A putative G-quadruplex structure in the proximal promoter of vegfr-2 has implications for drug design to inhibit tumor angiogenesis

Yaping Liu¹,²,³, Wenxian Lan¹,², Chunxi Wang¹,² and Chunyang Cao¹,²,³*

¹State Key Laboratory of Bioorganic and Natural Product Chemistry, Center for Excellence in Molecular Synthesis, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, 345 Lingling Road, Shanghai, 200032, China.
²University of Chinese Academy of Sciences, No.19(A) Yuquan Road, Shijingshan District, Beijing, 100049, China
³Collaborative Innovation Center of Chemistry for Life Sciences, Chinese Academy of Sciences, 345 Lingling Road, Shanghai, 200032, China.

Running title: Structure of G-quadruplex formed in VEGFR-2 promoter

*Corresponding author:
Chunyang Cao, Tel. 86-21-54925491; fax: 86-21-64166128; email: ccao@mail.sioc.ac.cn

These authors contributed equally to the paper as first authors

Keywords: angiogenesis, VEGFR-2, G-quadruplex, structure, nuclear magnetic resonance

Abstract
Tumor angiogenesis is mainly regulated by vascular endothelial growth factor (VEGF), produced by cancer cells. It is active on the endothelium via VEGF receptor 2 (VEGFR-2). G-quadruplexes are DNA secondary structures formed by guanine-rich sequences, for example, within gene promoters where they may contribute to transcriptional activity. The proximal promoter of vegfr-2 contains a G-quadruplex, which has been suggested to interact with small molecules that inhibit VEGFR-2 expression and thereby tumor angiogenesis. However, its structure is not known. Here, we determined its NMR solution structure, which is composed of three stacked G-tetrads containing three syn guanines. The first guanine (G1) is positioned within the central G-tetrad. We also observed that a noncanonical, V-shaped loop spans three G-tetrad planes, including no bridging nucleotides. A long and diagonal loop, which includes six nucleotides, connects reversal double chains. With a melting temperature of 54.51°C, the scaffold of this quadruplex is stabilized by one G-tetrad plane stacking with one nonstandard base pair, G₃-C₈, whose bases interact with each other through only one hydrogen bond. In summary, the NMR solution structure of the G-quadruplex in the proximal promoter region of the VEGFR-2 gene reported here has uncovered its key features as a potential anticancer drug target.

Introduction
The development of new blood vessels from a pre-existing vasculature, known as angiogenesis, plays an important role in a lot of physiological and pathological processes including tumor growth¹. It is tightly regulated...
by a balance between pro- and antiangiogenic factors. During tumor growth, this balance favors angiogenic switch 2. The tumor angiogenesis is a key process for the cancer progression, as new blood vessels nourish the growth of tumor cells and facilitate their metastasis 3-4. The proangiogenic factors are then produced and secreted by tumor, including basic fibroblast growth factor 5, vascular endothelial growth factor (VEGF)6-8, epidermal growth factor9, platelet-derived growth factor 10 and the related receptors, including VEGF receptor 1 (VEGFR-1), VEGFR-2 and VEGFR-3 11-12, and the endothelial cells (EC) of nearby blood vessels are activated. These proangiogenic molecules are thus the targets for anti-angiogenic drugs. Among them, VEGF-A is considered as the most important factor implicated in tumor angiogenesis 13 through binding to tyrosine kinase receptors expressed on the surface of EC with high-affinities, such as VEGFR-1, VEGFR-2 and the receptor functionally more relevant in the transduction of proangiogenic stimuli incoming from tumor cells 14-16. The interactions between VEGF-A and VEGFR-2 lead to the receptor auto-phosphorylation, and the phosphorylated tyrosine residues activate signaling cascades, eventually resulting in cellular processes involved in angiogenesis 17, such as vesicle trafficking, cytoskeleton regulation, microtubule dynamics, cell polarity and membrane transport 18. So far, a few of drugs had been developed in preclinical models to target VEGF/VEGFR pathway for anticancer therapy, including engineered proteins that mimic VEGF receptors and small molecules inhibitors that preferentially target VEGFR-2, and antibodies against VEGF-A or its receptors.

The G-quadruplex is generally formed in DNA and RNA sequences containing repeated short guanine-rich tracts by the stacked interaction of successive G•G•G•G tetraads (G-tetrad) and stabilized by the bound monovalent Na+ or K+ or NH4+ cation. Its arrangement can be tetramolecular, bimolecular or intramolecular, which are possible by virtue of the changes in strand polarities and also the sequence and topologies of the loops 19-20. The G-quadruplex structures were found in vitro in human telomere21-24 and the promoter regions of different oncogenes, such as c-Myc 25, c-kit26, 27, VEGF28 and Bcl-2 29-30. These G-quadruplexes in gene promoters have physiological and structural characteristics, making them become drug targets. Their structural diversity further provides a high possibility to selectively design anticancer drug. At the same time, with the application of stabilizing ligands as a new class of anticancer agents, G-quadruplex structures can be visualized in human cells.

Previously, a guanine rich sequence was found within the proximal promoter region of vegfr-2, and was suggested to form an antiparallel G-quadruplex structure 31. This G-quadruplex structure can be efficiently stabilized by small molecule quarfloxin, which has been progressed in clinical trials. Here, we report the NMR solution structure of a unique G-quadruplex formed in the proximal promoter region of vegfr-2, which provides not only the molecular details of this G-quadruplex, but also important insights for its loop conformations and interactions with the core tetrad structures.

