underscore the importance of considering the effects of point mutations on structure formation from the G- and C-rich sequences and provide further evidence for the involvement of both strands and associated structures in the control of gene expression.

Graphical Abstract

INTRODUCTION

The precise regulation of PDGFR- β expression is essential for organism development and maintenance of homeostasis in adults.^{1–3} Aberrant expression of PDGFR- β and persistent high levels of PDGFR- β signaling are causative factors for a variety of pathologies, including vascular disease, fibrotic disorders, and especially cancer.^{4–8} Accordingly, the PDGFR- β signaling pathway is a validated and important target for the treatment of certain malignant and nonmalignant pathologies.^{1,6,9,10} The two primary therapeutic approaches used to inhibit the PDGFR- β pathway focus on directly targeting the PDGFR- β protein or its cognate ligand.¹ Although these strategies are effective, there is still a need for the development of alternative therapeutic approaches.

In vivo models of PDGFR- β -driven disease show that an attenuating PDGFR- β transcript causes similar effects to targeting the protein, such as decreased tumor growth, survival, angiogenesis, and metastasis.¹¹ One strategy for achieving targeted downregulation of transcript with small drug-like molecules involves ligand-mediated stabilization of G-quadruplexes.^{12–14} These structures are found with high frequency around the transcription start sites of many proto-oncogenes,^{15–17} including MYC,^{12–14} hTERT,¹⁸ PDGF-A,¹⁹ and PDGFR- β .²⁰ Accordingly, if the G-quadruplex has a unique folded topology, it can be targeted more specifically by small molecules to modulate oncogene transcription.^{12–14}

Similar to G-quadruplexes, promoter-based i-motifs can nucleate from duplex DNA in response to negative super-helicity,^{21,22} using the complementary C-rich sequence to form intercalated hemiprotonated cytosine⁺–cytosine (C⁺–C) base pairs.^{23,24} Ex vitro, i-motifs are only stable at an acidic pH and demonstrate increasing single-strand dynamics as pH increases toward neutrality due to their requisite hemiprotonated C⁺–C building block.^{24–28} However, in a recent paper Wright and co-workers have demonstrated that a number of genomic sequences can form stable i-motif structures at neutral pHs.²⁹ In addition, molecular crowding conditions using single-walled carbon nanotubes, which were intended to mimic the cellular environment, demonstrate stable i-motif formation from telomeric sequences at physiological pH.³⁰ Last, Ag⁺ also stabilizes i-motifs in the telomeric sequence at a pH of 7.4.³¹ Although the existence of i-motifs has been known since 1993,²³ their potential biological significance has only recently been investigated.^{21,22,24,32–34} It can be expected that the genomic distribution of i-motifs is similar to that of G-quadruplexes, as they form from the complementary sequence.²⁴ Importantly, a recent report investigating i-motif formation from the BCL2 promoter demonstrates that hnRNP LL recognizes and binds to the lateral loops of the BCL2 i-motif to modulate transcription.³² Furthermore, two small molecules were identified that exclusively bind to the BCL2 C-rich sequence in either the i-motif (IMC48) or the flexible hairpin (IMC76) conformation and exert their activity by modulating the amount of the i-motif available for hnRNP LL binding.³³ In addition, studies on the MYC promoter demonstrate that hnRNP K recognizes the constrained loops in the i-motif. Furthermore, the amount of negative superhelicity induced by SP1-induced transcription activation determines the extent of unwinding of the duplex that controls access of hnRNP K to the three KH binding domains in the fully unwound promoter sequence.²²

Here we describe the critical importance of taking into consideration the effects of point mutations on DNA secondary structure formation from the C-rich sequence in the PDGFR- β NHE, in addition to their effects on the G-quadruplexes on the opposite strand, in order to fully understand the effect on transcription silencing or activation by these promoter sequences.

We have previously shown that an upstream region of the PDGFR- β proximal core promoter (–166 to –132 base pairs) contains a GC-rich NHE that can undergo strand separation in response to negative supercoiling to form at least four G-quadruplexes, referred to here as 5'-end, 5'-mid, 3'-mid, and 3'-end.^{20,35} Moreover, the G-quadruplex-stabilizing ligands telomestatin^{36,37} and TMPyP4^{14,38} exhibit differential binding to the G-quadruplex structures formed from this 34-base-pair region to negatively affect PDGFR- β promoter activity.^{20,39} Results from these studies suggest an opportunity to develop optimized G-quadruplex ligands that demonstrate selective G-quadruplex targeting to downregulate PDGFR- β expression and suppress PDGFR- β -related phenotypes.^{20,39} Therefore, we further investigated the formation and biological role of the different G-quadruplexes in the PDGFR- β NHE with the ultimate goal of demonstrating an alternate and effective strategy for molecularly targeting the PDGFR- β pathway. To further demonstrate the importance of taking into account both strands in the NHE, we also identified, through fluorescence resonance energy transfer (FRET) screening, a small molecule that targets the i-motif to most likely destabilize binding of hnRNP K, which also downregulates PDGFR- β transcription.

RESULTS

The GC-Rich NHE Located in the PDGFR- β Core Promoter Is Responsible for Over 60% of Gene Activity and Consists of a Dynamic Equilibrium of Four Unique and Distinct G-Quadruplexes

The PDGFR- β NHE Is a Critical Promoter-Based Regulatory Element for PDGFR- β Transcription Activity—Within the –166- to –132-base-pair region of the PDGFR- β core promoter is a GC-rich NHE consisting of seven polyguanine/polycytosine tracts that are proposed to form an equilibrating set of at least four G-quadruplexes from overlapping sequences (Figure 1A).^{20,35} To directly evaluate the biological importance of this dynamic element for PDGFR- β promoter activity, two luciferase constructs were prepared: WT-NHE and NO-NHE (Figure 1A). These constructs were each transiently cotransfected with pRL-TK into HEK293 cells, and luciferase activity was evaluated 24 h post-transfection. Promoter activity of these constructs is shown in Figure 1B as normalized luciferase expression relative to the WT-NHE construct. Both constructs showed increased luciferase activity relative to the negative control pGL3 luciferase reporter vector (data not shown), indicating that the PDGFR- β promoter fragments were indeed functional in the respective vectors. Loss of the NHE results in a significant decrease (approximately 60%, *** $p < 0.001$) in luciferase activity for the NO-NHE construct relative to the WT-NHE (Figure 1B). These findings demonstrate that the NHE is an important positive regulatory element for PDGFR- β promoter activity, because loss of this element leads to a negative effect. This is presumably because removing this WT-NHE region leads to loss of potential

binding sites for transcription factors. At the duplex level, SP1 binding sites activate transcription, while at the secondary DNA structure level, i-motifs can also activate transcription.^{32,33,40–43}

Circular Dichroism (CD) Spectra of G-Quadruplex Sequences Indicate the Formation of Four Stable G-Quadruplexes

—We have shown, using a polymerase stop assay, that the G-rich strand of the PDGFR- β NHE produces four distinct stop products, which suggests the formation of at least four stable G-quadruplexes from this sequence.^{20,44} Guided by our previous studies, we further characterized the formation of these four equilibrating G-quadruplexes by subdividing the full-length (FL Pu41) NHE into four putative G-quadruplex-forming subfragments, which we refer to as the 3'-end (R1–R4), 3'-mid (R2–R6), 5'-mid (R3–R6), and 5'-end (R4–R7) (Figure 1A).^{20,44} It is well established that CD spectroscopy can be effectively used to determine G-quadruplex formation and stability while also providing some useful structural insight, such as strand directionality.^{45,46} Accordingly, the FL Pu41 and the four subfragments were evaluated for G-quadruplex formation by CD spectral and thermal analyses in the absence (data not shown) and presence of buffered K⁺. We have previously reported on the formation and NMR-based folding pattern of a G-quadruplex from the 5'-mid sequence.³⁵ All four sequences demonstrated the characteristic CD spectra, indicating the formation of G-quadruplex DNA structures that are thermally stable at or above physiologically relevant temperatures.^{45,46} The FL and constituent sequences all adopted a predominantly parallel folding pattern, as characterized by a distinct maximum positive molar ellipticity between 260 and 265 nm with a slight shoulder around 290 nm (Figure 1C).^{45,46} The hierarchy of thermal stability for these G-quadruplexes, ordered from most stable to least stable, is 5'-mid (66.1 °C), 3'-mid (57.3 °C), 3'-end (55.6 °C), and 5'-end (52.6 °C) in 50 mM K⁺-containing solution. Somewhat surprisingly, the 3'-end G-quadruplex, which contains the 3' GGA sequence, is not the least stable isoform, even though it appeared to contain a “broken” tetrad (AGGG rather than GGGG). Collectively, these findings confirm our initial postulate that the PDGFR- β NHE is highly complex and can exist as an equilibrating set of at least four distinct overlapping G-quadruplexes.^{20,44}

DMS Footprinting of Each G-Quadruplex Sequence Reveals a Unique Pattern of Protection

—We further investigated structural aspects of the FL and constituent G-quadruplex-forming sequences through chemical footprinting analysis using DMS. The DMS footprinting pattern of the three sequences that include the run of two guanines in R5 (3'-mid, 5'-mid, 5'-end) all showed protection of one or both of these guanines to DMS (Figure 1D(b–d)). The 5'-mid showed the expected DMS footprint based on the known snap-back structure previously determined by NMR.³⁵ The 3'-mid can also be predicted to have a snap-back structure, but this would need to be determined by NMR studies. The DMS footprint for the 3'-end sequence (Figure 1D(a)), which has an adenine as the 3'-most purine in R1, displayed clear protection of guanines in R2, R3, and R4, while R1 is only partially protected, suggesting that this end is frayed, presumably because of instability produced by the 3'-terminal adenine. Finally, the DMS footprinting pattern of the FL sequence is shown in Figure 1D(e). While this is presumably a composite of all four G-quadruplexes, it is most indicative of the 5'-mid and 3'-mid DMS folding patterns. This was

based on a comparison of each of the DMS protection patterns of the constituent G-runs R1–R7 to the corresponding protection patterns of the four different G-quadruplexes that can be formed. For example, the 5′-mid protection pattern is found in R3, R5, and R7, whereas the 3′-mid protection pattern is found in R2 and R4.

Formation of the 3′-End G-Quadruplex Appears To Utilize the Adenine To Contribute to the Thermodynamic Potential and Repressor Function of This Isoform

Investigation into the structure of the 3′-end G-quadruplex began by evaluating the importance of the 3′-adenine in R1 using the three different point-mutated oligomers (A-to-G, A-to-T, A-to-C) by CD (Figure 2A). Our results show that 3′-end G-quadruplex stability is sensitive to mutation of the 3′-adenine, as purine-to-pyrimidine (A-to-T and A-to-C) mutations caused a decrease in stability of approximately -5°C , while the purine-to-purine (A-to-G) mutation markedly increased G-quadruplex stability by $\sim 16^{\circ}\text{C}$ (Figure 2A). These same R1 mutations made to the FL sequence resulted in a similar trend in stability (data not shown), suggesting the importance of this G-quadruplex in the equilibrium mixture, as also suggested by the DMS footprinting of the FL sequence. Since the A-to-G mutation caused a dramatic increase in the stability of the 3′-end G-quadruplex, we then asked how that would affect the repressor activity of the FL Pu41. Surprisingly, we found that this single-base mutation resulted in an increase in the luciferase relative to the control (Figure 2B). So, despite the fact that this A-to-G mutation resulted in an increase in stability of the 3′-end G-quadruplex, the overall effect of this mutation is loss of repressor activity, suggesting that this particular mutant G-quadruplex and corresponding i-motif (see later) might play an important role in the control of PDGFR- β gene expression. In contrast, the mutants in the R5 G-to-T contained in the single run of two guanines showed no significant effect on luciferase activity, and likewise the R6 G-to-T mutation in the run of seven guanines showed the expected reduction in luciferase activity (Figure 2B). In subsequent studies, we demonstrate that the apparent conflicting results obtained in the case of the R1 A-to-G mutation can be rationalized if we also take into account the effect of the corresponding R1 T-to-C mutation on the two equilibrating i-motifs on the complementary strand, which we show are involved in transcription activation of PDGFR- β (see later).