Results and discussion

Sequence isolation of VEGFR-17T G-
Previously, four consecutive guanine repeats (Figure 1A) in the G-rich region of vegfr-2 promoter was suggested to form an intramolecular G-quadruplex in the presence of K⁺ solution. However, its imino proton NMR spectrum in Figure 1B demonstrates very broad peaks, suggesting the presence of multiple G-quadruplex forms in the wild-type sequence of the proximal promoter of vegfr-2. Its truncated forms, VEGFR-25mer and VEGFR-24mer, still have broad signals, although VEGFR-25mer has much narrower signals than VEGFR-24mer does. To obtain the sequence of vegfr-2 proximal promoter suitable for NMR G-quadruplex structure studies, we mutated the nucleotide residue flanking the forth G-tract and got the sequence VEGFR-24T, which demonstrates more divergent signals than VEGFR-25mer and VEGFR-24mer do. To investigate which guanines are essential to G-quadruplex formation, further single-site mutations from guanine to adenine (A) and to thymine (T) were carried out (Supplementary Figure S1 and Table S1), indicating that only base G₁₇ is not involved in G-quadruplex formation. Mutants VEGFR-17T and VEGFR-17A have much better signal diversities in their imino proton NMR spectra than VEGFR-24T (Figure 1). Obviously, VEGFR-17T looks more suitable for NMR studies than VEGFR-17A, due to its fewer overlapping signals than those of VEGFR-17A. Moreover, its imino proton NMR spectra displays almost no changes upon the experimental temperature being increased from 10 °C to 20 °C (Supplementary Figure S2), and the concentrations of G-quadruplex and K⁺ ion being varied (Supplementary Figures S3 and S4). Therefore, the VEGFR-17T sequence was chosen for NMR structure determination.

**Folding topology of VEGFR-17T G-quadruplex**

Generally, the CD spectrum of a parallel-type G-quadruplex is always absent in Cotton effects at around λ = 290 nm, but with a strong positive band at 263 nm and a negative band at 240 nm. To probe the folding of VEGFR-17T G-quadruplex containing 5 mM, 10 mM, 20 mM and 100 mM K⁺ solution, respectively, CD spectroscopies were performed, which exhibits positive absorptions at λ = 290 nm and 263 nm as well as a smaller negative band at about 240 nm (supplementary Figure S5), identical to the previously reported results, typical of the hybrid structure. Therefore, VEGFR-17T has ideally mixed parallel/antiparallel folding in different K⁺ solution.

Then, NMR experiments were further carried out to obtain the detailed information of the folding of VEGFR-17T G-quadruplex. The signals in imino proton region of one-dimensional ¹H-NMR spectrum of VEGFR-17T display narrow spectral line widths (2 – 5 Hz for the sharpest peaks at 20 °C), indicative of a monomeric intramolecular structure. This is supported by the concentration-independent and the well-resolved quality of ¹H-NMR spectra for VEGFR-17T (supplementary Figure S3), where the line widths of imino proton signals are almost identical at the concentrations of 0.1 mM, 0.5 mM and 1 mM of VEGFR-17T NMR samples. The guanine imino and H8 protons of VEGFR-17T were unambiguously assigned using the site-specific low enrichment ¹⁵N-labeled and natural abundance VEGFR-17T samples through bond correlation strategies (Figure 2). The non-
exchangeable base and sugar protons in VEGFR-17T sequence were assigned using standard protocols by 2D $^1$H-$^1$H NOESY, COSY, TOCSY and $^{31}$P-$^1$H HETCOR (supplementary Figure S6) experiments acquired in H$_2$O or D$_2$O solutions, respectively. The expanded $^1$H-$^1$H NOESY cross-peaks (mixing time 250 ms) including the classical H8/H6-H1′ sequential connectivities were shown in Figure 3A. The intensities of intra-residue H8-H1′ NOE cross-peaks (Figure 3B) indicate syn glycosidic conformations of base G$_1$, G$_9$ and G$_{20}$, in contrast to other residues, which adopt anti conformation.

The characteristic NOEs between imino and H8 protons indicates that the formation of an intramolecular G-quadruplex involves three G-tetrads, G$_1$•G$_{22}$•G$_{10}$•G$_{15}$, G$_2$•G$_{21}$•G$_9$•G$_{13}$ and G$_{20}$•G$_{23}$•G$_{11}$•G$_{16}$ (Figures 4A and 4B), which is further confirmed by the NOEs in the imino protons (i.e., H1) region (Figure 4C). In this G-quadruplex fold, residues G$_1$, G$_{22}$, G$_{10}$ and G$_{15}$ are in the central G-tetrad, consistent with the imino protons of them being the most protected from exchange with water (Figure 3C). The glycosidic conformations of guanines around these three G-tetrads are syn (G$_{20}$)-anti (G$_{23}$)-anti (G$_{11}$)-anti (G$_{16}$), syn (G$_1$)-anti (G$_{22}$)-anti (G$_{10}$)-anti (G$_{15}$), anti (G$_2$)-anti (G$_{21}$)-syn (G$_9$)-anti (G$_{13}$), respectively, as shown in Figure 5.

Solution structure of VEGFR-17T G-quadruplex

The three-dimensional structure of the VEGFR-17T quadruplex was calculated on the basis of NMR restraints (Table 1) using program XPLOR$^{24}$. Finally, twenty refined structures of the VEGFR-17T G-quadruplex with the lowest energy are shown in Figure 5A.
of the G-quadruplex conformation. When the amino group of G3 is removed by replacing G3 by inosine (i.e., I3), the signals belonging to the imino protons of bases G2, G13, G9 and G21 were disappeared or shifted to overlap with other signals (supplementary Figure S8), suggesting that this G-tetrad is unstable. When base G3 is substituted by adenine (i.e., A3), the whole G-quadruplex folding is collapsed, indicating that the carbonyl group of G3 is involved in hydrogen-bond with amino group of base C8, and stabilizes the whole folding of VEGFR-17T G-quadruplex.

The first guanine is positioned in the central G-tetrad of VEGFR-17T

In contrast to the fact that the first guanine is always part of a terminal G-tetrad in almost all unimolecular G-quadruplexes19,35, in VEGFR-17T, base G1 is positioned in the central G-tetrad, thus, the G20 and G1-G2 that are parts of the same strand, that are not connected to each other, relative to the unbroken linkage for the other three G9-G10-G11, G13-G15-G16 and G21-G22-G23 strands of VEGFR-17T G-quadruplex. Currently, there are only two reported structures of G-quadruplexes in which base G1 locates in the central G-tetrad36-37 (supplementary Figure S9), one is that formed in human chl1 intronic region, the other is that formed in d(G3T4G4). All these three G-quadruplexes have a V-shaped loop (i.e., zero-nt) spanning three G-tetrad planes, due to the position of G1 base, indicating a possible relationship in G-quadruplex folding between V-shaped loop and G1 position at the central G-tetrad.

A V-shaped loop in VEGFR-17T spans three G-tetrad planes

The VEGFR-17T G-quadruplex contains one V-shaped loop, which contains zero-nt, and spans three G-tetrad planes. This kind of V-shaped loop was also ever observed within the G-quadruplex scaffold linked by two adjacent G-tetrads38, the dimeric G-quadruplexes formed by d(G17T4G4)37, the dimeric G-quadruplex adopted by d(GLGLT4GLGL)2 where L is a locked nucleic acid39, the G-quadruplex in chl1 intronic sequence40, and the G-quadruplex in hCEB1 minisatellite G-rich sequence41, as shown in Figure 6. Among these G-quadruplexes, VEGFR-17T is unique with a 6-nt loop, in which base G20 belongs to the top G-tetrad of one strand (this strand is broken by G1), while residue G21 is in the bottom G-tetrad of an adjacent strand. Thus, it seems that G1 is partially intercalated between bases G20 and G21 within the folding topology of the VEGFR-17T G-quadruplex, similar to that observed in the chl1 G-quadruplex. In contrast, the chl1 G-quadruplex contains four loops, each of them includes less than three bases. The G-quadruplex formed by d(G17T4G4)2 contains one V-shaped loop (i.e., G8-G9) only in one strand, while the G-quadruplex formed by d(GLGLT4GLGL)2 includes four specific V-shaped loops, i.e., G1-L2 in one strand and G9-L10 in another strand, all spanning three G-tetrad planes. The bases G1, L10, L2 and G9 alternately form two diagonal columns G1-L10-L2-G9 of the G-quadruplex.

A 6-nt long and diagonal loop destabilizes VEGFR-17T G-quadruplex

Generally, due to the conformational flexibility, the longer the loop contained in G-quadruplex, the more unstable the G-quadruplex. For examples, the G-quadruplex form observed in hCEB1 mini-satellite G-rich sequence40.
contains a 4-nt side loop (C5T6G7A8), which causes the instability of this G-quadruplex with a Tm value of only 46.18 °C. The G-quadruplex adopted by ODN sequence, containing a 5-nt diagonal loop, exhibits low stability with a Tm value of 58 °C. This accounts for why there’re very few reported structures of G-quadruplex with a loop longer than 6-nt in RCSB protein data bank (www.rcsb.org). However, when the loop is anchored by some interactions between loop and G-tetrads, the G-quadruplex will become stable. As shown in Figure 7, although the G-quadruplex formed in human CEB25 mini-satellite locus has a 9-nt side loop (T16G17T18A14G15T16G17T18)42, its structure is still stable with a Tm value of 76.5 °C in a 90 mM K+ solution, because the conformation of this long loop is fixed by π-π stacking interactions between G-tetrads and Watson-Crick base pair A2•T18 and potential A4•G17 base pair. Another example is the G-quadruplex formed in human BCL2 promoter region (i.e., Bcl2mid)43 with a 7-nt long loop (A10G11G12A13A14T15T16). The structure of this G-quadruplex is stable with a Tm value of 65 °C, due to π-π stacking interactions between base pair A10•T15 in this loop and G-tetrad G1•G17•G17•G21. The current VEGFR-17T G-quadruplex contains a 6-nt loop (G7T4A6C6C7C8), connecting two diagonal columns (G20-G1-G2 and G9-G10-G11). Although base-pair G3•C8 in this loop has π-π stacking interactions with G-tetrad G2•G13•G9•G21, but this base pair has only one hydrogen-bond interaction between them, this base-pair thus belongs to non-classical Watson-Crick base-pair. Therefore, VEGFR-17T G-quadruplex is more stable than the G-quadruplex formed in hCEB1 minisatellite G-rich sequence, but more unstable than those formed in human BCL2 promoter region and CEB25 mini-satellite locus.