To determine whether the 3′-adenine in the 3′-end G-quadruplex was involved in quasi base stacking by substituting for a guanine, we designed two oligomers with this base substituted for 2-aminopurine (2AP) (3′-end Pu22 2AP and FL Pu41 2AP) and monitored their fluorescence in the presence of increasing K^{+} concentration (0–100 mM) (Figure 2C, left and right panels, respectively). This experiment was designed based on the observation that when the adenine analog 2AP is incorporated into secondary DNA structures such as G-quadruplexes, the fluorescent properties are diminished.^{47–49} Our results show a significant $[\text{K}^{+}]$ -dependent decrease in 2AP fluorescence for the 3′-end Pu22 2AP (50–100 mM $[\text{K}^{+}]$) and FL Pu41 2AP (100 mM $[\text{K}^{+}]$). These results suggest that at physiologically relevant $[\text{K}^{+}]$ the R1 3′-AP is likely recruited to form the 3′-end G-quadruplex. However, how this 3′-adenine is utilized in the formation of the 3′-end G-quadruplex in both FL (Pu41) and constituent (Pu22) sequences is not clear from this experiment. It is noted that 2AP shares more structural similarity with guanine than adenine; therefore, 2AP may fit better in a tetrad formation than adenine.

In an attempt to distinguish the relative effect of the adenine substitution for the guanine in the 3'-tetrad in the WT, in contrast to the constraints imposed by a 2:1:1 loop structure, which is expected to produce some instability in the R1 A-to-G mutant Pu22, we carried out comparative DMS footprinting on the two sequences (Figure 2D). Under physiologically relevant K^+ concentration, the 3'-end R1 A-to-G mutant Pu22 (Figure 2D, right) showed dissimilar patterns of protection in R1 and R4 when compared to the WT Pu22 (Figure 2D, left). The incomplete protection against DMS cleavage in all four guanine runs that make up the 2:1:1 loop structure in the mutant sequence is probably because of the two single-base loops in this G-quadruplex. For R1, the WT sequence shows less protection than for the mutant sequence, probably because of the instability of the WT sequence relative to the mutant sequence. This suggests some perturbation of structure due to the GGA sequence at the 3'-end of the G-quadruplex. As expected, the most pronounced differences between the WT and mutant sequences are in R1, where clearly both guanines are cleaved in the GGA sequence. Overall, these DMS footprinting and 2AP results suggest that the 3'-terminal adenine plays an important but as yet undefined role in formation of an unusually folded G-quadruplex. The unique effects of the 3'-terminal adenine in the formation of an unusual 3'-AGGG tetrad in the 3'-end G-quadruplex along with the unexpected increase in luciferase expression of the R1 A-to-G mutation (Figure 2B) led us to ask what effect the corresponding R1 T-to-C mutation would have on the i-motifs that we suspected would form on the opposite strand.

The C-Rich Strand Forms an i-Motif, and While the R1 T-to-C Mutation Results in a Large Increase in Stability, the Other Mutations Result in Lower Stabilities

CD spectroscopy is commonly used to assess i-motif DNA secondary structure formation based on the distinctive spectral signature associated with hemiprotonated C^+-C base pairs, which is characterized by a maximum positive molar ellipticity (λ^{MAX}) within 280–288 nm and a negative molar ellipticity (λ^{MIN}) within 260–267 nm.^{50–53} CD spectral and thermal analyses were initially carried out on the PDGFR- β NHE Py41 WT sequence over a pH gradient to determine i-motif structure formation and pH dependence and to show that the Py41 WT sequence at pH 6.0 has a spectral signature with a λ^{MAX} at 286 nm and a λ^{MIN} within 260–267 nm, which is indicative of i-motif formation (Figure 3A(a)).⁵⁴ CD thermal melt analysis was carried out on the WT sequence at its transition pH (pH 6.6) (Figure 3A(b)) and pH 6.0. The results from these studies show that under more acidic conditions, the thermal stability (T_m) is increased 7 °C from pH 6.6 (28.5 °C) and pH 6.0 (35.5 °C), indicating that the WT i-motif T_m is pH-dependent (Figure 3B).

Since the effect on PDGFR- β expression of the mutations on the G-quadruplex-forming strand had produced apparently contradictory results (Figure 2B), we compared the T_m 's and λ^{MAX} of the three different mutants shown in Figure 3A(c) with the WT (see Figure 3A,B). Compared to the spectral signature of the WT sequence, the R1 T-to-C has a more pronounced λ^{MAX} at 286 nm and λ^{MIN} within 260–267 nm, whereas the R5 C-to-A mutant has a decreased λ^{MAX} at 282 nm and a λ^{MIN} within 260–267 nm, and R6 C-to-A has a decreased λ^{MAX} at 280 nm and a λ^{MIN} within 260–267 nm (Figure 3A(a)). Together, these results suggest that the R1 mutation causes an increase in i-motif formation, the R5 mutation modestly affects structure formation, and the R6 mutation dramatically decreases the i-motif

structure formed compared to the WT sequence. The results in Figure 3A(b) show that the R1 mutant has no effect on transition pH, while the structures formed by the R5 and R6 mutant sequences transition to unordered structures under more acidic conditions at pH 6.4 and 6.2, respectively. Last, the change in T_m (ΔT_m) was calculated relative to the WT i-motif. Our results in Figure 3B show that the R1 mutation markedly increases i-motif thermal stability ($\Delta T_m = +10.1^\circ\text{C}$ and $+13.2^\circ\text{C}$), the R5 mutation only slightly decreased thermal stability ($\Delta T_m = -1.8^\circ\text{C}$ and -0.2°C), and the R6 mutation dramatically decreased the thermal stability of the structure ($\Delta T_m = -8.1^\circ\text{C}$ and -6.5°C). Overall, these results indicate that creating an additional 5'-CCC-3' tract with the R1 mutation allows for the formation of a more stable i-motif, while the R6 mutation, which disrupts a critical run of seven cytosines, has dramatic effects on i-motif formation. As expected, the R5 mutation, which disrupts a cytosine run presumably unimportant for i-motif formation, had negligible effects on i-motif stability. In subsequent studies we concentrated our efforts on comparing the R1 T-to-C mutant, which unexpectedly had increased PDGFR- β gene expression relative to the WT.

The R1 T-to-C Mutant NHE in the Py41 Forms a Much More Defined i-Motif Than the WT Sequence, Which Appears To Form Multiple Folded Topologies

The R1 T-to-C mutation creates a 5'-CCC-3' run from the natural sequence 5'-TCC-3'. This has some similarities to the 3' A-to-G mutation in the 3'-end G-quadruplex where a 3'-GGG-5' run is created. In parallel with the findings on the purine-rich strand, we suspected that this newly created 5' cytosine run would be involved in the formation of a more stable i-motif than can be formed from the WT sequence (Figure 4A). To characterize the importance and utilization of individual cytosines in i-motif formation and stability, we designed a series of oligomers that contain single C-to-T point mutations in cytosine runs R1–R7 for both the WT and R1 mutant sequences (Figure 4A), and each of these mutated sequences was evaluated by CD thermal melt analysis. Each of the individual C-to-T mutations that are involved in C⁺–C base pairing in the i-motif structure should show a decrease in T_m relative to the parent sequences. The results from the WT sequence show that only three cytosine runs (2, 3, and 7) show protection, while the remaining runs show minimal decrease in T_m relative to the parent sequences (Figure 4A, left). This data suggests that at least two equilibrating i-motifs can be formed from the WT Py41 through the differential use of cytosine runs. In contrast, the comparative T_m 's relative to the parent sequences of the C-to-T point mutations in the i-motif formed from the R1 mutant Py41 sequence show this mutated sequence uses four much more defined runs of cytosines (R1 (C0–C2), R3 (C8–C10), R6 (C21–C23), and R7 (C32–C34)) for i-motif formation (Figure 4A, right). Using this data we can predict the putative folding patterns for the R1 mutant i-motif, which is formed by six C⁺–C hemiprotonated base pairs and has loops of 5:10:8 (Figure 4B).

Rationale for Drug Targeting the 3'-End G-Quadruplex and the 5:10:8 Loop i-Motif To Repress PDGFR- β Gene Expression

Although the results of mutational conversion of the 3'-AGG-5'/5'-TCC-3' to the 3'-GGG-5'/5'-CCC-3' sequence show stabilization of both the 3'-end G-quadruplex and the 5:10:8 loop i-motif, the functional consequences on transcription activation are probably

different. In the case of the R1 A-to-G mutant, which forms a conventional G-tetrad, this form probably does not exist naturally, and we suspect that for repression to occur, the WT broken-tetrad 3'-end G-quadruplex carries special recognition properties that are required for binding of a repressor protein that will specifically recognize this G-quadruplex and lower transcription. On the other hand, the i-motif that is favored by the R1 T-to-C mutation is one of a number of equilibrating i-motifs that exist naturally and may be the one favored for binding of the protein that results in transcription activation. If this is true, we would aim to identify compounds that selectively bind to the natural 3'-end G-quadruplex by recognizing the broken tetrad and thus mimic the selectivity of the putative repressor protein, while for the i-motif-forming strand, we would aim to identify compounds that prefer to bind to the i-motif recognized by the activator protein, which we suspect is the mutant i-motif. In the latter case, we would want to identify compounds that disrupt the binding of the activator protein to the mutant form of the i-motif. We therefore set out to identify compounds that would bind selectively to the 3'-end G-quadruplex containing the broken tetrad (AGGG) over the other G-quadruplexes in the PDGFR- β promoter that have snap-back structures and that also have regular tetrads containing four guanines, together with compounds that would bind to the R1 T-to-C mutant i-motif. The compounds that specifically recognize the naturally occurring 3'-end G-quadruplex should mimic the effect of the naturally occurring repressor protein. On the other hand, we aim to identify compounds that bind preferentially to the mutated form of the i-motif that would prevent the binding of the naturally occurring activator protein and thus also inhibit transcription activation.

The 3'-End G-Quadruplex Is Identified as the Target for a Compound That Will Produce Transcriptional Silencing of PDGFR- β Expression and Subsequent Identification of an Ellipticine That Binds Selectively to This G-Quad-ruplex

The Ellipticine NSC338258 Demonstrates Selectivity for the PDGFR- β 3'-End G-Quadruplex—Our previous studies have shown that of the two structurally distinct groups of compounds we have tested (porphyrins and telomestatin), only telomestatin produced a decrease in PDGFR- β , and this compound was unique in that it appeared to bind selectively to the 3'-end G-quadruplex.²⁰ We were also drawn to the 3'-end G-quadruplex as the molecular target because it has a unique 3'-end GGA sequence, in contrast to the other three G-quadruplexes, which all seemed likely to have snap-back structures, providing quite different molecular recognition properties for a protein or drug that would selectively bind to this G-quadruplex. Last, it appeared from the data in Figure 2B that mutation of the 3'-end sequence from a GGA to a GGG resulted in loss of suppression of luciferase, suggesting the critical involvement of the 3'-end G-quadruplex in repression of PDGFR- β expression.