**VEGFR-17T G-quadruplex is a potential drug target to inhibit tumor angiogenesis**

In a word, in the presence of K+, the proximal promoter of vegfr-2 can form G-quadruplex. Our NMR data indicates that the G-quadruplex of VEGFR-17T has a melting temperature of 54.51°C. It has a unique folding with three striking features. Its first guanine is positioned in the central G-tetrad plane, it has a V-shaped loop spanning three G-tetrad planes, as well as a 6-nt loop connecting two anti-parallel columns. Single-site mutations on G1, G2 and G11 (in the sequence of VEGFR-24T) into adenine, a new, parallel G-quadruplex (confirmed by their CD spectra, shown in supplementary Figure S5) with two G-tetrads is generated with a melting temperature of nearly 80°C (supplementary Figure S7), indicating that they are much more stable than VEGFR-17T G-quadruplex. The imino proton NMR spectra of these mutants display only eight signals (supplementary Figure S10), revealing that they are identical to one another, all with only two G-tetrad planes. This stable G-quadruplex is one of multiple forms in wild-type G-quadruplex formed in the proximal promoter of vegfr-2. Therefore, the current VEGFR-17T G-quadruplex can be a potential anti-cancer drug target, and will be efficiently stabilized by small molecules to inhibit vegfr-2 expression, resulting in a turning off of signaling components, which further mediates the cellular events leading to endothelial cell proliferation, migration and differentiation. This assumption had been confirmed by the previously reported data31, in which angiogenic process is strongly inhibited in vitro and in vivo by the ligands targeting G-quadruplex formed...
in the proximal promoter of vegfr-2, resulting in impairment of the endothelial cell function. In conclusion, we determine the solution structure of a unique G-quadruplex form adopted by the proximal promoter of vegfr-2, which can be potentially stabilized by G-quadruplex ligands, thereby revealing a promising way to block VEGFR-2 expression as target for anticancer therapy.

**Experimental procedures**

**G-quadruplex sequence isolation and sample preparation**

Isotope unlabeled single-strand DNA, such as VEGFR, VEGFR-25mer, VEGFR-24mer, VEGFR-24T, VEGFR-17A and VEGFR-17T (Figure 1) and their variants (Table S1), and site-specific low-enrichment (6% $^{15}$N-labeled guanine) VEGFR-17T strands were commercially synthesized at an Ultra-PAGE grade from Shanghai Boshang Biotechnology Co, China. They were dialyzed successively three times against water and 100 mM K$^+$ solution (20 mM KH$_2$PO$_4$, 80 mM KCl, pH 6.8), then annealed at 98°C for five minutes, and cooled down to room temperature. The concentrations of the DNA samples for NMR experiments were typically 0.1 - 1 mM in NMR buffer (20 mM KH$_2$PO$_4$, 80 mM KCl, pH 6.8).

**Circular dichroism (CD) spectroscopy**

Circular dichroism spectrum was recorded at 25°C on JASCO-815 spectropolarimeter using a 1-cm path length quartz cuvette with a reaction volume of 350 μL. DNA concentration was 20 - 30 μM. The DNA oligonucleotides were prepared in a pH 6.8 buffer containing 20 mM KH$_2$PO$_4$ and 80 mM KCl, and were heated to 98°C for 5 min and cooled down to room temperature overnight. An average of three scans was taken, the spectrum of the buffer was subtracted.

**Differential scanning calorimetry (DSC)**

To probe the stability of VEGFR-17T, VEGFR-1A, VEGFR-2A and VEGFR-11A, differential scanning calorimetry (DSC) measurements were carried out using an auto VP-CAP-DSC Microcalorimeter (Malvern Inc, USA). The experiments were performed at the single-stranded DNA concentrations in the range 50 - 70 μM. Scans were performed at 1.5°C/min in the 20 - 110°C temperature range. A buffer-buffer scan was subtracted from the buffer-sample scans and linear-polynomial baselines were drawn for each scan. Baseline corrected thermograms were normalized with respect to the single strand molar concentration to obtain the corresponding molar heat capacity curves. The enthalpy for the overall unfolding of G-quadruplex structures were estimated by integrating the area under the heat capacity versus temperature curves. Tm values were obtained as the temperatures corresponding to the maximum of each thermogram peak. Entropy values were calculated by integrating the curve $C_p/T$ versus T (where $C_p$ is the molar heat capacity and T is the temperature in Kelvin) and the free energy values were computed by the equation $\Delta G = \Delta H - T\Delta S$.

**NMR data collection and spectra analysis**

NMR experiments were performed on 600 MHz and 800 MHz Varian and Bruker spectrometers at 10 and 20°C, respectively. Resonances of DNA were assigned unambiguously using site-specific low enrichment $^{15}$N-labelled samples and through-
bond correlations at natural abundance. The NMR experiments for samples in water solution were performed with watergate or jump-and-return water suppression techniques. The acquisition data points were set to 2048 x (250-512) (complex points). All spectra were processed with the program NMRPipe. The 45° or 60° shifted sine-squared functions were applied to two dimensional (2D) $^1$H-$^1$H NOESY and TOCSY spectra. The fifth-order polynomial functions were employed for the baseline corrections. The final spectral sizes are 2048x 1024. The $^1$H chemical shifts were referenced to 2, 2-dimethylsilapentane-5-sulfonic acid (DSS). Peak assignments and integrations were achieved using Sparky (http://www.cgl.ucsf.edu/home/sparky/). The NOE peaks were integrated using the peak fitting function and volume integration of Sparky.