We had previously shown that the ellipticine derivative NSC338258 binds almost as well to the PDGFR- β NHE as to the MYC promoter G-quadruplex.¹³ In a subsequent screen we then tested the relative affinity of NSC338258 for the different NHE subfragments (Figure 5). Our data show that NSC338258 binds preferentially to the 3'-end G-quadruplex at approximately 1 mol equiv and increases its thermal stability by 4 and 9 °C at 1 and 2 mol equiv, respectively. We also examined the effect of NSC338258 on the CD spectra of the four different constituent G-quadruplexes (Figure 5B) and the FL (Supporting Information,

Figure S2). Consistent with the results in Figure 5A, only in the case of the 3'-end G-quadruplex did we see a significant change. On the basis of the large selective increase in molecular ellipticity at just one equivalent of added drug, it appears that the most favored site for binding of the ellipticine-based NSC338258 is the 3'-end G-quadruplex.

The Design and Synthesis of Ellipticine Analogs To Increase Selectivity and Potency for PDGFR- β Transcription Inhibition

Our approach to the modification of the lead molecule (ellipticine) was to probe the different interaction sites with the G-quadruplex target. Our interaction model consisted of stacking interactions between pyridocarbazole and the G-tetrad and interactions of side-chain substituents with the G-quadruplex loop regions. Our lead modification strategy consisted of selecting side-chains with varying linker lengths and basicity at two different positions on the ellipticine scaffold. On the basis of this strategy, we then designed and synthesized the compounds listed in Table 1, each containing different side-arms as well as different side-chain lengths with basic substituents. The synthesis of these compounds was carried out as shown in Schemes 1–3.

GSA1129 was synthesized starting from methoxyellipticine, as shown in Scheme 1. 9-Methoxyellipticine was treated with bromochloropropane and potassium *tert*-butoxide, with *tert*-butanol as a solvent, to obtain 6-chloropropyl-9-methoxyellipticine. GSA1129 was obtained from the reaction of 6-chloropropyl-9-methoxyellipticine and 1-methylpiperazine.

As shown in Scheme 2, other ellipticine analogs were prepared by alkylation of ellipticine or methoxyellipticine using corresponding alkyl bromides or chlorides with potassium *tert*-butoxide in *tert*-butanol solvent.

GSA1125 was synthesized by alkylation of 9-hydroxyellipticine using dimethylaminopropyl chloride and potassium *tert*-butoxide in *tert*-butanol, as shown in Scheme 3.

The cytotoxicity results for the ellipticine analogs are shown in Table 1. On the basis of just the cytotoxicity against the human kidney cell line HEK293, it is difficult to make any firm conclusions about structure–activity relationships. For this reason, we moved all nine compounds directly into the screen, in which we examined their selectivity for binding to the 3'-end G-quadruplex versus the other three G-quadruplexes in the PDGFR- β promoter element.

GSA1129 Shows Selectivity for the 3'-End G-Quad-ruplex in the PDGFR- β Promoter Element and Requires the NHE for Decreases in Luciferase Activity

Using the different oligomers previously used for the four G-quadruplexes found in the PDGFR- β promoter, we added 2 equiv of the nine different ellipticines and determined the T_m for each ellipticine analog/oligomer. The results shown in Figure 6A demonstrate that GSA1125 and GSA1129 show the best selectivity for the 3'-end G-quadruplex over the other three G-quadruplexes.

To further differentiate the selectivity of the different ellipticine analogs in a cell-based system, we compared the luciferase activity in two different constructs, one containing the

WT (WT-NHE) and a second in which the G-quadruplex-containing element was eliminated (NO-NHE). In some cases, such as NSC338258 (Figure 6B(a)), there was a moderate degree of selectivity that correlated with the selectivity we saw in the T_m measurements shown in Figure 6A; however, in others, such as GSA1125 (Figure 6B(f)), which appeared to show selectivity in the T_m measurements, there was no selectivity between the empty vector and the WT. Significantly, GSA1129 showed an impressive degree of selectivity at both the 50% IC₅₀ dose and the IC₅₀ dose (Figure 6B(g)) and was therefore chosen as the compound to move forward.

GSA1129 Selectively Stabilizes the 3'-End G-Quad-ruplex in the PDGFR- β Promoter NHE and Moves the Equilibrium to This G-Quadruplex

To examine the selectivity of GSA1129 for the different G-quadruplexes formed in the PDGFR- β promoter NHE, we performed a Taq polymerase stop assay using increasing concentrations of GSA1129 in the absence of K⁺ (Figure 7A). Our results show that in the absence of potassium GSA1129, at a concentration of 2.5 μ M, was able to selectively induce the formation of the 3'-end G-quadruplex. It is important to note that while the mid-stop product was also seen, this was not induced by GSA1129 because it was present even in the absence of GSA1129, whereas the 3'-end product only appeared in the presence of GSA1129 (Figure 7A). The selectivity of GSA1129 for the different G-quadruplexes formed in the PDGFR- β promoter NHE was further examined using DMS footprinting of the WT FL Pu41 in the absence of K⁺ (Figure 7B). GSA1129 was able to selectively induce the formation of the 3'-end G-quadruplex, as shown by the clear protection of the four 3'-end guanine runs (R1–R4) in the DMS footprinting. Taken together, our results show that GSA1129 was not only able to selectively stabilize the 3'-end G-quadruplex, it can also induce the formation of this G-quadruplex among the different G-quadruplexes formed in the PDGFR- β promoter NHE. To illustrate the importance of the 3'-AGG in the WT 41-mer in comparison to the 3'-GGG found in the R1 A-to-G mutant for binding of GSA1129, we compared the increased melting points of these two oligomers in the presence of this drug. The results shown in Figure 7C clearly show that the WT undergoes a much greater increase in T_m than the R1 mutant (9 °C versus 3 °C).

While GSA1129 Has Selectivity for Targeting the 3'-End G-Quadruplexes over the G-Quadruplexes in the PDGFR- β Promoter NHE, It Does Not Appear To Show Complete Selectivity over G-Quadruplexes in Other Promoters

While we had demonstrated selectivity for the 3'-end G-quadruplex over the other G-quadruplexes in the PDGFR- β promoter NHE, this does not necessarily carry over to G-quadruplexes in other promoters. To initially test this idea, we used the c-Myc promoter, which has a parallel G-quadruplex, as a test case because the initial hit compound NSC338258 binds equally well to this G-quadruplex and the 3'-end G-quadruplex.¹³ Somewhat surprisingly, we found that although GSA1129 has selectivity for the 3'-end G-quadruplex in the PDGFR- β promoter, it, like its parent compound NSC338258, binds equally well to this G-quadruplex and the 3'-end G-quadruplex (Figure 8A). To determine whether this lack of selectivity is also exhibited more generally at the transcript level, we examined the effect on a variety of promoters known to have G-quadruplexes in their promoters. The results show that while there was good selectivity in comparison to most of

the promoters, c-Myc, which contains a parallel G-quadruplex, showed little if any selectivity (Figure 8B). It is also possible that the apparent lack of selectivity for those promoters where the effects are less pronounced is due to indirect effects related to downstream effects from PDGFR- β and c-Myc.

Identification of a Small Molecule That Binds Selectively to the i-Motif in the PDGFR- β NHE To Destabilize the Folded Form

FRET probes of the WT and R1 mutant sequences were designed as previously described³³ for use in medium-throughput screening to identify small molecules that interact with the WT or R1 mutant i-motif-forming sequences. Using these FRET probes, we screened the NCI Diversity Set III and Mechanistic Set II (2869 compounds) for compounds that either stabilize or destabilize the i-motif DNA secondary structure, as indicated by a decrease or increase in fluorescence intensity, respectively (Figure 9A(a,b)). Compounds that decreased fluorescence by approximately 50% or increased fluorescence by approximately 200% were considered potential hits. This cutoff provided a hit rate of 1.1% (WT) and 0.8% (R1 mutant) for i-motif-stabilizing compounds and a hit rate of 0.5% (WT) and 0.7% (R1 mutant) for i-motif-destabilizing compounds. To eliminate compounds that also interact with the PDGFR- β G-quadruplex, all hits were further screened by CD spectral and thermal melt analysis using the FL Pu41. Through these screening methods, we identified the benzothiophene-2-carboxamide compound NSC309874 (Figure 9A(a)), which increased the FRET signal from the WT and R1 mutant by 170% and 218%, respectively (Figure 9A(c)). We next investigated the effects of NSC309874 on the WT and R1 mutant PDGFR- β i-motifs as well as the PDGFR- β G-quadruplex-forming sequence using CD spectral and thermal melt analysis. The spectral results for the WT i-motif in Figure 9B(a) show that an increase and shift in λ^{MAX} (from 286 to 279 nm) is associated with increasing concentrations of NSC309874, while there is little change in the CD of the R1 mutant (Figure 9B(b)). In each case there was a small increase in T_m . These data suggest that NSC309874 may preferably bind to the R1 mutant rather than the WT conformation since less perturbation of structure is required. Similar CD experiments were carried out with the G-quadruplex-forming sequence and NSC309874 that showed negligible interaction (Figure 9B(c)). These data show that NSC309874 selectively interacts with DNA secondary structures formed from the C-rich sequence of the PDGFR- β NHE.

Targeted Stabilization of the 3'-End G-Quadruplex in the PDGFR- β Promoter with GSA1129 Downregulates PDGFR- β Transcription and Inhibits Cell Proliferation and Motility

GSA1129 Potently Induces Differential Cytotoxicity in the Neuroblastoma Cell Lines SK-N-SH and SH-SY-5Y—We evaluated cell-line sensitivity to GSA1129 and methoxyellipticine with the MTS assay using an in-house panel of PDGFR- β -expressing cell lines (Figure 10A). Our results show that GSA1129 exhibits the greatest amount of cytotoxicity in the neuroblastoma cell lines SK-N-SH ($1.2 \pm 0.8 \mu\text{M}$) and SH-SY-5Y ($1.8 \pm 0.6 \mu\text{M}$), which have elevated levels of PDGFR- β mRNA (Figure 10A). Importantly, both cell lines are approximately 12-fold more sensitive to GSA1129 compared to methoxyellipticine (Figure 10A). These results led us to suspect that the differential cytotoxicity observed for GSA1129 is caused by a decrease in PDGFR- β transcript.

GSA1129 Reduces Both Nascent and Mature PDGFR- β RNA Transcript in the SK-N-SH Cell Line—To directly assess the effect of GSA1129 on PDGFR- β transcription, we utilized two sets of primers that specifically amplify either the nascent or mature PDGFR- β RNA transcript by quantitative real-time PCR. Our results show that GSA1129 decreases both nascent and mature PDGFR- β transcripts in a dose- and time-dependent manner in SK-N-SH (Figure 10B(a,b)) and SH-SY-5Y (Figure 10B(c,d)). Importantly, only SK-N-SH showed an equivalent decrease of both nascent and mature PDGFR- β transcript (Figure 10B(a,b)), which suggests that there are no alternate stabilities between nascent and mature PDGFR- β transcripts in this cell line.

GSA1129 Inhibits PDGFR- β -Related Phenotypes at Sub-cytotoxic

Concentrations in SK-N-SH—It is well-known that PDGFR- β signaling activates potent signaling cascades that are responsible for promoting cell survival, growth, proliferation, and motility.^{1–3} Accordingly, studies have shown that targeting PDGFR- β in neuronal-derived tumor cells inhibits proliferation and cell migration.^{55–59} To determine if targeted downregulation of PDGFR- β transcription by GSA1129 decreases these mitogenic signals in SK-N-SH, we evaluated cell viability and proliferation over a 96-h time course (Figure 10C). Our results show that IC₅₀ concentrations prevent cell proliferation and lead to a greater than 75% loss of cell viability by 48 h (Figure 10C(d)), while 50% (Figure 10C(e)) and 25% (Figure 10C(f)) concentrations inhibit cell proliferation similar to the anti-PDGFR- β siRNA-positive control through 48- and 96-h time points, respectively (Figure 10C(c)). These results show that targeted knockdown of PDGFR- β transcript by either siRNA or GSA1129 at 25% or 50% IC₅₀ concentrations inhibits cell proliferation.