The $^{31}$P NMR spectrum was collected on a DNA sample at 1.5 mM in D$_2$O at 20°C and was referenced to an external standard of 85% H$_3$PO$_4$, including the one dimensional proton-decoupled phosphorus spectrum, and two dimensional heteronuclear $^{31}$P-$^1$H Correlation Spectroscopy ($^{31}$P-$^1$H HETCOR). Assignments of the individual $^{31}$P resonance were accomplished by a combination of 2D $^1$H-$^1$H NOESY, COSY, TOCSY and heteronuclear $^{31}$P-$^1$H HETCOR spectra.

**Solution structure determination**

The distances between non-exchangeable protons were estimated based on the NOE cross-peak volumes at 50, 200 and 250 ms mixing times, and were divided into strong (1.8 – 2.9Å), medium (1.8 - 3.5Å) and weak (1.8 – 6.0Å) groups, respectively, using the proton H5-H6 distance (2.45Å) in cytosine base as a reference. Exchangeable proton restraints are based on NOESY data sets at two mixing times (50 and 250 ms) in H$_2$O. Cross-peaks involving exchangeable protons were classified as strong (strong intensity at 50 ms), medium (weak intensity at 50 ms) and weak (observed only at 250 ms) NOEs. The G-tetrads within the G-quadruplex were restrained with distances corresponding to ideal hydrogen bond geometry. Each individual hydrogen bond was restrained using two distance restraints (heavy atom–heavy atom and heavy atom–proton, respectively). Hydrogen bond distance restraints were also applied to the carbonyl oxygen (in residue G3) and the -NH$_2$ group (in residue C8), based on the single-site mutation studies on the residue G3 into adenine (i.e., A3) and inosine (i.e., I3). The aromatic rings of guanines in each G-tetrad were also restrained into one plane during calculation. 24 dihedral angle restraints were used to restrict the glycosidic torsion angle (χ) for the experimentally assigned syn configuration, i.e. G1, G9 and G20 tetrad-guanines [60° ± 35°], as well as for the experimentally assigned anti-configuration bases, i.e. G2, G3, T4, A5, C6, C7, C8, G10, G11, T12, G13, A14, G15, G16, T17, G18, C19, G21, G22, G23 and T24 [240° ± 70°].

The structural calculations of VEGFR-17T were carried out using a standard simulated annealing protocol implemented in the program XPLOR-2.37 (NIH version). A total of 618 NOE-distance restraints (Table 1), of which 216 are from inter-residue NOE interactions, were incorporated into the NOE-restrained structure calculation. A total of ten iterations (50 structures in the initial ten iterations) were performed. 100 structures were computed in the last five iterations, 20 conformers with the lowest energy and
minimal restrain violations are used to represent the three dimensional structures. In the ensemble of the simulated annealing 20 structures, there was no distance constraint violation more than 0.3 Å and no torsion angle violation more than 5°.

Acknowledgements

This work was supported by National Key R&D Program of China (2017YFE0108200 and 2016YFA0502302), by the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB 20000000), by National Science Foundation of China (NSFC) under No. 91753119, 21472229 and 21778065. The authors also thank facility team members in National Center of Protein Sciences Shanghai (NCPSS) and the NMR center of Shanghai Institute of Materia Medica, Chinese Academy of Sciences for their great help with NMR spectra acquirement.

Authors’ contribution

C. C conceived the experimental designs, wrote the manuscript and assisted with data analysis and structural calculation. Y. L made NMR samples, acquired and analyzed data, calculated structure, draw figures. W.L helped to acquire most NMR spectra and analyze them. C.W performed and analyzed some enzymatic experiments.

Conflict of interest

The authors declare that they have no conflicts of interest with the contents of this article.

References

1. Folkman, J., Angiogenesis in cancer, vascular, rheumatoid and other disease. Nat Med 1995, 1 (1), 27-31.
2. Bergers, G.; Benjamin, L. E., Tumorigenesis and the angiogenic switch. Nat Rev Cancer 2003, 3 (6), 401-10.
3. Liotta, L. A.; Kleinerman, J.; Saidel, G. M., Quantitative relationships of intravascular tumor cells, tumor vessels, and pulmonary metastases following tumor implantation. Cancer Res 1974, 34 (5), 997-1004.
4. Folkman, J., What is the evidence that tumors are angiogenesis dependent? J Natl Cancer Inst 1990, 82 (1), 4-6.
5. Rifkin, D. B.; Moscatelli, D., Recent developments in the cell biology of basic fibroblast growth factor. J Cell Biol 1989, 109 (1), 1-6.
6. Martiny-Baron, G.; Marme, D., VEGF-mediated tumour angiogenesis: a new target for cancer therapy. Curr Opin Biotechnol 1995, 6 (6), 675-80.
7. Folkman, J., Role of angiogenesis in tumor growth and metastasis. Semin Oncol 2002, 29 (6 Suppl 16), 15-8.
8. Jain, R. K., Tumor angiogenesis and accessibility: role of vascular endothelial growth factor. Semin Oncol 2002, 29 (6 Suppl 16), 3-9.
9. Gleave, M. E.; Hsieh, J. T.; Wu, H. C.; Hong, S. J.; Zhai, H. E.; Guthrie, P. D.; Chung, L. W., Epidermal growth factor receptor-mediated autocrine and paracrine stimulation of human transitional cell carcinoma. Cancer Res 1993, 53 (21), 5300-7.
10. Nicosia, R. F.; Nicosia, S. V.; Smith, M., Vascular endothelial growth factor, platelet-derived growth factor, and insulin-like growth factor-1 promote rat aortic angiogenesis in vitro. Am J Pathol 1994, 145 (5),
11. Davis, S.; Aldrich, T. H.; Jones, P. F.; Acheson, A.; Compton, D. L.; Jain, V.; Ryan, T. E.; Bruno, J.; Radziejewski, C.; Maisonpierre, P. C.; Yancopoulos, G. D., Isolation of angiopoietin-1, a ligand for the TIE2 receptor, by secretion-trap expression cloning. *Cell* 1996, 87 (7), 1161-9.
12. Maisonpierre, P. C.; Suri, C.; Jones, P. F.; Bartunkova, S.; Wiegand, S. J.; Radziejewski, C.; Compton, D.; McClain, J.; Aldrich, T. H.; Papadopoulos, N.; Daly, T. J.; Davis, S.; Sato, T. N.; Yancopoulos, G. D., Angiopoietin-2, a natural antagonist for Tie2 that disrupts in vivo angiogenesis. *Cancer Cell* 2002, 1 (2), 193-202.
13. Inoue, M.; Hager, J. H.; Ferrara, N.; Gerber, H. P.; Hanahan, D., VEGF-A has a critical, nonredundant role in angiogenic switching and pancreatic beta cell carcinogenesis. *Cancer Cell* 2002, 1 (2), 193-202.
14. Mustonen, T.; Alitalo, K., Endothelial receptor tyrosine kinases involved in angiogenesis. *J Cell Biol* 1995, 129 (4), 895-8.
15. Ferrara, N.; Gerber, H. P.; LeCouter, J., The biology of VEGF and its receptors. *Nat Med* 2003, 9 (6), 669-76.
16. Shibuya, M., Vascular endothelial growth factor receptor-1 (VEGFR-1/Flt-1): a dual regulator for angiogenesis. *Angiogenesis* 2006, 9 (4), 225-30; discussion 231.
17. Olsson, A. K.; Dimberg, A.; Kreuger, J.; Claesson-Welsh, L., VEGF receptor signalling - in control of vascular function. *Nat Rev Mol Cell Biol* 2006, 7 (5), 359-71.
18. Heasman, S. J.; Ridley, A. J., Mammalian Rho GTPases: new insights into their functions from in vivo studies. *Nat Rev Mol Cell Biol* 2008, 9 (9), 690-701.
19. Patel, D. J.; Phan, A. T.; Kuryavyi, V., Human telomere, oncogenic promoter and 5'-UTR G-quadruplexes: diverse higher order DNA and RNA targets for cancer therapeutics. *Nucleic Acids Res* 2007, 35 (22), 7429-55.
20. Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S., Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res* 2006, 34 (19), 5402-15.
21. Wang, Y.; Patel, D. J., Solution structure of a parallel-stranded G-quadruplex DNA. *J Mol Biol* 1993, 234 (4), 1171-83.
22. Zhang, N.; Phan, A. T.; Patel, D. J., (3 + 1) Assembly of three human telomeric repeats into an asymmetric dimeric G-quadruplex. *J Am Chem Soc* 2005, 127 (49), 17277-85.
23. Luu, K. N.; Phan, A. T.; Kuryavyi, V.; Lacroix, L.; Patel, D. J., Structure of the human telomere in K+ solution: an intramolecular (3 + 1) G-quadruplex scaffold. *J Am Chem Soc* 2006, 128 (30), 9963-70.
24. Phan, A. T.; Luu, K. N.; Patel, D. J., Different loop arrangements of intramolecular human telomeric (3+1) G-quadruplexes in K+ solution. *Nucleic Acids Res* 2006, 34 (19), 5715-9.
25. Siddiqui-Jain, A.; Grand, C. L.; Bearss, D. J.; Hurley, L. H., Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc Natl Acad Sci U S A* 2002, 99 (18), 11593-8.
26. Shirude, P. S.; Okumus, B.; Ying, L.; Ha, T.; Balasubramanian, S., Single-molecule conformational analysis of G-quadruplex formation in the promoter DNA duplex of the proto-oncogene c-kit. *J Am Chem Soc* 2007, 129 (24), 7484-5.
27. Kuryavyi, V.; Phan, A. T.; Patel, D. J., Solution structures of all parallel-stranded monomeric and dimeric G-quadruplex scaffolds of the human c-kit2 promoter. *Nucleic Acids Res* 2010, 38 (19), 6757-73.
28. Sun, D.; Guo, K.; Rusche, J. J.; Hurley, L. H., Facilitation of a structural transition in the
polyurine/polypurine tract within the proximal promoter region of the human VEGF gene by the presence of potassium and G-quadruplex-interactive agents. *Nucleic Acids Res* 2005, 33 (18), 6070-80.

29. Nambiar, M.; Goldsmith, G.; Moorthy, B. T.; Lieber, M. R.; Joshi, M. V.; Choudhary, B.; Hosur, R. V.; Raghavan, S. C., Formation of a G-quadruplex at the BCL2 major breakpoint region of the t(14;18) translocation in follicular lymphoma. *Nucleic Acids Res* 2011, 39 (3), 936-48.