Scratch assays are commonly used to evaluate the efficacy of compounds that inhibit PDGFR- β -mediated cell migration.^{60,61} We evaluated the ability of subcytotoxic concentrations of GSA1129 (300 nM) to inhibit cell migration using the Cell Comb scratch assay (Millipore, MD) (Figure 10D). In the presence of GSA1129, cell migration is inhibited through 48 h. Collectively, these results show that targeting the 3'-end G-quadruplex in the PDGFR- β promoter effectively down-regulates PDGFR- β transcript, resulting in inhibition of cell proliferation and motility.

NSC309874 Interacts with the i-Motif in the PDGFR- β NHE To Downregulate Both Luciferase and PDGFR- β Gene Expression

In vitro analysis by luciferase assay in HEK 293 shows that NSC309874 downregulates luciferase activity in a PDGFR- β NHE-dependent manner from the WT-NHE and R1 NHE constructs by approximately 45% and 50%, respectively (Figure 11A(b,c)). In contrast, NSC309874 does not inhibit transcription from the normal B-DNA-containing luciferase constructs, as evident from the negative results using the empty NO-NHE construct (Figure 11A(a)). These data suggest that NSC309874 elicits its transcriptional inhibitory activities through binding to the PDGFR- β i-motif. To extend these findings, the neuroblastoma cell line SK-N-SH, which overexpresses PDGFR- β (see before), was treated with 0.5 IC₅₀ (50 μ M) and IC₅₀ (100 μ M) concentrations of NSC309874, which reduced PDGFR- β expression by approximately 25% and 50%, respectively (Figure 11B). Taken together, these results

demonstrate that molecularly targeting the PDGFR- β i-motif with NSC309874 downregulates PDGFR- β gene activity.

Drug-Targeting PDGFR- β Transcription in in Vivo Models of Acute Lung Injury

With an established preclinical murine model of acute lung injury,^{62–66} we have recently evaluated multiple therapeutic drug candidates (unpublished results). GSA1129 (10 mg/kg, ip) attenuates endotoxin-mediated (LPS, 1 mg/kg, 18 h) acute lung inflammation. Figure 12A–C depicts protein levels, total white blood cell counts, and polymorphonuclear neutrophil (PMN) percentage in the bronchoalveolar lavage (BAL) that are increased by LPS challenge, and these inflammatory parameters are significantly attenuated by GSA1129 treatment. In addition, similar findings have been made in lung histology analysis: GSA1129 markedly reduces infiltration of inflammatory leukocytes into the alveoli (Figure 12D). In addition, we tested GSA1125 as a control compound, which does not lower transcription from the PDGFR- β NHE (see Figure 6B(f)), for effects on BAL protein and white blood cell count, and as expected, no activity was seen (Supporting Information, Figure S2). These findings strongly suggest that GSA1129 is an effective drug candidate for acute lung inflammation.

DISCUSSION

The PDGFR- β signaling pathway is a validated and important target for the treatment of certain malignant and nonmalignant pathologies.^{1,6,9,10} In vivo studies have demonstrated that down-regulation of aberrant PDGFR- β expression effectively diminishes PDGFR- β -driven tumor phenotypes.¹¹ Despite the importance of PDGFR- β in the pathophysiology of certain malignancies, few strategies exist for targeting the PDGFR- β pathway other than targeting the receptor itself or its cognate ligand.¹

Because the PDGFR- β promoter element involves four overlapping G-quadruplexes, it is quite distinct from that of the MYC and other promoter systems previously characterized. Of the four classes of G-quadruplex-forming promoter elements, the majority that have been characterized belong to promoter elements in Classes I–III.⁴¹ Although these three classes are quite distinct, they share a common feature in that their constituent poly-GC tracts are not shared between multiple isoforms and give rise to a single primary G-quadruplex, as is seen in the VEGF,^{67,68} HIF-1 α ,⁶⁹ RET,^{70,71} and PDGF-A¹⁹ promoters,⁴¹ although there may be complexities we do not appreciate at this point. The Class IV promoter-based G-quadruplex control elements comprise multiple contiguous guanine runs that form a dynamic set of G-quadruplexes from overlapping sequences, as found in the BCL2,^{72,73} ADAM-15,⁷⁴ and PDGFR- β ²⁰ promoters.⁴¹

The PDGFR- β core promoter NHE comprises seven contiguous guanine runs, which nucleate to form a dynamic equilibrating set of at least four G-quadruplexes, and is the most complex Class IV promoter studied to date. Three of the four G-quadruplexes are highly polymorphic and likely depend on a 2 + 1 snap-back structure,³⁵ while the fourth G-quadruplex is unique in that it may have an external GGGA tetrad. The unique folding patterns and topologies of the G-quadruplexes that result from this promoter system provided a basis for suspecting that there may be distinct proteins or compounds that

selectively bind to one or more of these structures.^{75,76} Results from a previous study on the PDGFR- β promoter element using the molecular probes telomestatin and TMPyP4 suggested that targeted stabilization of the 3'-end G-quadruplex had inhibitory effects on PDGFR- β transcription.²⁰

Apparently conflicting results emerged from our initial mutational analysis of the PDGFR- β NHE in which we only considered the consequences of mutation of the G-quadruplex-forming strand. However, considering both the structural and isoform equilibrating effects that a point mutation has on the 3'-end G-quadruplex and the i-motifs, respectively, has allowed us to reconcile these discrepancies. First, the presence of a naturally occurring i-motif in which the 5'-end TCC run of pyrimidines lacks a critical terminal cytosine for stability mimics the effect of the opposite strand where the 3'-end G-quadruplex has the same type of deficiency. This suggested to us that just as the broken-tetrad 3'-end G-quadruplex found in the WT sequence could be uniquely recognized by a protein that might repress PDGFR- β transcription, this TCC-containing isomer of the WT i-motif could also be uniquely recognized and stabilized by a member of the hnRNP family of proteins known to bind to i-motifs in promoters.^{22,32} Indeed, for the PDGFR- β i-motif that forms from the R1 mutant, a more stable i-motif structure is formed that contains two lateral loops having CT elements (CCCT), as shown in Figure 4B. Two such CT elements found in the lateral loops of the i-motif should be recognized by hnRNP K in an analogous fashion to similar CT elements contained in the MYC promoter. Thus, a potential mechanism for transcription activation by hnRNP K is identified. Indeed, knockdown of hnRNP K in SK-N-SH cells using siRNA leads, as expected, to an increase in PDGFR- β transcription (unpublished results).

Collectively, these results demonstrate that the formation of i-motifs and G-quadruplexes in the PDGFR- β NHE is sensitive to point mutations that shift the dynamic equilibrium to favor an active (R1 mutant, favoring i-motif formation) or inactive (R6 mutant, favoring G-quadruplex formation) transcription status. This dynamic interplay of G-quadruplex and i-motif DNA secondary structure formation represents a third case in which a "molecular switch" mechanism for regulating promoter activity has been demonstrated. Specifically, formation of the G-quadruplex(es) represents an "OFF switch," while formation of an i-motif represents an "ON switch," as it can be recognized by transcription activators such as hnRNP K. This now extends the examples of ON/OFF switches that contain G-quadruplexes and i-motifs to three promoter systems, the other being MYC²² and BCL2.³² In these two other cases, mutual exclusivity has been demonstrated between the two secondary DNA structures.

On the basis of the potential ON/OFF switch function for the 5'-end i-motif and the 3'-end G-quadruplex structures, we should be able to identify small molecules that specifically target these unique secondary DNA structures to mimic the effect of the proteins that bind to these structures. However, since from a therapeutic perspective we desired compounds that target both structures, which would turn off PDGFR- β expression while we set out to mimic the effect of the protein binding to the 3'-end G-quadruplex, we needed to find a compound that would prevent binding of the activating protein to the i-motif. Thus, we used two different strategies for identifying compounds that target the two different structures. For

targeting of the G-quadruplex, we started out with a known G-quadruplex-binding compound that had reasonable drug-like properties. However, because there is much less insight into i-motif-binding compounds, we used an established FRET assay so we could identify compounds that disrupt the i-motif, which we predicted would prevent binding of the activating transcription factor and also downregulate PDGFR- β expression.

For targeting the 3'-end G-quadruplex, the initial hit compound NSC338258 was found to bind selectively to the PDGFR- β 3'-end G-quadruplex. This is not entirely surprising, as the ellipticine core possesses certain physiochemical properties, such as a conjugated crescent-shaped π -system,^{39,77,78} which has been shown to target the external tetrads of G-quadruplexes, and an earlier study from our laboratory has demonstrated that NSC338258 directly targets the MYC G-quadruplex to decrease its promoter activity in non-Hodgkin's lymphoma using an exon-specific assay.¹³ In the present study, we demonstrate that enhanced selectivity for the 3'-end G-quadruplex over the other snap-back G-quadruplex structures in the PDGFR- β NHE can be achieved through traditional medicinal chemistry hit-to-lead optimization of the ellipticine scaffold. While GSA1129 could clearly differentiate between the different G-quadruplexes in the PDGFR- β NHE, presumably because of quite different folding patterns of the three upstream G-quadruplexes, which have snap-back structures, in contrast to the 3'-end G-quadruplex, which has an AGGG tetrad, at the same time there was no selectivity for other G-quadruplex-containing promoters, such as c-Myc. Thus, although we were able to knock down PDGFR- β transcription, we cannot claim that other G-quadruplex promoter-containing genes were not also repressed. This is likely to be a common feature of compounds such as ellipticines that bind to external tetrads in G-quadruplexes.

For targeting the 5'-end i-motif to decrease PDGFR- β gene expression, we used insight we had previously obtained from targeting the BCL2 i-motif system.^{32,33} These studies found that enhancing the stability of the BCL2 i-motif using IMC48 increased BCL2 transcription, while destabilizing by IMC76 produced the opposite effect. In this case, there is an equilibrating folded i-motif and a flexible hairpin to which IMC76 binds, decreasing BCL2 transcription.³³ IMC48 stabilizes the i-motif to which hnRNP LL binds to activate BCL2 transcription.³² Thus, we identified compounds that increased the FRET measurements of the oligomer representing the C-rich strand of the PDGFR- β NHE. The benzothiophene-2-carboxamide compound NSC309874 (Figure 9A(a)) was identified from this screen and evaluated in cancer cell lines that overexpress PDGFR- β .

Genetic studies in vitro and in vivo have shown that PDGFR- β is a key regulator of cell growth, proliferation, and migration.^{1,3,4,79–81} Herein, we show that treating SK-N-SH with GSA1129 or NSC309874, which selectively bind to either the 3'-end G-quadruplex or the 5'-end i-motif, respectively, results in a decrease in PDGFR- β transcription, which is accompanied by reduced cell proliferation and migration in the case of GSA1129. In addition, using a well-established endotoxin-mediated in vivo model for acute lung inflammation, it was found that GSA1129 is well tolerated and of significant utility in diminishing the inflammation associated with inhibition of the PDGFR- β pathway. While there is a clear correlation between the effects we see at the transcriptional level and the

downstream biological effects in cancer cells and acute injury models, it is premature to claim that the effects are mediated exclusively through effects on PDGFR- β transcription.

CONCLUSIONS

The results from this study demonstrate for the first time that to fully understand the effects of mutations on transcription in promoters that can form both G-quadruplexes and i-motifs, both strands must be considered. A parallel conclusion is that G-quadruplexes and i-motifs are both important in controlling gene expression from promoters that can form these secondary DNA structures under transcriptionally induced negative superhelicity. A second major insight from these studies of a promoter sequence that contains multiple runs of guanines and cytosines that can form an equilibrating set of G-quadruplexes and i-motifs is that it is vital to first determine which of these structures is responsible for downregulation of the target gene. Additionally, guanine runs beyond the four necessary to form a G-quadruplex are less likely to be “spare tires,” as proposed recently,⁸² but are more likely to be part of the complete system in which equilibrating G-quadruplex structures exist. In addition, the number of guanine runs must take into account the potential need for the necessary number of cytosine runs required for formation of the i-motif on the opposite strand. Last, we demonstrate for the first time that it is possible to target either the G-quadruplex or the i-motif in the same promoter with distinctly different small molecules to lower gene transcription. An alternative approach is to target the two equilibrating structures found on the C-rich strand (i-motif and flexible hairpin) in the BCL2 system, but in this case this results in opposite effects on transcription by either increasing (i-motif) or decreasing (flexible hairpin) gene expression. These examples therefore extend the opportunities to target gene transcription in these secondary DNA structure systems beyond just the G-quadruplex.

MATERIALS AND METHODS

DNA Oligonucleotides

DNA oligonucleotides were synthesized by Eurofins MWG Operon (Huntsville, AL). Strand concentration for each oligonucleotide was calculated using the Beer–Lambert law ($A = \epsilon lc$). The extinction coefficient for each oligonucleotide was determined by the nearest neighbor method. Since these particular oligonucleotides exhibit higher order structures, absorbance at 260 nm was recorded at 95 °C to ensure the presence of single-stranded DNA.

Circular Dichroism (CD) Spectroscopy

CD spectra were recorded in triplicate at room temperature on a Jasco J-810 spectropolarimeter (Easton, MD) using a quartz cell with an optical path length of 1 mm. The instrument was set at a scanning speed of 100 nm/min, with a response time of 1 s, over a wavelength range of 230–330 nm for G-quadruplex-forming oligos and 230–350 nm for i-motif-forming oligos. Triplicate readings were averaged, smoothed, and baseline-corrected for signal contributions from buffers and/or compounds used. Molar ellipticities for melting curves were recorded at the maximum molar ellipticity for the G-quadruplex (262 nm) or i-motif (286 nm). Monitoring a single wavelength, the sample was heated from 20 to 95 °C at

a rate of 1 °C/min, and molar ellipticities were recorded every minute. Melt curves were plotted as molar ellipticity versus temperature, and T_m values were determined with GraphPad Prism software using nonlinear regression modeling. All T_m values were calculated within ± 1 °C error. All G-quadruplex oligonucleotide stocks were diluted to a final concentration of 5 μ M in 50 mM Tris-HCl, pH 7.4, ± 25 –100 mM K^+ concentration. To induce a G-quadruplex, samples were heated at 95 °C for 10 min and slowly cooled to room temperature. Unless otherwise indicated, all i-motif-forming oligonucleotides were prepared at a 5 μ M strand concentration in a 50 mM sodium cacodylate buffer adjusted to the proper pH. For G-quadruplex and i-motif studies performed with compounds, oligonucleotide stocks were diluted to a final concentration of 5 μ M in 50 mM Tris-HCl with 50 mM K^+ concentration, pH 7.4 (G-quadruplex), or 10 mM sodium cacodylate buffer at pH 6.5 (i-motif), heated at 95 °C for 5 min, then slowly cooled to room temperature before adding molar equivalents of compound or DMSO and incubating 30 min. CD spectra and T_m measurements were performed as described above and compared to DMSO controls. Unless otherwise indicated, the oligo sequences used for CD were as follows: WT Pu41, GCTGGGAGAAGGGGGGCGGCGGGGCAGGGAGGGTGGACGC; 5'-end, GCTGGGAGAAGGGGGGCGGCGGGGCAG; 5'-mid: AAGGGGGGCGGCGGGGCAGGGAG; 3'-mid: AAGGGGGGCGGCGGGGCAGGGAGGGTG; 3'-end: GCGGGGCAGGGAGGGTGGACGC; WT Py41, GCGTCCACCCTCCCTGCCCCGCCGCCCCCCTTCTCCCAGC; R1 T-to-C mutant Py41, GCGCCACCCCTCCCTGCCCCGCCGCCCCCCTTCTCCCAGC; R5 C-to-A mutant Py41, GCGTCCACCCTCCCTGCCCCGAAGCCCCCCTTCTCCCAGC; and R6 C-to-A mutant Py41, GCGTCCACCCTCCCTGCCCCGCCGACACCTTCTCCCAGC.

Purification of DNA Oligonucleotides by Polyacrylamide Gel Electrophoresis

All DNA oligonucleotides used for footprinting experiments were first diluted to a final concentration of 50 μ M in 50 mM Tris-HCl, heated to 95 °C for 10 min, and snap-cooled to 4 °C on ice. DNA oligonucleotides were further purified by 16% denaturing polyacrylamide gel with 7 M urea Tris-Borate-EDTA buffer, and electrophoresis was conducted at 100 V. DNA bands were cut from the gel and soaked in double-distilled water with agitation overnight to elute DNA. DNA was isolated using ethanol precipitation, excess salts were removed using 70% ethanol, and all purified DNA oligonucleotides were resuspended to a final concentration of 5 μ M in 50 mM Tris-HCl $\pm$ 25–100 mM K^+ concentration to a final volume of 90 μ L per reaction.

Dimethylsulfate (DMS) Footprinting

Following purification of DNA, samples were heated to 95 °C for 10 min and slowly cooled to room temperature to induce a G-quadruplex. For drug treatments, the samples were further incubated in different buffers with or without GSA1129 at room temperature for 1 h. DMS reaction was carried out by the addition of 1 μ L of calf thymus DNA (0.1 μ g/ μ L) to the reaction mixture followed by the addition of 1 μ L of 25% dimethylsulfate solution (DMS/ethanol; 1:4, vol/vol) for 5–10 min, and then the reaction was stopped by adding 18 μ L of stop buffer (3 M β -mercaptoethanol/water/NaOAc; 1:6:7, vol/vol). The monomolecular G-quadruplex structure was isolated by 12% native polyacrylamide gel electrophoresis.

conducted at 100 V in the presence of Tris-Borate-EDTA buffer. The different mobility shift DNA bands were cut from the nondenatured gel and soaked in double-distilled water with agitation overnight to elute DNA. DNA was isolated using ethanol precipitation, excess salts were removed using 70% ethanol, and all purified DNA oligonucleotides were resuspended to a final concentration of 1 μ M in 10 mM Tris, 1 mM EDTA, pH 7.5. Piperidine cleavage was performed at 90 °C for 30 min using 30 μ L of 20% piperidine. Following piperidine treatments, DNA samples were completely dried and resuspended with alkaline sequencing gel-loading dye, then applied to a 20% denatured polyacrylamide gel with 7 M urea. Electrophoresis was conducted at 1200 V, and the gel was visualized using a Storm 820 phosphorimager and quantified using ImageQuant 5.1 software (Molecular Dynamics).

Fluorescent 2-Aminopurine (2AP) Studies

2AP-labeled DNA oligonucleotides (Pu22 2AP and Pu41 2AP) were synthesized by Eurofins MWG Operon (Huntsville, AL). Strand concentration for each oligonucleotide was calculated using the Beer–Lambert law. 2AP-labeled oligonucleotide stocks were diluted to a final concentration of 5 μ M in 50 mM Tris-HCl, pH 7.4, $\pm$ 0–100 mM K⁺ concentration (as indicated). To induce G-quadruplex formation, samples were heated at 95 °C for 10 min and slowly cooled to room temperature. 2AP oligonucleotides were excited at 310 nm, and fluorescent emission was recorded at 370 nm using a BioTek Synergy HT microtiter plate reader at 25 °C in experimental triplicate of technical triplicates. Fluorescent emission contributions for the buffer were corrected using no DNA blanks to yield corrected fluorescent emission values. All corrected fluorescent emission values were normalized to their respective 0 mM control to obtain relative fluorescent emission intensities. Emission values were averaged from technical triplicate samples to generate individual experimental replicates. Significance was determined using GraphPad Prism software by one-way ANOVA with a post hoc Tukey.

Cell Assay Products

MTS (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium) kit, Multi-Tox-Fluor Multiplex Cytotoxicity and cell proliferation assay, Caspase-Glo 3/7 assay system, Dual-Glo dual luciferase assay kit, pRL-TK renilla luciferase control plasmid, and pGL3-basic firefly luciferase plasmid were purchased from Promega (Madison, WI). Silencer Select anti-PDGFR- β siRNA, Silencer Select Negative Control #1 siRNA, and Lipofectamine 3000 transfection reagent were purchased from Life Technologies (Carlsbad, CA).

Cell Culture

All cell lines were purchased from American Type Culture Collection (Manassas, VA) and maintained in a 37 °C, 5% CO₂ incubator, in exponential growth, for the duration of experimentation. Unless otherwise specified, all cell lines were cultured in FBS and penicillin/streptomycin-supplemented media as specified by ATCC. Cells were assessed for viability (>90%) by trypan blue exclusion prior to use for experimental purposes.

Fluorescence Resonance Energy Transfer (FRET) Assay

Each FRET probe was gel-purified, placed in a 50 mM sodium cacodylate buffer, pH 6.5, at a concentration of 20 nM, and heated to 95 °C for 5 min and then incubated at room temperature (25 °C) for 1 h to allow for i-motif formation. Samples were prepared in triplicate, and 100 μ L was placed into each well of a black 96-well OptiPlate (PerkinElmer). For the high-throughput screen, a FRET probe (20 nM) was incubated with compounds from the NCI compound libraries (200 nM) for 30 min at pH 6.5. Samples were prepared in single wells according to the NCI plate setup. Following fluorescent measurements, NCI plate maps were used to determine compound identity. Fluorescence measurements were recorded by a Synergy HT microplate reader (BioTek) at 25 °C with 20 nm bandwidths. The excitation and emission wavelengths were set at 495 and 528 nm, respectively. End point fluorescence/quenching was plotted as the average relative fluorescence units of the triplicate wells after correction for background. Positive and negative controls of HCl and NaOH were used to evaluate the effects of acidic pH on the emission spectra of the FRET probes. WT FRET probe: 5'-[FAM]GCGTCCACCCTCCCTGCCCCGCCGCCCCCCCCTTCTCCCAGC[Quencher]-3'; R1 Mutant FRET Probe: 5'-[FAM]-GCGCCACCCCTCCCTGCCCCGCCGCCCCCCCCTTCTCCCAGC[Quencher]-3'.

MTS Colorimetric Cytotoxicity Assay

Cytotoxicity of compounds was determined in biological triplicate of technical triplicate samples using the MTS assay with compound that was diluted over a 5–6 log range in 0.5 log steps. Cells were seeded at a density of 10 000 cells/90 μ L in a 96-well plate format and allowed to equilibrate overnight. The following day, cells were treated with 10 μ L of 10 $\times$ compound to achieve a final concentration range of 0.001–100 μ M; cells were treated with normal media serving as a “not-treated” control. After the desired incubation time (24–96 h) MTS plus PBS (2 mg MTS/mL in PBS with a MTS/PBS ratio of 20:1) was added to the culture media according to the manufacturer’s instructions, and absorbance was read. Absorbance was recorded, following a 2 h incubation, using a BioTek Synergy HT microtiter plate reader at 25 °C at 490 nm. Negative “no-cell” control absorbance wells were prepared with culture media with and without compound only to account for background interference. The data were corrected with the absorbance of compounds in the negative “no-cell” control and normalized to the “not-treated” control to obtain the relative absorbance intensity. Absorbance values were averaged from technical triplicate samples to generate individual biological replicates. IC₅₀ values were determined from the plot of relative absorbance intensity versus compound concentration with GraphPad Prism software using nonlinear regression modeling; data are presented as mean $\pm$ standard error from at least three independent experiments.

Luciferase Reporter Assay

Luciferase reporter constructs were prepared in pGL3-Basic and sequence confirmed by Eurofins MWG Operon (Huntsville, AL). Cells were seeded at a density of 1.5×10^4 cells per well in a 12-well plate and allowed to equilibrate overnight. Cells were cotransfected with the promoter containing pGL-3 Basic (firefly) and pRL-TK (renilla) using

Lipofectamine 3000 (Life Technologies), as recommended by the manufacturer. Cells were allowed to transfect for 8 h and then were treated with desired concentration of compound or an equal volume of DMSO or normal media. Luciferase activity was measured 24 h after treatment with compound using the Stop & Glo luciferase assay system (Promega), as recommended by the manufacturer. Basal luciferase activity was similarly evaluated 24 h after transfection, as recommended by the manufacturer (Promega). Controls for these experiments included negative “no-cell” media with and without compound only and “not-treated” cell controls, which served for later background subtraction. All luciferase experiments were performed in biological triplicate of technical duplicates. Technical duplicate samples were averaged to generate individual biological replicates. The results of the luciferase assay are presented as firefly luminescence relative to renilla luminescence normalized to a no-drug control. Statistical significance ($p < 0.01$) was evaluated using GraphPad Prism software by two-way ANOVA with a post hoc Tukey; data are presented as mean $\pm$ standard error from at least three independent experiments.

Taq Polymerase Stop Assay

The 5'-end labeled DNA primer (5'-TCGACTCTAAGCAAATGCGTCGAG-3') was annealed to DNA template (5'-GCTGGGAGAAGGGGGGGCGGCGGGGCAGGGAGGGTGGACGCGTTAGTCAGACC TCGACGCATTTGCTTAGAGTCGA-3') and subjected to the Taq polymerase stop assay, as previously described.²⁰ Sequencing reactions were carried out using a Thermo Sequenase Cycle Sequencing Kit (Affymetrix).

Quantitative Real-Time PCR (qRT-PCR)

All qRT-PCR experiments were performed in biological triplicate of technical duplicates using untreated wild-type and DMSO treated as controls for basal PDGFR- β expression levels. Total RNA was isolated with a Qiagen RNeasy Kit according to the manufacturer's specifications. Reverse transcription was performed using the QuantiTect reverse transcription kit (Qiagen) with 500 ng of total RNA, as per the manufacturer's protocol. TaqMan probes were used for GAPDH (Hs02758991) and PDGFR- β mature transcript (Hs01019589_m1), and primers for nascent transcript were designed using primer quest from IDT technologies: TGGAAGTGTGCCCACACCAGAAG (Anti-Sense), TCCAGCACTTACCTTGCTGCTGAT (Sense Probe), CTACTC-GAAGAGATGCAGGTTAG (Sense), GAPDH (Hs02758991) PCR amplification. Real-time PCR was conducted using the Rotor-Gene Q (Qiagen), and Ct values were normalized to GAPDH and compared to the untreated controls using the 2^{-Ct} calculation for relative expression. Statistical significance ($p < 0.01$) was evaluated using GraphPad Prism software by two-way ANOVA with a post hoc Tukey; data are presented as mean $\pm$ standard error from at least three independent experiments.

siRNA Knockdown

Silencer Select anti-PDGFR- β siRNA (Life Technologies) and Silencer Select Negative Control #1 siRNA (Life Technologies) were diluted to 50 nM as a final concentration. For the untreated control, only media were used. SK-N-SH cells (1.5×10^4 per well of a 12-well plate) cultured in 10% FBS and 1% penicillin/streptomycin-supplemented RPMI were

treated with anti-PDGFR- β siRNA or Negative Control #1 with Lipofectamine 3000 transfection reagent (Life Technologies) for the time indicated. For determining the effect on PDGFR- β mRNA of the anti-PDGFR- β or Negative Control #1 siRNAs or normal media, a parallel experiment was conducted using SK-N-SH cells seeded at 1.5×10^4 cells/well of a 12-well plate. Cells were harvested at the appropriate time points, and knockdown was evaluated by qRT-PCR as indicated earlier.

MultiTox-Fluor Cytotoxicity/Cell Viability

The MultiTox-Fluor Cytotoxicity Assay simultaneously measures the relative number of live and dead cells in culture by detecting changes in cell membrane integrity. Viable cells were fluorescently detected using the cell-permeable live-cell protease substrate GF-AFC. The cytotoxic cell populations were fluorescently detected using the dead-cell protease substrate bis-AAF-R110, which is not cell permeable. Proliferation is demonstrated as an increase in only the live-cell signal with respect to the control. Live and dead cell protease activity was evaluated in biological triplicate of technical duplicates in a 96-well plate format, as per the manufacturer's specifications. Cells were seeded at a density of 5500 cells/90 μ L in a 96-well plate format and allowed to equilibrate overnight. The following day, cells were treated with siRNA or the desired concentration of compound. Controls for these experiments included negative "no-cell" media with and without compound only and "not-treated" cell controls, which served for later background subtraction and normalization. After the desired incubation time, live and dead cell substrate was added to the culture media according to the manufacturer's specifications, and fluorescence was evaluated using a GloMax microtiter plate reader at 25 °C (Promega). Negative "no-cell" control fluorescence wells were prepared with culture media with and without compound only to account for background interference. The data were corrected with the fluorescence of compounds in the negative "no-cell" control and normalized to the "not-treated" control to obtain the relative live and dead cell fluorescence intensities. Live and dead cell fluorescence was normalized to indicated time-matched controls and relative to the $t = 0$ time point, which was evaluated 24 h post-cell-seeding.

Scratch Assay

Using the Cell Comb scratch assay (Millipore), SK-N-SH cells were plated, grown to confluence, and wounded, according to the manufacturer's specifications. After wounding, detached cells were removed by washing three times with $1 \times$ PBS. Media was replaced with either normal media, normal media with 300 nM DMSO, or normal media with 300 nM GSA1129. At time points 0, 24, and 48 h, cells were visualized and images captured using a Zeiss 510 microscope (Carl Zeiss Micro Imaging).

Model of LPS-Induced Murine Acute Lung Injury

An established murine model^{62–66} of acute lung injury was used to examine the effects of GSA1129 on LPS-induced lung inflammation. Briefly, C57/B6 mice (10–12 weeks of age) were challenged with an intratracheal administration of *E. coli* LPS solution (1.0 mg/kg) or sterile saline via a 20-gauge catheter. Simultaneously, mice received either GSA1129 (10 mg/kg) or same amount of vehicle via ip injection. The animals were allowed to recover for 18 h, and bronchoalveolar lavage (BAL) fluid (by 1 mL of Hank's balanced salt solution)

and lungs were collected for analysis. The BAL fluid was used to measure total protein according to the manufacturer's manual (BCA Protein Assay Kit, Bio-Rad). BAL cell pellets were examined for total number of white blood cells and cell differential analyses, using cytocentrifugation and Diff-Quik staining. Mouse lungs were made into paraffin sections for H&E staining.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Saad Sammani and Carrie Kempf for technical support for the animal experiments. We are grateful to Dr. David Bishop for preparing, proofreading, and editing the final version of the manuscript and figures. This research was supported by the National Institutes of Health (Grant 5R01CA153821 to L.H.H. and D.Y. and Grant P01HL126609 to J.G.N.G.).

References

1. Andrae J, Gallini R, Betsholtz C. *Genes Dev.* 2008; 22:1276. [PubMed: 18483217]
2. Betsholtz C. *Cytokine Growth Factor Rev.* 2004; 15:215. [PubMed: 15207813]
3. Demoulin JB, Essaghir A. *Cytokine Growth Factor Rev.* 2014; 25:273. [PubMed: 24703957]
4. Alvarez RH, Kantarjian HM, Cortes JE. *Mayo Clin Proc.* 2006; 81:1241. [PubMed: 16970222]
5. Bonner JC. *Cytokine Growth Factor Rev.* 2004; 15:255. [PubMed: 15207816]
6. Östman A, Heldin CH. *Adv Cancer Res.* 2007; 97:247. [PubMed: 17419949]
7. Pietras K, Sjoblom T, Rubin K, Heldin CH, Ostman A. *Cancer Cell.* 2003; 3:439. [PubMed: 12781361]
8. Raines EW. *Cytokine Growth Factor Rev.* 2004; 15:237. [PubMed: 15207815]
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. *J Biol Chem.* 2005; 280:13936. [PubMed: 15695519]
10. Heldin CH, Rubin K, Pietras K, Östman A. *Nat Rev Cancer.* 2004; 4:806. [PubMed: 15510161]
11. Furuhashi M, Sjoblom T, Abramsson A, Ellingsen J, Micke P, Li H, Bergsten-Folestad E, Eriksson U, Heuchel R, Betsholtz C, Heldin CH, Ostman A. *Cancer Res.* 2004; 64:2725. [PubMed: 15087386]
12. Brooks TA, Hurley LH. *Genes Cancer.* 2010; 1:641. [PubMed: 21113409]
13. Brown RV, Danford FL, Gokhale V, Hurley LH, Brooks TA. *J Biol Chem.* 2011; 286:41018. [PubMed: 21956115]
14. Siddiqui-Jain A, Grand CL, Bearss DJ, Hurley LH. *Proc Natl Acad Sci U S A.* 2002; 99:11593. [PubMed: 12195017]
15. Huppert JL. *FEBS J.* 2010; 277:3452. [PubMed: 20670279]
16. Huppert JL, Balasubramanian S. *Nucleic Acids Res.* 2005; 33:2908. [PubMed: 15914667]
17. Huppert JL, Balasubramanian S. *Nucleic Acids Res.* 2007; 35:406. [PubMed: 17169996]
18. Palumbo SL, Ebbinghaus SW, Hurley LH. *J Am Chem Soc.* 2009; 131:10878. [PubMed: 19601575]
19. Qin Y, Rezler EM, Gokhale V, Sun D, Hurley LH. *Nucleic Acids Res.* 2007; 35:7698. [PubMed: 17984069]
20. Qin Y, Fortin JS, Tye D, Gleason-Guzman M, Brooks TA, Hurley LH. *Biochemistry.* 2010; 49:4208. [PubMed: 20377208]
21. Sun D, Hurley LH. *J Med Chem.* 2009; 52:2863. [PubMed: 19385599]
22. Sutherland C, Cui Y, Mao H, Hurley LH. *J Am Chem Soc.* 2016; 138:14138.

23. Gehring K, Leroy JL, Guéron M. *Nature*. 1993; 363:561. [PubMed: 8389423]
24. Day HA, Pavlou P, Waller ZA. *Bioorg Med Chem*. 2014; 22:4407. [PubMed: 24957878]
25. Choi J, Kim S, Tachikawa T, Fujitsuka M, Majima T. *J Am Chem Soc*. 2011; 133:16146. [PubMed: 21882887]
26. Dettler JM, Buscaglia R, Cui J, Cashman D, Blynn M, Lewis EA. *Biophys J*. 2010; 99:561. [PubMed: 20643075]
27. Lieblein AL, Buck J, Schlepckow K, Fürtig B, Schwalbe H. *Angew Chem, Int Ed*. 2012; 51:250.
28. Lieblein AL, Krämer M, Dreuw A, Fürtig B, Schwalbe H. *Angew Chem, Int Ed*. 2012; 51:4067.
29. Wright EP, Huppert JL, Waller ZAE. *Nucleic Acids Res*. 2017; 45:2951. [PubMed: 28180276]
30. Li X, Peng Y, Ren J, Qu X. *Proc Natl Acad Sci U S A*. 2006; 103:19658. [PubMed: 17167055]
31. Day HA, Huguin C, Waller ZA. *Chem Commun*. 2013; 49:7696.
32. Kang HJ, Kendrick S, Hecht SM, Hurley LH. *J Am Chem Soc*. 2014; 136:4172. [PubMed: 24559432]
33. Kendrick S, Kang HJ, Alam MP, Madathil MM, Agrawal P, Gokhale V, Yang D, Hecht SM, Hurley LH. *J Am Chem Soc*. 2014; 136:4161. [PubMed: 24559410]
34. Kendrick S, Hurley LH. *Pure Appl Chem*. 2010; 82:1609. [PubMed: 21796223]
35. Chen Y, Agrawal P, Brown RV, Hatzakis E, Hurley L, Yang D. *J Am Chem Soc*. 2012; 134:13220. [PubMed: 22866911]
36. Gowan SM, Harrison JR, Patterson L, Valenti M, Read MA, Neidle S, Kelland LR. *Mol Pharmacol*. 2002; 61:1154. [PubMed: 11961134]
37. Shin-ya K, Wierzbka K, Matsuo K, Ohtani T, Yamada Y, Furihata K, Hayakawa Y, Seto H. *J Am Chem Soc*. 2001; 123:1262. [PubMed: 11456694]
38. Han H, Langley DR, Rangan A, Hurley LH. *J Am Chem Soc*. 2001; 123:8902. [PubMed: 11552797]
39. Yang D, Okamoto K. *Future Med Chem*. 2010; 2:619. [PubMed: 20563318]
40. Balasubramanian S, Hurley LH, Neidle S. *Nat Rev Drug Discovery*. 2011; 10:261. [PubMed: 21455236]
41. Brooks TA, Kendrick S, Hurley L. *FEBS J*. 2010; 277:3459. [PubMed: 20670278]
42. Brown RV, Hurley LH. *Biochem Soc Trans*. 2011; 39:635. [PubMed: 21428953]
43. Neidle S. *J Med Chem*. 2016; 59:5987. [PubMed: 26840940]
44. Sun D, Hurley LH. *Methods Mol Biol*. 2010; 608:65. [PubMed: 20012416]
45. Kypr J, Kejnovska I, Renciuk D, Vorlickova M. *Nucleic Acids Res*. 2009; 37:1713. [PubMed: 19190094]
46. Vorlickova M, Kejnovska I, Sagi J, Renciuk D, Bednarova K, Motlova J, Kypr J. *Methods*. 2012; 57:64. [PubMed: 22450044]
47. Gray RD, Petraccone L, Buscaglia R, Chaires JB. *Methods Mol Biol*. 2010; 608:121. [PubMed: 20012419]
48. Gray RD, Petraccone L, Trent JO, Chaires JB. *Biochemistry*. 2010; 49:179. [PubMed: 19961180]
49. Lee HT, Waters L, Olsen CM, Khutsishvili I, Marky LA. *Acta Chim Slov*. 2012; 59:443. [PubMed: 24061296]
50. Cashman DJ, Buscaglia R, Freyer MW, Dettler J, Hurley LH, Lewis EA. *J Mol Model*. 2008; 14:93. [PubMed: 18087730]
51. Khan N, Aviñó A, Tauler R, González C, Eritja R, Gargallo R. *Biochimie*. 2007; 89:1562. [PubMed: 17850948]
52. Manzini G, Yathindra N, Xodo LE. *Nucleic Acids Res*. 1994; 22:4634. [PubMed: 7984411]
53. Pataskar SS, Dash D, Brahmachari SK. *J Biomol Struct Dyn*. 2001; 19:307. [PubMed: 11697735]
54. Cantor CR, Warshaw MM, Shapiro H. *Biopolymers*. 1970; 9:1059. [PubMed: 5449435]
55. Beppu K, Jaboine J, Merchant MS, Mackall CL, Thiele CJ. *J Natl Cancer Inst*. 2004; 96:46. [PubMed: 14709738]
56. Kaneko M, Yang W, Matsumoto Y, Watt F, Funa K. *Exp Cell Res*. 2006; 312:2028. [PubMed: 16624290]

57. Nilsson MB, Zage PE, Zeng L, Xu L, Cascone T, Wu HK, Saigal B, Zweidler-McKay PA, Heymach JV. *Oncogene*. 2010; 29:2938. [PubMed: 20208561]
58. Wetterskog D, Moshiri A, Ozaki T, Uramoto H, Nakagawara A, Funa K. *Mol Cancer Res*. 2009; 7:2031. [PubMed: 19952113]
59. Yang W, Wetterskog D, Matsumoto Y, Funa K. *Int J Cancer*. 2008; 123:2020. [PubMed: 18697203]
60. Eccles SA, Box C, Court W. *Biotechnol Annu Rev*. 2005; 11:391. [PubMed: 16216785]
61. Hulkower KI, Herber RL. *Pharmaceutics*. 2011; 3:107. [PubMed: 24310428]
62. Camp SM, Bittman R, Chiang ET, Moreno-Vinasco L, Mirzapioazova T, Sammani S, Lu X, Sun C, Harbeck M, Roe M, Natarajan V, Garcia JG, Dudek SM. *J Pharmacol Exp Ther*. 2009; 331:54. [PubMed: 19592667]
63. Mirzapioazova T, Moitra J, Moreno-Vinasco L, Sammani S, Turner JR, Chiang ET, Evenoski C, Wang T, Singleton PA, Huang Y, Lussier YA, Watterson DM, Dudek SM, Garcia JG. *Am J Respir Cell Mol Biol*. 2011; 44:40. [PubMed: 20139351]
64. Moreno-Vinasco L, Quijada H, Sammani S, Siegler J, Letsiou E, Deaton R, Saadat L, Zaidi RS, Messana J, Gann PH, Machado RF, Ma W, Camp SM, Wang T, Garcia JG. *Am J Respir Cell Mol Biol*. 2014; 51:223. [PubMed: 24588101]
65. Singleton PA, Mirzapioazova T, Guo Y, Sammani S, Mambetsariev N, Lennon FE, Moreno-Vinasco L, Garcia JG. *Am J Physiol Lung Cell Mol Physiol*. 2010; 299:L639. [PubMed: 20709728]
66. Wang L, Sammani S, Moreno-Vinasco L, Letsiou E, Wang T, Camp SM, Bittman R, Garcia JG, Dudek SM. *Crit Care Med*. 2014; 42:e189. [PubMed: 24335440]
67. Guo K, Gokhale V, Hurley LH, Sun D. *Nucleic Acids Res*. 2008; 36:4598. [PubMed: 18614607]
68. Sun D, Liu WJ, Guo K, Rusche JJ, Ebbinghaus S, Gokhale V, Hurley LH. *Mol Cancer Ther*. 2008; 7:880. [PubMed: 18413801]
69. De Armond R, Wood S, Sun D, Hurley LH, Ebbinghaus SW. *Biochemistry*. 2005; 44:16341. [PubMed: 16331995]
70. Guo K, Pourpak A, Beetz-Rogers K, Gokhale V, Sun D, Hurley LH. *J Am Chem Soc*. 2007; 129:10220. [PubMed: 17672459]
71. Shin YJ, Kumarasamy V, Camacho D, Sun D. *Oncogene*. 2015; 34:1292. [PubMed: 24662821]
72. Agrawal P, Lin C, Mathad RI, Carver M, Yang D. *J Am Chem Soc*. 2014; 136:1750. [PubMed: 24450880]
73. Dai J, Chen D, Jones RA, Hurley LH, Yang D. *Nucleic Acids Res*. 2006; 34:5133. [PubMed: 16998187]
74. Brown RV, Gaerig VC, Simmons T, Brooks TA. *Molecules*. 2013; 18:15019. [PubMed: 24317528]
75. Bryan TM, Jarstfer MB. *Methods*. 2007; 43:332. [PubMed: 17967703]
76. Sissi C, Gatto B, Palumbo M. *Biochimie*. 2011; 93:1219. [PubMed: 21549174]
77. Chen Y, Yang D. *Curr Protoc Nucleic Acid Chem*. 2012; 50:17.5.1.
78. Dai J, Carver M, Hurley LH, Yang D. *J Am Chem Soc*. 2011; 133:17673. [PubMed: 21967482]
79. Abouantoun TJ, MacDonald TJ. *Mol Cancer Ther*. 2009; 8:1137. [PubMed: 19417143]
80. Kim KJ, Jung JM, Cho JY, Woo SY, Cho KA, Ryu KH, Yoo ES. *Exp Ther Med*. 2011; 2:557. [PubMed: 22977540]
81. Timeus F, Crescenzo N, Doria A, Foglia L, Pagliano S, Ricotti E, Fagioli F, Tovo PA, Cordero di Montezemolo L. *Oncol Rep*. 2012; 27:734. [PubMed: 22159894]
82. Fleming AM, Zhou J, Wallace SS, Burrows CJ. *ACS Cent Sci*. 2015; 1:226. [PubMed: 26405692]

Figure 1.

PDGFR- β core promoter NHE forms four predominantly parallel G-quadruplexes from overlapping sequences that control about 40% of the transcriptional firing. (A) Cartoon representation of the PDGFR- β proximal promoter region depicting the NHE, transcription start site (TSS), and coding region of the PDGFR- β gene. The brackets at the top indicate the regions included in the WT-NHE and NO-NHE pGL-3 luciferase constructs. The sequence of the NHE and downstream CCAAT box (blue) is shown in the center; the seven runs of poly-GC are highlighted in red and numbered R1–R7. Brackets at the bottom indicate overlapping sequences used to form the equilibrating set of constitutive G-quadruplexes: 3'-end; 3'-mid; 5'-mid; 5'-end. (B) Decrease in promoter activity of the NO-NHE relative to the WT-NHE at 24 h by luciferase assay; data presented as luciferase (L) normalized to cotransfected renilla (R) relative to the WT-NHE (60% decrease, ***p < 0.001). (C) CD spectra and thermal stability for the full-length and constitutive G-quadruplex-forming sequences in 100 mM [K⁺]. The inset table shows the sequence/color identity of the spectral traces as well as their associated thermal stabilities (T_m) in 100 mM or 50 mM [K⁺] buffer. (D) DMS chemical footprints of the (a) 3'-end, (b) 3'-mid, (c) 5'-mid, (d) 5'-end, and (e) full-length G-quadruplex-forming sequences in the absence (lanes 1 and 2) and presence (lanes 3 and 4) of DMS; lane 3 serves as a sequencing lane. Open,

closed, and partially closed circles indicate protection, cleavage, and partial protection, respectively, from DMS of guanines in the presence of 100 mM $[K^+]$ buffer (lane 4).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 2.

Stable formation of the 3'-end G-quadruplex carrying a 3'-GGA sequence. (A) Importance of the R1 3'-adenine for WT 3'-end G-quadruplex formation and stability as evaluated by CD spectra using the WT and mutant sequences shown below the CD spectra, with color identity and thermal stability (T_m) in 50 mM [K⁺] buffer indicated. The Pu22 sequence's point-mutated bases are highlighted in blue, while runs of guanines are highlighted in red. (B) An R1 A-to-G mutation in the PDGFR- β NHE unexpectedly produced a significant increase in promoter activity relative to the WT-NHE by luciferase assay 24 h post-transfection, while the R6 G-to-T mutation produces the expected decrease in promoter activity (** $p < 0.01$). (C) Determination of the role of the adenine in the 3'-GGA sequence using the fluorescent adenine analog 2AP in place of the R1 3'-adenine as determined by [K⁺]-dependent quenching of the 2AP fluorescence in the 3'-end Pu22 2AP (left) and FL Pu41 2AP (right). Data presented as relative fluorescence (RFU) relative to 0 mM [K⁺] for each respective sequence (** $p < 0.01$). (D) DMS chemical footprints comparing effects of base composition of G-quadruplexes formed from the WT 3'-end Pu22 (left) and R1 A-to-G mutant (right) 3'-end Pu22 G-quadruplex-forming sequences in the absence (lanes 1 and 2) and presence (lanes 3 and 4) of DMS; lane 3 serves as a sequencing lane. Open, closed, and partially closed circles indicate protection, cleavage, and partial protection, respectively, from DMS of guanines in the presence of 100 mM [K⁺] buffer (lane 4).

Figure 3.

Mutations in critical runs of cytosine differentially effect i-motif formation from the WT PDGFR- β NHE Py41. (A) The PDGFR- β Py41 WT, R1 mutant, R5 mutant, and R6 mutant sequences were subjected to a pH gradient for CD spectra and thermal-melt analysis. (a) The R1 mutant (blue line), R5 mutant (orange line), and R6 mutant (green line) have altered spectra relative to the WT sequence (red line) at pH 6.0. (b) The transitional pH was determined from the plot of pH versus the corresponding molar ellipticity at 286 nm and is indicated on the inset table. (c) PDGFR- β NHE Pu41/Py41 sequence with color-matched sites of mutated bases for the R1 mutant (blue), R5 mutant (orange), and R6 mutant (green) with runs of poly-GC highlighted in red. (B) The effect of each mutation on thermal stability (T_m , °C), relative to the WT sequence (T_m , °C), was determined for each oligo at their respective transition pH and at pH 6.0 by CD spectra and thermal-melt analysis.

Figure 4.

Mutational analysis of the PDGFR- β NHE WT and R1 mutant Py41 sequences identifies cytosines used for i-motif formation. (A) Graphical representation of CD thermal-melt analysis (T_m , °C) at pH 6.5 for sequential C-to-T point mutations of the WT (left) and R1 mutant (right) PDGFR- β NHE Py41 sequences. Data is represented as T_m of each C-to-T point mutation normalized to its respective parent sequence. (B) Folding pattern for the predicted major i-motif in the R1 mutant (5:10:8 loop isomer). Putative hnRNP K consensus CT-elements (CCCT), located in the lateral loops, are indicated by red lines. Colored circles represent nucleobases: cytosine (yellow), adenine (green), guanine (red), and thymine (blue).

Figure 5.

(A) Stoichiometric binding analysis monitoring fraction change in molar ellipticity at 262 nm as a function of increasing molar equivalents, demonstrating that NSC338258 has binding selectivity for the 3'-end G-quadruplex (blue line), the 5'-end (magenta line), the 5'-mid (green line), and the 3'-mid (red line). Structure of NSC338258 is shown on the right-hand side of the panel. (B) The CD scans of the four different constituent G-quadruplexes. The CD scan the FL 41-mer with NSC338258 is shown in Supplemental Figure 1A.

Figure 6.

Structure-activity relationships of novel ellipticine analogs shows that GSA1129 preferentially stabilizes the 3'-end G-quadruplex and specifically downregulates PDGFR- β promoter activity in vitro. (A) Comparative binding of ellipticines to each of the four constituent PDGFR- β G-quadruplexes to identify ellipticine derivatives that stabilize the 3'-end G-quadruplex using CD thermal analysis (T_m) at 2 mol equiv of each compound. (B) Target-based inhibition of promoter activity evaluated in vitro by luciferase assay at 24 h using G-quadruplex-containing WT-NHE and G-quadruplex-deficient NO-NHE; $\pm$ compound at either IC_{50} concentrations (blue bars) or 50% IC_{50} concentrations (magenta bars) of (a) NSC338258, (b) GSA1116, (c) GSA1137, (d) methoxyellipticine, (e) GSA1112, (f) GSA1125, and (g) GSA1129. Data are presented as luciferase (L) normalized to cotransfected renilla (R) relative to no-drug controls (green bars). GSA1129 decreases luciferase activity in a G-quadruplex-dependent manner by approximately 60% ($***p < 0.001$; $**p < 0.01$; $*p < 0.05$).

Figure 7. GSA1129 selectively binds to and stabilizes the 3'-end G-quadruplex in the PDGFR- β promoter NHE. (A) The Taq polymerase stop assay was used to show the selectivity of GSA1129 for the different G-quadruplexes formed in the PDGFR- β promoter by increasing concentrations of the drug (0, 0.05, 0.5, 1, and 2.5 μ M) in the absence of K⁺. The two left most lanes are sequencing reactions on the template. The primer extension products are designated 3'-end, 3'-mid, 5'-mid, and 5'-end. The corresponding arrest sites are indicated on the NHE Pu41 sequence below the figure. The Taq polymerase stop assay shows that GSA1129 selectively arrests the stop site for the 3'-end G-quadruplex. (B) DMS footprinting of Pu41 in the presence (lane 4) or absence (lane 3) of GSA1129 in the absence of KCl. Lanes 1 and 2 serve as control experiments in the absence of DMS. Autoradiogram densitometric scans comparing G-quadruplex formation in the absence and presence of GSA1129 are shown to the right of the gel. GSA1129 appears to selectively induce formation of the 3'-end G-quadruplex in the absence of K⁺. (C) GSA1129 binds to and increases the thermal stability of the WT 41-mer (red) and R1 A-to-G mutant 41-mer (blue) relative to their no drug controls (WT, green; R1, black). GSA1129 preferentially stabilizes the WT 41-mer as demonstrated by the 3-fold increase in T_m relative to the R1 A-to-G mutant 41-mer.

Figure 8.

While GSA1129 shows selectivity in the PDGFR- β promoter for the 3'-end G-quadruplex, this is not exhibited at a more general level in promoters containing G-quadruplexes. (A) GSA1129 increases the thermal stability of the MYC G-quadruplex by approximately 30 °C (blue), relative to MYC G-quadruplex without GSA1129 (red), as determined by CD thermal-melt analysis. (B) qRT-PCR demonstrates that IC₅₀ concentrations of GSA1129 differentially affect the transcript levels of other oncogenes with promoter-based G-quadruplexes. The largest effect is observed for PDGFR- β and MYC in SK-N-SH at 24 h relative to DMSO controls.

Figure 9.

Identification and validation of PDGFR- β i-motif-interactive compounds. (A) Molecular-beacon i-motif FRET probes were designed using the PDGFR- β Py41 WT and R1 mutant sequences for high-throughput screening as diagrammed in part a, with FAM and BHQ1 at the 5' - and 3' -ends, respectively. The WT Py41mer is shown in part a. (b) Representative data from FRET (pH 6.6) high-throughput screening of the NCI Diversity and Mechanistic compound libraries for the WT (black bars) and R1 mutant (green bars) FRET probes (left panel). An asterisk indicates “hits,” such as compound NSC309874, found on plate 4727, which increased the FRET of the R1 mutant more than 2-fold (c). (B) CD spectra and thermal-melt analysis demonstrate that NSC309874 stabilizes the WT and R1 mutant (a, b) but does not interact with or stabilize the PDGFR- β G-quadruplex-forming sequence (c). Arrows indicate effect of NSC309874 on CD spectral signal.

Figure 10.

Molecularly targeting the 3'-end G-quadruplex with GSA1129 attenuates PDGFR- β transcription and pathway-related phenotypes in vitro. (A) Comparative sensitivity to GSA1129 (red bars) compared to methoxyellipticine (yellow bars) as evaluated by MTS cytotoxicity assay at 24 h in PDGFR- β -expressing cancer cell lines. (B) Time- and [GSA1129]-dependent decrease of PDGFR- β promoter activity and mRNA as evaluated by qRT-PCR of nascent transcript (blue bars) and mature transcript (red bars) in SK-N-SH (a, b) and SH-SY-5Y (c, d). Alternate mechanisms of mRNA regulation in SH-SY-5Y are highlighted by differences in nascent transcript and mature transcript (mRNA) at 24 and 48 h 25% IC₅₀ and 50% IC₅₀. All PDGFR- β expression at 24 and 48 h is normalized to WT SK-N-SH or WT SH-SY-5Y untreated, time-matched controls (data not included on graphs). (C) Biological consequences of attenuating PDGFR- β transcript by (c) anti-PDGFR- β siRNA, (d) [GSA1129] IC₅₀, (e) 50% [GSA1129] IC₅₀, and (f) 25% [GSA1129] IC₅₀ on viability (dead cells; red line) and proliferation (live cells; blue line) over a 0–96-h time-course in SK-N-SH; controls include (a) not treated and (b) scrambled siRNA. Data are presented as live or dead cell activity relative to $t = 0$ h for each individual treatment condition. (D) Migration study by scratch assay in SK-N-SH using subcytotoxic, 25% IC₅₀ [GSA1129], as monitored over a 0–48-h time-course compared to no drug and DMSO controls.

Figure 11.

In vitro analysis localizes the transcriptional inhibitory activity of NSC309874 to the PDGFR- β NHE. NSC309874 elicits inhibitory effects on transcription in an NHE-dependent fashion. (A) Luciferase assay in HEK 293 with 100 μ M NSC309874 ($IC_{50} > 100 \mu$ M) shows (a) the NO-NHE construct is not affected, while the NHE-containing (b) WT and (c) R1 mutant constructs exhibit decreased luciferase activity. Luciferase data is presented as the 24-h post-transfection ratio of firefly to renilla luciferase activity normalized to a DMSO control for each construct. *P* values (** $p < 0.001$, *** $p < 0.0001$) were determined by ANOVA with post hoc Tukey, relative to DMSO control. (B) qRT-PCR demonstrates that NSC309874 reduces PDGFR- β gene expression by approximately 25% and 50% in SK-N-SH at 24 h using 0.5 IC_{50} (50 μ M) and IC_{50} (100 μ M) concentrations of compound, respectively, relative to DMSO control. *P* values (* $p < 0.05$, ** $p < 0.001$) were determined by ANOVA with post hoc Tukey, relative to untreated control.

Figure 12.

GSA1129 attenuates LPS-induced lung injury. (A) BAL protein levels, (B) BAL total white blood cell counts, and (C) BAL PMN percentage were increased by LPS challenge (1 mg/kg) and attenuated by GSA1129 treatment (10 mg/kg). A lower dose of GSA1129 (1 mg/kg) does not exhibit significant effects. (D) Lung histology analysis indicates that GSA1129 attenuates LPS-mediated leukocyte infiltration into the alveoli. $n = 3-5$. $*p < 0.05$ between two compared groups.

Scheme 1.
Synthesis of GSA1129

Scheme 2.
Synthesis of Ellipticine Analogs

Scheme 3.
Synthesis of GSA1125

Table 1

Structures and Activity Data on Ellipticine Analogs

Compound	R ₁	R ₂	Cytotoxicity HEK293 IC ₅₀ μM
NSC338258	–H		3.16
Methoxyellipticine	–H	–OCH ₃	6.70
GSA1116		–OCH ₃	1.62
GSA1112			2.20
GSA1121		–OCH ₃	13.92
GSA1129		–OCH ₃	1.87
GSA1137		–H	9.46
GSA1134		–H	22.47

Compound	R ₁	R ₂	Cytotoxicity HEK293 IC ₅₀ μM
GSA1125	-H		12.6