30. Dai, J.; Dexheimer, T. S.; Chen, D.; Carver, M.; Ambrus, A.; Jones, R. A.; Yang, D., An intramolecular G-quadruplex structure with mixed parallel/antiparallel G-strands formed in the human BCL-2 promoter region in solution. *J Am Chem Soc* 2006, 128 (4), 1096-8.

31. Salvati, E.; Zizza, P.; Rizzo, A.; Iachettini, S.; Cingolani, C.; D’Angelo, C.; Porru, M.; Randazzo, A.; Pagano, B.; Novellino, E.; Pisanu, M. E.; Stoppacciaro, A.; Spinella, F.; Bagnato, A.; Gilson, E.; Leonetti, C.; Biroccio, A., Evidence for G-quadruplex in the promoter of vegfr-2 and its targeting to inhibit tumor angiogenesis. *Nucleic Acids Res* 2014, 42 (5), 2945-57.

32. Phan, A. T.; Patel, D. J., A site-specific low-enrichment (15)N,(13)C isotope-labeling approach to unambiguous NMR spectral assignments in nucleic acids. *J Am Chem Soc* 2002, 124 (7), 1160-1.

33. Phan, A. T., Long-range imino proton-13C J-couplings and the through-bond correlation of imino and non-exchangeable protons in unlabeled DNA. *J Biomol NMR* 2000, 16 (2), 175-8.

34. Brunger, A. T., X-PLOR Version 3.1: A System for X-ray Crystallography and NMR. Book X-PLOR Version 3.1: A System for X-ray Crystallography and NMR. *Yale University Press, New Haven, CT. 1993*.

35. Neidle, S., The structures of quadruplex nucleic acids and their drug complexes. *Curr Opin Struct Biol* 2009, 19 (3), 239-50.

36. Kuryavyi, V.; Patel, D. J., Solution structure of a unique G quadruplex scaffold adopted by a guanosine-rich human intronic sequence. *Structure* 2010, 18 (1), 73-82.

37. Crnuigelj, M.; Sket, P.; Plavec, J., Small change in a G-rich sequence, a dramatic change in topology: new dimeric G-quadruplex folding motif with unique loop orientations. *J Am Chem Soc* 2003, 125 (26), 7866-71.

38. Zhang, N.; Gorin, A.; Majumdar, A.; Kettani, A.; Chernichenko, N.; Skripkin, E.; Patel, D. J., V-shaped scaffold: a new architectural motif identified in an A x (G x G x G x G) pentad-containing dimeric DNA quadruplex involving stacked G(anti) x G(syn) tetrads. *J Mol Biol* 2001, 311 (5), 1063-79.

39. Nielsen, J. T.; Arar, K.; Petersen, M., Solution structure of a locked nucleic acid modified quadruplex: introducing the V4 folding topology. *Angew Chem Int Ed Engl* 2009, 48 (17), 3099-103.

40. Piazza, A.; Cui, X.; Adrian, M.; Samazan, F.; Heddi, B.; Phan, A. T.; Nicolas, A. G., Non-Canonical G-quadruplexes cause the hCEB1 minisatellite instability in Saccharomyces cerevisiae. *Elife* 2017, 6.

41. Marusic, M.; Sket, P.; Bauer, L.; Viglasky, V.; Plavec, J., Solution-state structure of an intramolecular G-quadruplex with propeller, diagonal and edgewise loops. *Nucleic Acids Res* 2012, 40 (14), 6946-56.

42. Amrane, S.; Adrian, M.; Heddi, B.; Serero, A.; Nicolas, A.; Mergny, J. L.; Phan, A. T., Formation of pearl-necklace monomorphic G-quadruplexes in the human CEB25 minisatellite. *J Am Chem Soc* 2012, 134 (13), 5807-16.

43. Dai, J.; Chen, D.; Jones, R. A.; Hurley, L. H.; Yang, D., NMR solution structure of the major G-quadruplex structure formed in the human BCL2 promoter region. *Nucleic Acids Res* 2006, 34 (18), 5133-44.

44. Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A., NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 1995, 6 (3), 277-93.

45. Kuszewski, J.; Clore, G. M., Sources of and solutions to problems in the refinement of protein NMR
structures against torsion angle potentials of mean force. *J Magn Reson* **2000**, **146** (2), 249-54.

**Figures and tables legends**

**Table 1** Structural statistics for the VEGFR-17T G-quadruplex structure

NMR distance and dihedral constraints

Distance restraints

|               |       |
|---------------|-------|
| Total NOEs    | 618   |
| Intraresidue  | 402   |
| Interresidue  | 216   |
| Sequential(|i-j|=1) | 82    |
| Non-sequential(|i-j|>1) | 134   |
| Hydrogen bonds| 50    |
| Total dihedral angle restraints | 24    |

Structural statistics

Violations (mean and SD)

|                  |       |
|------------------|-------|
| Distance constraints(Å) | 0.051±0.0049 |
| Dihedral angle constraints(°) | 0.90±0.015    |

Deviations from idealized geometry

|               |       |
|---------------|-------|
| Bond lengths (Å) | 0.0031±0.000077 |
| Bond angle(°)   | 1.03±0.013   |
| Impropers(°)    | 0.35±0.0052  |

Average pairwise r.m.s.d. of all heavy atoms (Å)

|                |       |
|----------------|-------|
| All residues   | 0.39±0.07 |
| All heavy atoms| 0.45±0.08 |
**Figure 1** G-quadruplex sequence isolation of vegfr-2 proximal promoter for NMR studies. (A) The top sequence is the wild-type (WT) VEGFR containing the reported four consecutive G-stretches; The four G-stretches are colored in red and numbered using Roman numerals; Two GC-boxes are labeled. VEGFR-25mer and VEGFR-24mer are its truncated forms; VEGFR-24T, VEGFR-17A and VEGFR-17T are the variants of VEGFR-24mer. (B) The imino proton regions of 1D $^1$H NMR spectra of wild-type VEGFR, VEGFR-25mer, VEGFR-24mer, VEGFR-24T, VEGFR-17A and VEGFR-17T at 0.3 mM sample concentrations. NMR buffer conditions: 20 °C, 20 mM K-phosphate, 80 mM KCl, pH 6.8.
**Figure 2** Resonance assignments of the protons in the bases. (A) Imino proton spectra with assignments indicated over the reference spectrum on the top. Guanine imino protons were assigned in $^{15}$N-filtered spectra of samples, 6.0% $^{15}$N-labeled guanines at the indicated positions. (B) A schematic indicating long-range J couplings used to correlate imino proton and H8 proton within the guanosine base. (C) H8 proton assignments of VEGFR-17T sequence by through-bond correlations between guanosine imino and H8 protons via $^{13}$C (at 5-position) at natural abundance, using long-range J couplings shown in (B).
Figure 3 Non-exchangeable proton assignments of VEGFR-17T Sequence. (A) Expanded $^1$H-$^1$H NOESY spectrum (200ms mixing time) correlating base and sugar H1’ protons. The line connectivities trace NOEs between a base proton (H8 or H6) and its own and 5’-flanking sugar H1’ protons. Intra-residue base to sugar H1’ NOEs are labeled with residue numbers. (B) Stacked plot of short mixing time NOESY spectrum. The strong intraresidue guanosine H8-H1’ cross-peaks (syn glycosidic bonds) are labeled and can be distinguished from weak cross-peaks (anti glycosidic bonds). (C) Imino proton NMR spectrum of VEGFR-17T sequence after 60 min in D$_2$O solution. Assignments of slowly exchanging imino protons are listed over the spectrum.
Figure 4 Determination of G-quadruplex topology for the proximal promoter sequence VEGFR-17T in K⁺ solution. (A) Characteristic guanosine imino-H8 NOE connectivity patterns around a Gα•Gβ•Gγ•Gδ tetrad as indicated by arrows. (B) Guanosine imino-H8 connectivities observed for the G-tetrad G20•G23•G11•G16, G1•G22•G10•G15 and G2•G21•G9•G13 planes. (C) Interstrand NOEs between imino H1 and aromatic H8 of unimolecular guanine bases within the same layer of G-tetrads mentioned in (B).
Figure 5 (A) Stereoview of the 20 superimposed refined structures of VEGFR-17T G-quadruplex (generated by program MOLMOL), (B) Stereoview of a representative structure in ribbon mode (generated by program SYBYL-X). Guanosine bases in G-tetrad core are colored cyan (anti) and magenta (syn). Bases in loops are in yellow for G3-C8 base pair, violet for T4A5C6C7, green for T24, blue for T17G18C19; For clarity, all bases are only shown the backbone in (B). The orientation of G-quadruplex in (A) is similar to those shown in (B). (C) Schematic structure of VEGFR-17T G-quadruplex. The anti and syn guanines are colored cyan and magenta, and G3-C8 base pair is colored in yellow, respectively. (D) The non-classical G3-C8 (in yellow) base pair has π-π stacked interactions with the G-tetrad G2•G21•G9•G13. The hydrogen-bond between G3 and C8 base is displayed as a dashed line.
Figure 6 Schematic structures of V-shaped loop contained in (A) the G-quadruplex formed by VEGFR-17T in K\+ solution (this work), (B) the dimeric G-quadruplexes formed by d(G\_T\_G\_G)_2, (C) the dimeric G-quadruplex adopted by d(GLGLT\_GLGL)_2, where L is a locked nucleic acid, (D) the G-quadruplex in chl1 intronic sequence, and (E, F) the G-quadruplexes in hCEB1 minisatellite G-rich sequence. Loops are colored in black, except that the V-shaped loops are colored in red. The anti and syn guanines are colored in cyan and in magenta, respectively.
Figure 7 Schematic structures of long loop contained in (A) the G-quadruplex formed by VEGFR-17T in K+ solution (this work), (B) the G-quadruplex adopted by ODN sequence, containing a 5-nt diagonal loop, (C) the G-quadruplex formed in human CEB25 mini-satellite locus, with a 9-nt side loop, (D) the G-quadruplex formed in human BCL2 promoter region (i.e., Bcl2mid) with a 7-nt long loop. Loops are colored black, anti and syn guanines are colored cyan and magenta, respectively. The V-shaped loop in VEGFR-17T is colored in red.
A putative G-quadruplex structure in the proximal promoter of vegfr-2 has implications for drug design to inhibit tumor angiogenesis
Yaping Liu, Wenxian Lan, Chunxi Wang and Chunyang Cao

*J. Biol. Chem.* *published online April 17, 2018*

Access the most updated version of this article at doi: 10.1074/jbc.RA118.002666

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts