J|A|C|S ARTICLES

Published on Web 11/24/2010

An RNA G-Quadruplex Is Essential for Cap-Independent Translation Initiation in Human VEGF IRES

Mark J. Morris,[†] Yoichi Negishi,[‡] Cathy Pazsint,[†] Joseph D. Schonhoft,[§] and Soumitra Basu*,[†]

Department of Chemistry and Biochemistry, Kent State University, Kent, Ohio 44242, United States, School of Pharmacy, Tokyo University of Pharmacy and Life Sciences, Tokyo, Japan, and Program in Molecular Biophysics, Johns Hopkins University, Baltimore, Maryland 21218, United States

Received July 22, 2010; E-mail: sbasu@kent.edu

Abstract: RNA G-quadruplexes located within the 5′-UTR of mRNA are almost always known to be associated with repression of cap-dependent translation. However, in this report we present functional as well as structural evidence that sequence redundancy in a G-rich segment within the 5′-UTR of human VEGF mRNA supports a 'switchable' RNA G-quadruplex structure that is essential for IRES-mediated translation initiation. Additionally, utilization of a specific combination of G-tracts within this segment allows for the conformational switch that implies a tunable regulatory role of the quadruplex structure in translation initiation. A sequence engineered from a functionally handicapped mutant moderately rescued the activity, further indicating the importance of G-quadruplex structure for VEGF IRES-A function. This to our knowledge is the first report of a conformationally flexible RNA G-quadruplex which is essential for IRES-mediated translation initiation.

Introduction

Structures present within the 5′-untranslated region (5′-UTR) of mRNAs, including cellular mRNAs, have been shown to regulate translation.[1] Many mRNAs have been reported to initiate translation in a 5′ cap-independent manner.[2−6] Such cap-independent translation is initiated by internal ribosomal entry sites (IRESs) found within the 5′-UTR of these RNAs. Originally, IRES elements were reported in viruses,[5−9] but later these elements were discovered in a wide variety of cellular RNAs occurring in organisms ranging from fruit flies to humans. A list of reported cellular IRESs show that they can be found in mRNAs from proto-oncogenes, growth factors, transcription factors, and translation factors to cell cycle genes, among others.[10−15] Besides the IRES, other secondary structural elements are known to regulate translation of mRNA. For example, 5′-UTR of bacterial mRNAs contains RNA motifs termed 'riboswitches', which in response to metabolite or protein factor binding undergo a conformational switch to regulate translation.[16−20]

Human vascular endothelial growth factor (hVEGF) is a key angiogenic growth factor whose overexpression is critical for tumor outgrowth, metastasis, and prognosis of several cancers.[21] The 5′-UTR region of VEGF mRNA is unusually long (1038 nt) and GC rich and can initiate translation via a cap-independent mechanism.[10,11,22] This alternative translation mode can be particularly critical for VEGF expression under hypoxia, a condition often encountered during neovascularization, which is essential to tumor outgrowth and metastasis.[23,24] Interestingly,

[†] Kent State University.
[‡] Tokyo University of Pharmacy and Life Sciences.
[§] Johns-Hopkins University.

(1) Kozak, M. *J. Biol. Chem.* **1991**, *266*, 19867–19870.
(2) Hellen, C. U. T.; Sarnow, P. *Gene. Dev.* **2001**, *15*, 1593–1612.
(3) Gilbert, W. V.; Zhou, K.; Butler, T. K.; Doudna, J. A. *Science* **2007**, *317*, 1224–1227.
(4) Macejak, D. G.; Sarnow, P. *Nature* **1991**, *353*, 90–94.
(5) Jang, S. K.; Krausslich, H. G.; Nicklin, M. J.; Duke, G. M.; Palmenberg, A. C.; Wimmer, E. *J. Virol.* **1988**, *62*, 2636–2643.
(6) Pelletier, J.; Sonenberg, N. *Nature* **1988**, *334*, 320–325.
(7) Pestova, T. V.; Kolupaeva, V. G.; Lomakin, I. B.; Pilipenko, E. V.; Shatsky, I. N.; Agol, V. I.; Hellen, C. U. T. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 7029–7036.
(8) Pfingsten, J. S.; Kieft, J. S. *RNA* **2008**, *14*, 1255–1263.
(9) Kieft, J. S. *Trends. Biochem. Sci.* **2008**, *33*, 274–283.
(10) Akiri, G.; Nahari, D.; Finkelstein, Y.; Le, S. Y.; Elroy-Stein, O.; Levi, B. Z. *Oncogene* **1998**, *17*, 227–236.
(11) Huez, I.; Creancier, L.; Audigier, S.; Gensac, M. C.; Prats, A. C.; Prats, H. *Mol. Cell. Biol.* **1998**, *18*, 6178–6190.
(12) Vagner, S.; Waysbort, A.; Marenda, M.; Gensac, M.-C.; Amalric, F.; Prats, A.-C. *J. Biol. Chem.* **1995**, *270*, 20376–20383.
(13) Nanbru, C.; Lafon, I.; Audigier, S.; Gensac, M.-C.; Vagner, S.; Huez, G.; Prats, A.-C. *J. Biol. Chem.* **1997**, *272*, 32061–32066.
(14) Stoneley, M.; Paulin, F. E.; Le Quesne, J. P.; Chappell, S. A.; Willis, A. E. *Oncogene* **1998**, *16*, 423–428.
(15) Bernstein, J.; Sella, O.; Le, S.-Y.; Elroy-Stein, O. *J. Biol. Chem.* **1997**, *272*, 9356–9362.
(16) Winkler, W.; Nahvi, A.; Breaker, R. R. *Nature* **2002**, *419*, 952–956.
(17) Serganov, A. *Curr. Opin. Struct. Biol.* **2009**, *19*, 251–259.
(18) Garst, A. D.; Batey, R. T. *BBA-Gene. Regul. Mech.* **2009**, *1789*, 584–591.
(19) Holbrook, S. R. *Curr. Opin. Struct. Biol.* **2005**, *15*, 302–308.
(20) Ray, P. S.; Jia, J.; Yao, P.; Majumder, M.; Hatzoglou, M.; Fox, P. L. *Nature* **2009**, *457*, 915–919.
(21) Harmey, J. *VEGF and Cancer*; Springer: New York, 2004.
(22) Miller, D. L.; Dibbens, J. A.; Damert, A.; Risau, W.; Vadas, M. A.; Goodall, G. J. *FEBS Lett.* **1998**, *434*, 417–420.
(23) Shweiki, D.; Itin, A.; Soffer, D.; Keshet, E. *Nature* **1992**, *359*, 843–845.
(24) Banai, S.; Shweiki, D.; Pinson, A.; Chandra, M.; Lazarovici, G.; Keshet, E. *Cardiovasc. Res.* **1994**, *28*, 1176–1179.

this region harbors two separate internal ribosomal entry sites, which are capable of independently initiating translation without the help of the 5′-cap. One of the sequences is a 463 nucleotide fragment (nt 91−554, IRES-B) that initiates translation at the CUG 499,[25,26] whereas the other, a 293 nucleotide portion (nt 745−1038, IRES-A), has been shown to initiate translation at the canonical AUG codon.[11,22] We focused our studies on IRES-A as it is known to maintain VEGF translation under hypoxia.[27] Also, it is known that the protein product synthesized from initiation at the AUG codon is a secreted form of VEGF and thus important for the autocrine and paracrine functions of the VEGF ligand.[26]

Although scores of cellular IRESes have been identified thus far, the effects of RNA structures on the cellular IRES function are not well defined. Upon the basis of the accumulated evidence it has been suggested that RNA structures in this region are beneficial to the translation process through binding with ribosome subunits and initiation factors.[7,28,29] However, the molecular details on the secondary and tertiary structures of cellular IRESs and how they regulate IRES function are not well understood, and there is paucity of experimental data on hVEGF IRES structures in particular.

Recent reports have shown the presence of G-rich sequences within the 5′-UTR of mRNAs.[30−33] Such sequences adopt G-quadruplex structures in vitro and have been shown to repress translation in vivo. The G-quadruplex structures consist of G-quartet planes, which is a planar array of four guanines joined by Hoogsteen hydrogen bonds, with a monovalent cation at the center.[34] Both Na^+ and K^+ can function as this monovalent cation, although K^+ has been shown to have a more stabilizing effect.[34] Because of significantly higher intercellular concentrations of K^+ compared to Na^+, the former is more physiologically relevant in terms of monovalent cation-dependent study of G-quadruplex structures. Four repeats of a minimum of two guanosines are sufficient to form an intramolecular G-quadruplex in vitro as well as in vivo.[34−36] These G-rich domains have been found in the context of various DNA regions,[37−48] and such sequences have been extensively studied for their ability to form G-quadruplex structures. In contrast, the roles of RNA G-quadruplexes are just beginning to emerge. RNA G-quadruplex-forming motifs have been found in the HIV-1 genomic RNA,[49] FMR1,[50,51] FGF-2,[52] NRAS,[30] Zic-1,[31] MT3-MMP,[32] and IGF-II[53] mRNAs. The current model is that solitary G-quadruplex structures present in the 5′-UTR of cellular mRNAs act as translational repressors. Herein, we show that a G-quadruplex in the 5′-UTR of hVEGF IRES-A alone is essential for its translation initiation activity. A 17 nt sequence identified by footprinting can potentially adopt multiple G-quadruplex structures, which is a deviation from the RNA quadruplex sequences reported in the literature, and such multiple structures are possible because of the redundancy in the number of G-stretches. Mutational analyses of the hVEGF IRES-A present in the context of a bicistronic dual-luciferase reporter system show that a 'switchable' sequence containing greater than four G-stretches (nts 774−790) provides enough redundancy to ensure the formation of RNA G-quadruplex structures that are critical for the cap-independent translation initiation. Furthermore, a sequence designed by modifying a crippled mutant sequence was able to partially rescue IRES-A activity. Our data suggest that by choosing different combinations of four G-stretches, a conformational diversity can be generated within the previously defined segment, which can potentially modulate efficiency of cap-independent translation initiation in VEGF IRES-A, while pointing toward a novel mechanism of control of gene expression at the translation level.

Materials and Methods

Preparation of Oligonucleotide Sequences. RNA sequences were synthesized by in vitro transcription[54] and purified via denaturing polyacrylamide gel electrophoresis (PAGE). The RNA was harvested via the crush and soak method by tumbling the gel slice at 4 °C in a solution of 300 mM NaCl, 10 mM Tris-HCl (pH 7.5), and 0.1 mM EDTA. Samples were concentrated with 2-butanol, and salt was removed by ethanol precipitation and washed twice with 70% ethanol. The RNA pellet was dissolved in 10 mM Tris-HCl (pH 7.5) and 0.1 mM EDTA. Concentrations of RNAs were determined based on their absorbance value at 260 nm and extinction coefficients calculated using nearest neighbor parameters.[55]

(25) Huez, I.; Bornes, S.; Bresson, D.; Creancier, L.; Prats, H. *Mol. Endocrinol.* **2001**, *15*, 2197–2210.
(26) Bornes, S. p.; Boulard, M.; Hieblot, C.; Zanibellato, C.; Iacovoni, J. S.; Prats, H.; Touriol, C. *J. Biol. Chem.* **2004**, *279*, 18717–18726.
(27) Stein, I.; Itin, A.; Einat, P.; Skaliter, R.; Grossman, Z.; Keshet, E. *Mol. Cell. Biol.* **1998**, *18*, 3112–3119.
(28) Cobbold, L. C.; Spriggs, K. A.; Haines, S. J.; Dobbyn, H. C.; Hayes, C.; de Moor, C. H.; Lilley, K. S.; Bushell, M.; Willis, A. E. *Mol. Cell. Biol.* **2008**, *28*, 40–49.
(29) Kieft, J. S.; Zhou, K.; Jubin, R.; Doudna, J. A. *RNA* **2001**, *7*, 194–206.
(30) Kumari, S.; Bugaut, A.; Huppert, J. L.; Balasubramanian, S. *Nat. Chem. Biol.* **2007**, *3*, 218–221.
(31) Arora, A.; Dutkiewicz, M.; Scaria, V.; Hariharan, M.; Maiti, S.; Kurreck, J. *RNA* **2008**, *14*, 1290–1296.
(32) Morris, M. J.; Basu, S. *Biochemistry* **2009**, *48*, 5313–5319.
(33) Balkwill, G. D.; Derecka, K.; Garner, T. P.; Hodgman, C.; Flint, A. P. F.; Searle, M. S. *Biochemistry* **2009**, *48*, 11487–11495.
(34) Neidle, S.; Balasubramanian, S. *Quadruplex Nucleic Acids*; RSC Biomolecular Sciences: Cambridge, U.K., 2006.
(35) Paeschke, K.; Simonsson, T.; Postberg, J.; Rhodes, D.; Lipps, H. J. *Nat. Struct. Mol. Biol.* **2005**, *12*, 847–854.
(36) Wieland, M.; Hartig, J. S. *Chem. Biol.* **2007**, *14*, 757–763.
(37) Sen, D.; Gilbert, W. *Nature* **1988**, *334*, 364–66.
(38) Williamson, J. R. *Annu. Rev. Biophys. Biomol. Struct.* **1994**, *23*, 703–30.
(39) Neidle, S.; Parkinson, G. N. *Curr. Opin. Struct. Biol.* **2003**, *13*, 275–283.
(40) Gilbert, D. E.; Feigon, J. *Curr. Opin. Struct. Biol.* **1999**, *9*, 305–14.
(41) Daniels, G. A.; Lieber, M. R. *Nucleic Acids Res.* **1995**, *23*, 5006–5011.
(42) Dunnick, W.; Hertz, G. Z.; Scappino, L.; Gritzmacher, C. *Nucleic Acids Res.* **1993**, *21*, 365–372.
(43) Hanakahi, L. A.; Sun, H.; Maizels, N. *J. Biol. Chem.* **1999**, *274*, 15908–15912.
(44) Siddiqui-Jain, A.; Grand, C. L.; Bearss, D. J.; Hurley, L. H. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11593–11598.
(45) Simonsson, T.; Pecinka, P.; Kubista, M. *Nucleic Acids Res.* **1998**, *26*, 1167–1172.
(46) Dexheimer, T. S.; Sun, D.; Hurley, L. H. *J. Am. Chem. Soc.* **2006**, *128*, 5404–5415.
(47) Cogoi, S.; Xodo, L. E. *Nucleic Acids Res.* **2006**, *34*, 2536–2549.
(48) Sun, D.; Guo, K.; Rusche, J. J.; Hurley, L. H. *Nucleic Acids Res.* **2005**, *33*, 6070–6080.
(49) Sundquist, W. I.; Heaphy, S. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 3393–3397.
(50) Schaeffer, C.; Bardoni, B.; Mandel, J. L.; Ehresmann, B.; Ehresmann, C.; Moine, H. *EMBO J.* **2001**, *20*, 4803–4813.
(51) Darnell, J. C.; Jensen, K. B.; Jin, P.; Brown, V.; Warren, S. T.; Darnell, R. B. *Cell* **2001**, *107*, 489–499.
(52) Bonnal, S.; Schaeffer, C.; Creancier, L.; Clamens, S.; Moine, H.; Prats, A.-C.; Vagner, S. *J. Biol. Chem.* **2003**, *278*, 39330–39336.
(53) Christiansen, J.; Kofod, M.; Nielsen, F. C. *Nucleic Acids Res.* **1994**, *22*, 5709–5716.
(54) Milligan, J. F.; Groebe, D. R.; Witherell, G. W.; Uhlenbeck, O. C. *Nucleic Acids Res.* **1987**, *15*, 8783–8798.
(55) Puglisi, J. D.; Tinoco, I.; James, E. D.; Abelson, J. N. In *Methods in Enzymology*; Academic Press: New York, 1989; Vol. 180, pp 304−325.

Radiolabeling of RNA Oligonucleotides. Calf-intestinal alkaline phosphatase-treated RNA was 5′-end radiolabeled by treatment with T4 polynucleotide kinase (Promega), [γ-^{32}P] ATP (Pelkin Elmer), and incubated for 45 min at 37 °C. The reaction was stopped by addition of an equal volume of stop buffer (7 M urea, 10 mM Tris-HCl, pH 7.5, and 0.1 mM EDTA). The radiolabeled full-length RNA was isolated by denaturing PAGE. The RNA was extracted from the gel via the crush and soak method as described above.

Circular Dichroism (CD) Studies. All measurements were recorded at room temperature. RNA was folded by heating the samples in 150 mM KCl, 10 mM Tris-HCl (pH 7.5), and 0.1 mM EDTA at 95 °C for 5 min followed by slow cooling to room temperature over a 90 min period. The circular dichroism (CD) spectra were recorded using a Jasco J-810 spectrophotometer with a 0.1 cm cell at a scan speed of 50 nm/min with a response time of 1 s. The spectra were averaged over five scans. For each sample, a buffer baseline was collected in the same cuvette and subtracted from the average scan.

Footprinting by RNase T1. The 5′-end radiolabeled RNA was folded by heating the samples in 10 mM Tris-HCl (pH 7.5), 0.1 mM EDTA, and in the presence of 150 mM KCl or 150 mM LiCl at 70 °C for 5 min and slow cooled to 37 °C over a 30 min period. Once reactions attained the appropriate temperature, the RNA was digested with 0.01 U of RNase T1 (Ambion) for 5 min at 37 °C. The reactions were terminated by using an equal volume of stop buffer (7 M urea, 10 mM Tris-HCl, pH 7.5, and 0.1 mM EDTA). Treated RNA was electrophoresed on a denaturing gel, dried on Whatman paper, and exposed to a phosphorimager screen. The gel images were visualized by scanning the screen on a Typhoon Phosphorimager 8600 (Molecular Dynamics).

Dimethyl Sulfate (DMS) Footprinting. 5′-End-radiolabeled IRES A was prepared in 10 mM Tris-HCl pH 7.5, 150 mM KCl, 1 mM $MgCl_2$. To fold the RNA, the reaction mix was heated for 5 min at 70 °C in a heating block. The block was then allowed to cool to room temperature on the benchtop over a period of 1 h. Separate samples were prepared in an identical way except 150 mM LiCl was added. Dimethyl sulfate was then added to a final concentration of 2%, and the reaction mix was incubated at room temperature for 10 min at which point the reaction was stopped with addition of stop buffer (2 M β-mercaptoethanol, 300 mM sodium acetate, 250 μg/mL sheared salmon sperm DNA). Three volumes of 100% ethanol were added immediately, and the RNA was precipitated on a dry ice/acetone bath followed by washing the pellet with 70% ethanol. The modified RNA was reduced with sodium borohydride and then treated with aniline to induce cleavage at the corresponding phosphodiester bond, according to a previously published procedure.[56]

Plasmid Construct for IRES-A Transcription. We subcloned the relevant portions from the hVEGF cDNA (hVEGF cDNA is a kind gift from Dr. Judith Abraham, Scios Inc., CA). The portions encompassing the IRES A sequence were amplified by PCR. PCR Primer sequences: IRES-A: 5′-GGTACCGCTAGCTCGGGCCGG-GAG (sense) and 5′-TCTAGAGGTTTCGGAGGCCCGACC (antisense). The restriction sites used are *Kpn 1* and *Xba 1*, and they were cloned into a pBluescriptII(+) vector (Stratagene) containing the T7 RNA polymerase promoter.

Plasmid Construction for the Dual-Luciferase Assay. The following four primers were used for subcloning of the VEGF IRES A fragment. F1 (Forward): 5′-**AATCTACTCGAG** *TCGCGGAG-GCTTGGGGCA*, R1 (Reverse): 5′-GGCGTCTTCCAT *GGTTTCG-GAGGCCCGACC*. F2 (Forward): 5′-*GCCTCCGAAACC* ATG-GAAGACGCCAAAAAC. R2 (Reverse): 5′-CTTATCATGTCT-GCTCGAAGCG.

VEGF IRES coding sequences are shown in italics, luciferase coding sequences are shown in bold, and *Xho* I restriction site is underlined. VEGF IRES-A PCR fragment (0.3 kb) was amplified with Pfu Turbo DNA polymerase (Stratagene) using pVEGFS-1 containing VEGF IRES A as a template and primers F1 and R1. Simultaneously, a PCR fragment (1.65 kb) encompassing the Firefly luciferase was amplified from pGL3-basic vector (Promega) using primers F2 and R2. The DNA fragment was digested with *Xho* I/*Xba* I restriction enzymes and cloned into *Xho* I/*Xba* I sites of pIRES vector (Clontech). The resulting plasmid was named (pvIRESLF). The plasmid was digested with *Xho* I site and then *Nhe* I/*Xba* I fragment containing the *Renilla* luciferase cDNA from pRL-CMV vector (Promega) were introduced into the *Xho* I site of pvIRESLF by blunt-end ligation. The resulting plasmid was named (pL*R*vIRESLF).

Mutations on IRES-A Sequence. Mutations were performed using a QuikChange site-directed mutagenesis kit (Stratagene). The incorporation of mutations within the sequences was verified by sequencing performed at the Ohio State University's Plant-Microbe Genomics Facility.

Cell Culture. HeLa cells were grown in 96-well plates in Dulbecco's modified Eagles medium (DMEM) supplemented with 10% fetal bovine serum and the antibiotics streptomycin and penicillin at 37 °C in 5% CO_2 in a humidified incubator.

Luciferase and Quantitative RT-PCR Assays. HeLa cells were transfected with vectors using Lipofectamine 2000 (Invitrogen) according to the manufacturer's protocol. Twenty-four hours after transfection, *Renilla* (RL) and firefly (FL) luciferase activities were measured using a Dual-Glo Luciferase assay system (Promega) as per the manufacturer's supplied protocol on a Synergy 2 microplate reader (BioTek Instruments). Total RNA was extracted from transfected HeLa cells using a NucleoSpin RNA II kit (Clontech). Prior to reverse transcription, DNA was removed from each RNA sample upon treatment with RQ1 RNase-Free DNase (Promega). *Renilla* and firefly mRNAs were reverse transcribed using AMV-RT (New England Biolabs), and cDNA was subjected to quantitative real-time PCR using a SYBR Green PCR Master Mix kit (Applied Biosystems) on an ABI PRISM 7000 Sequence Detection System in the presence of the appropriate set of primers: forward primers RL 5′-(GTAACGCTGCCTCCAGCTAC)-3′ and FL 5′-(TTCGCTAAGAGCACCCTGAT)-3′ and reverse primers RL 5′-(GTAGGCAGCGAACTCCTCAG)-3′ and FL 5′-(GCTGCAG-CAGGATAGACTCC)-3′.

Results

Footprinting Reveals a Highly Protected G-Rich Segment within hVEGF IRES-A. To determine secondary structural features within the hVEGF IRES-A, RNase T1 footprinting was performed on the entire 293 nt sequence in the presence of the divalent and monovalent ions (Figure 1A, refer to Supporting Information Figure 1 for the entire primary sequence of IRES-A). RNase T1 catalyzes cleavage at guanosines in single-stranded RNA but not guanosines that are involved in secondary and tertiary structures, including G-quadruplexes.[32] RNase T1 footprinting on the IRES-A sequence shows that in the presence of 150 mM K^+ a G-rich region between G774 and G790 (nucleotides are numbered based on entire 5′-UTR of human VEGF mRNA) is protected from cleavage (Figure 1A), suggesting at a minimum the presence of some form of secondary structure within this segment. Because of the highly G-rich nature of the sequence and the tandem G repeats there is a possibility of the formation of a quadruplex structure within this segment. Therefore, we wanted to explore whether the protection persists in the presence of Li^+, as the monovalent cation lithium is known to lack the ability to promote or stabilize G-quadruplex structures[34] and thus was used as a control for monovalent metal-ion-dependent quadruplex folding. We did not observe any protection in the presence of 150 mM Li^+ (Figure 1A). We also did not observe any protection patterns in the presence of 150 mM Na^+ (data not shown). This observation is consistent with a previous report on the effect of

(56) Peattie, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 1760–1764.

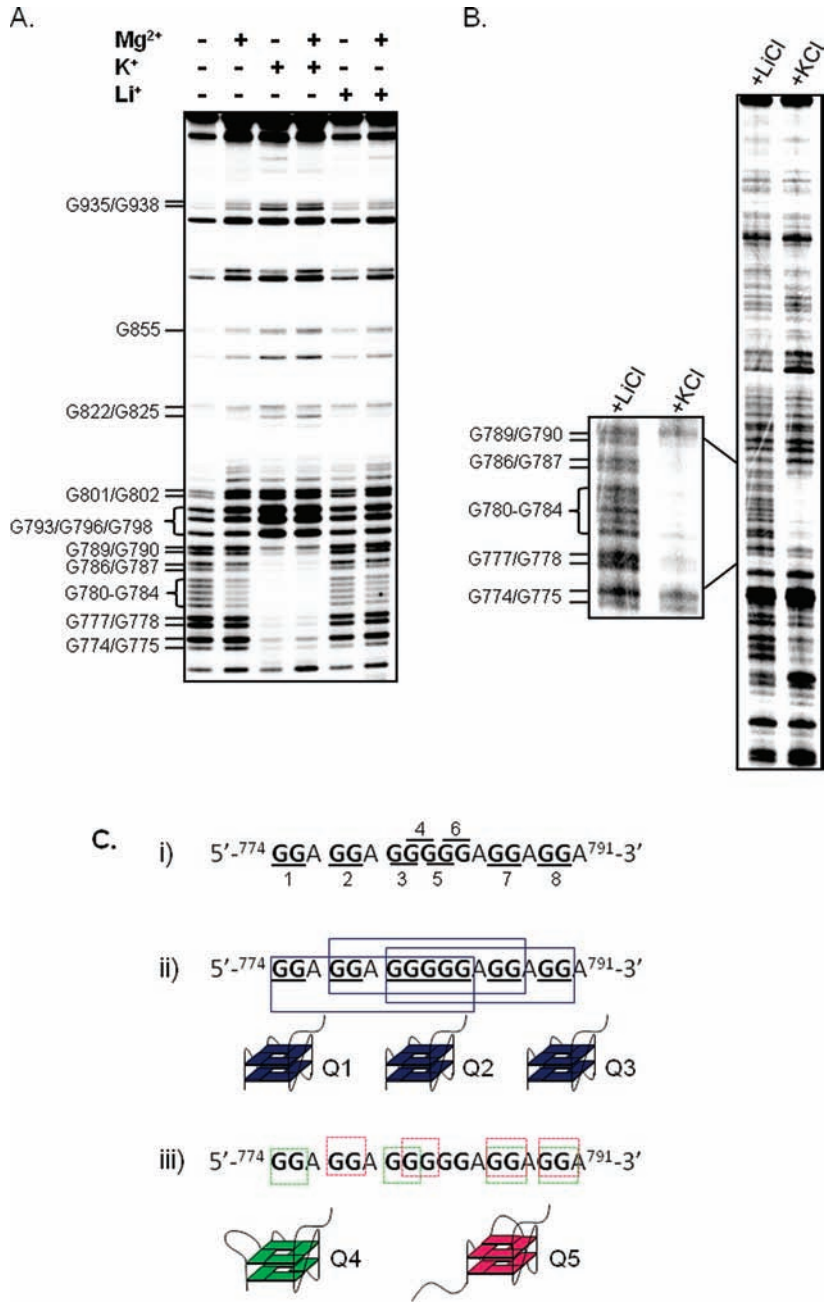

Figure 1. Scanned images of gels showing enzymatic and chemical footprinting of the 293 nt human VEGF IRES-A. (A) RNase T1 footprinting in the presence of 150 mM K^+, 150 mM Li^+, and 1 mM $MgCl_2$. (B) DMS footprinting in the presence of 150 mM K^+ or 150 mM Li^+. (C) Schematic of a subset of G-quadruplex structures that shows the different G-stretches (i) and the proposed utilization of adjacent (ii) and nonadjacent G-tracts (iii).

sodium on the RNA G-quadruplex present within the 5′-UTR of FGF-2.[52] The observation that the protection of the nucleotides only happens in the presence of K^+ ions and not in Li^+, combined with the fact that the primary sequence of this protected region harbors more than four tandem G-repeats, suggests that this segment forms an intramolecular G-quadruplex structure. Furthermore, the footprinting pattern in the presence of Mg^{2+} but lacking monovalent metal ions did not show any protection within the region, suggesting that the protected segment is a K^+-induced independently folded domain, most likely adopting a G-quadruplex structure, and overall divalent ion-dependent folding of this large RNA molecule is insufficient to cause folding of the G774−G790 segment (Figure 1A).

To further test the existence of such structures, the entire IRES-A sequence was subjected to DMS footprinting. DMS modifies the N7 of guanosine and thus can directly probe G-quadruplex structure formation.[34] Analysis of the footprinting pattern in Figure 1B shows that the regions that were protected from RNase T1 cleavage were also protected from DMS modification in the presence of K^+ ions (Figure 1B). However, the protection was lost when K^+ ions were replaced with an equimolar amount of Li^+ ions. The K^+ ion-specific DMS modification pattern strongly suggests the presence of a G-quadruplex structure within the G774−G790 section of IRES-A. The protected region consists of a five G-stretch that is flanked by two stretches of 2 Gs on both sides. Upon analysis of this primary sequence it is evident that greater than 20 two-tiered intramolecular G-quadruplexes are theoretically possible by utilizing different combinations of G-tracts. Figure 1C shows a schematic representation of a subset of possible structures;

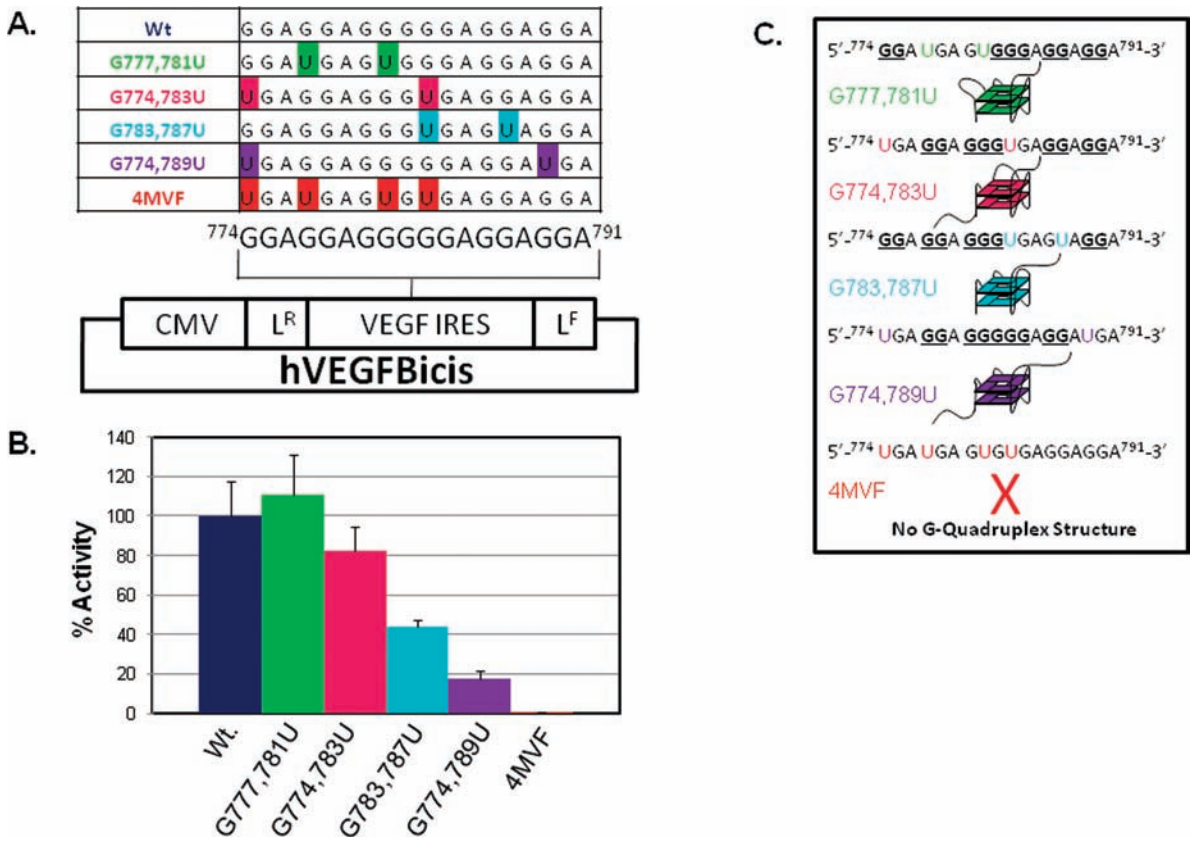

Figure 2. (A) Schematic of various dual-luciferase bicistronic constructs. (B) Histogram showing percent activity of the mutant constructs normalized to the wild-type construct. (C) Schematic of the proposed utilization of guanosines in formation of the 'switchable' G-quadruplex structure.

however, G-quadruplexes resulting from other combinations of G-stretches are possible. Thus, combining the observations made in the RNase T1 enzymatic footprinting and the DMS modification studies it can be concluded that the G774−G790 section of the VEGF IRES-A can adopt one or more quadruplex structures in a K^+-dependent manner, which is not contingent upon divalent metal-ion-mediated folding of the RNA.

In order to establish that the protected segment can also adopt a G-quadruplex conformation in isolation (G774-G790), we subjected it to circular dichroism (CD) analysis. CD experiments conducted in the presence of 150 mM K^+ showed a peak at ~260 nm and a trough around 240 nm (Supporting Information Figure 2), which are characteristics of a parallel quadruplex structure.[32,34] However, the CD spectral features were absent when analyzed in the presence of 150 mM Li^+ or Na^+ (Supporting Information Figure 2). Additionally, RNase T1 footprinting also indicated formation of a K^+-dependent G-quadruplex structure (data not shown). Thus, the G-rich segment (G774-G790) not only adopts a G-quadruplex conformation in the context of the entire 293 nt IRES but is also capable of adopting a G-quadruplex conformation in isolation.

G-Rich Protected Region Is Essential for Translation Initiation Activity of hVEGF IRES A. Thus far the data have shown that a highly G-rich portion within IRES-A RNA is strongly protected from DMS modification and RNase T1-mediated cleavage, suggesting the presence of a G-quadruplex structure, and the segment when studied in isolation adopts a parallel G-quadruplex conformation as determined by circular dichroism. An important and obvious question is whether the protected sequence is critical to translation initiation activity of IRES-A. To directly address this question, a dual-luciferase reporter construct was prepared in which the entire IRES-A sequence was placed just upstream of the firefly luciferase, while the *Renilla* luciferase was under the control of a CMV promoter (hVEGFBicis, Figure 2A). This plasmid construct allowed quantification of the translation initiation activity of the IRES-A by individually yet simultaneously measuring the dual-luciferase activities. To dissect the role of the putative G-quadruplex forming sequence, the protected segment within the IRES-A was mutated with a series of G to U substitutions (Figure 2A). Four G to U mutations were used to decisively eliminate any potential intramolecular G-quadruplex formation (G774,777, 781,783U, quadruple mutant, 4MVF). The activities of the wild-type IRES-A and 4MVF along with several other mutants (see below) are shown in Figure 2B. When the activity of the quadruple mutant was measured, the IRES-A function was found to be completely abolished (Figure 2B). Because the quadruple mutant lacks the ability to adopt a quadruplex, the observation made above suggested a potential role of a G-quadruplex structure in IRES-A-mediated translation initiation.

Since the putative G-quadruplex can form from any combination of four G-tracts, mutations were designed to investigate the importance of any specific G-quadruplex structure that may form. In a double-mutant (G777,781U), G777, and G781 were replaced with U's, making the second and third G-stretches unavailable for participation in G-quadruplex formation, which in turn can disrupt the formation of the G-quadruplex encompassed by the segment 774−785. However, the dual mutations (G777,781U) did not affect the activity of the IRES as it maintained its functionality on par with the wild-type IRES (Figure 2B; mutants G774,783U, G783,787U, and G774,789U will be discussed below). Presumably, the mutant remained fully

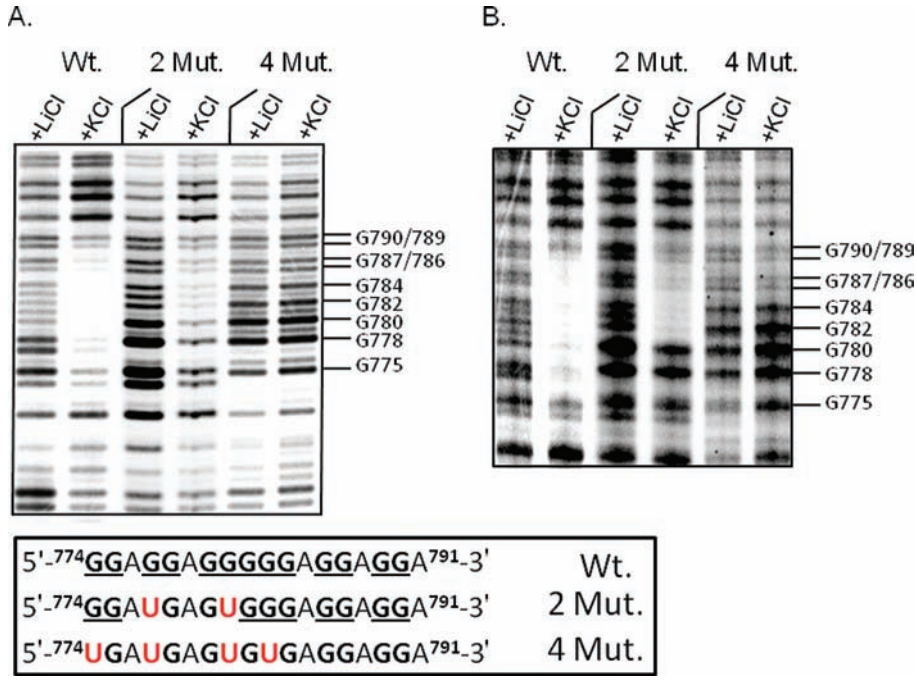

Figure 3. (A) Scanned images of gels showing RNase T1 footprinting of the 293 nt human VEGF IRES-A and its various mutants. (B) Scanned images of gels showing DMS footprinting of the 293 nt human VEGF IRES-A and its various mutants. The guanosines that can potentially participate in G-quadruplex formation are underlined. The mutated nucleotides are shown in red (bottom).

functional because four stretches of tandem G's in close enough proximity were still available, allowing formation of a G-quadruplex conformation for maintaining the IRES function (see discussion below). This combined with the observation that the quadruple mutant lacking functional activity as well the ability to form a G-quadruplex lead us to propose that the full loss of IRES activity in this mutant was most likely due to the nonformation of any G-quadruplex structure in that region. Thus, from these results it may be concluded that the quadruplex-forming sequence is necessary for IRES-A-mediated cap-independent initiation of translation. To confirm that this loss in activity was indeed occurring at the translational level, RT-PCR was performed on both the wild-type and 4MVF sequences, the results of which indicate that both constructs produced similar mRNA levels, and thus, the loss in activity of the quadruple mutant (4MVF) can be attributed to the loss of function at the translation level (Supporting Information Figure 3). The above observations can be explained by invoking the notion of the presence of a 'switchable' G-quadruplex. For example, assuming that typically contiguous G-stretches are used for G-quadruplex formation, a wild-type G-quadruplex would use any set of four G-stretches that are in tandem. However, in case of the double mutant such residues are unavailable because the second and third G-stretches from the 5′-end of the sequence are mutated. Thus, to adopt a quadruplex conformation and retain its translation initiation activity on par with the wild type, the structure most likely has to undergo a 'switch', forcing the first G-stretch to recruit distal G-stretches. While in the case of the wild-type sequence an ensemble of quadruplexes with similar energy are possible, in the case of the double mutant, many such structures cannot form; thus, a switch resulting in a subset of structures that are functionally as active as the wild type is likely to be formed.

Footprinting Indicates That the Double Mutant (G777,781U) Retains the Ability To Adopt G-Quadruplex Conformation. To ensure that the loss in activity associated with the quadruple mutant of IRES-A is truly due to the lack of formation of the quadruplex structure and not a result of disruption of the global folding or structures outside the G-rich segment, protection patterns in the presence of RNase T1 and DMS were monitored. We hypothesized that the lack of quadruplex formation in the inactive mutant (4MVF) would make all of the guanosines in the 774−790 region of IRES-A RNA (four mutations identical to the ones in 4MVF; named 4 Mut.) susceptible to RNase T1-mediated cleavage and DMS modification, whereas the double-mutant sequence (two mutations identical to the ones in G777,781U; named 2 Mut.) will show footprinting patterns similar to the wild-type sequence except at the two mutated guanosines (G777 and G781). These sequences (2 Mut. and 4 Mut.) were subjected to RNase T1-mediated cleavage and DMS modification in the presence of K^+ or Li^+ ions. In the case of the double mutant (2 Mut.), the protection pattern (774−790) corresponds to the presence of a quadruplex and is similar to that of the wild type; however, there was no discernible protection in the case of the quadruple mutant (4 Mut. Figure 3). These results directly corroborate the proposal that the double mutant (G777,781U) is active because of the formation of a G-quadruplex via utilization of four available G-stretches, whereas the quadruple mutant (4MVF) lacked the minimum number of G-stretches necessary to support intramolecular G-quadruplex formation. The quadruplex formation in the case of 2 Mut. perhaps occurs via a different set of guanosines compared to the wild-type quadruplex, which was sufficient to maintain IRES-A activity in the mutant G777,781U. Taken together, the protection patterns in the presence of K^+ and Li^+ provide strong evidence for quadruplex formation in the wild-type and double mutant (2 Mut.) of the IRES-A sequence. Another observation from the footprinting data was that the sites of cleavage/protection outside of the quadruplex-forming region in the wild-type (Wt.), double-mutant (2 Mut.), and quadruple-mutant (4 Mut.) sequences essentially remained unaltered, suggesting that the global fold of the RNAs were unperturbed

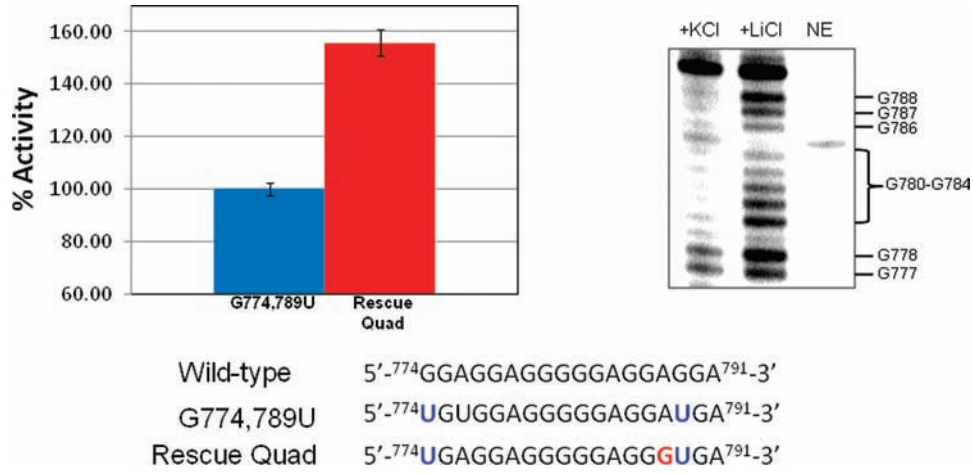

Figure 4. (A) Histogram showing percent activity of the dual-luciferase rescue mutant construct (Rescue Quad) normalized to its parent construct (G774, 789U). (B) Scanned image of a gel showing RNase T1 footprinting of the mutant Rescue Quad version of the transcribable 293 nt human VEGF IRES-A in the presence of K^+ and Li^+. NE is the no enzyme lane that contained K^+.

irrespective of the mutations (data not shown). Thus, formation or lack of formation of the G-quadruplex is apparently unrelated to the folding of the remainder of the molecule.

Utilization of Different Combinations of G-Stretches Can Lead to Differential Regulation of IRES-A Function. In order to precisely constrain the G-stretches being utilized, mutating specific Gs so as to allow determination of stretches that are preferentially used by the IRES was further investigated. For this purpose, three more double mutants (G774,783U, G783,787U, and G774,789U) of the wild-type dual-luciferase construct (hVEGFBicis) were prepared and their level of activities measured. The mutant G774,783U was ~82% active as the wild type (Figure 2B). This mutant was chosen to determine whether or not the mutations that differed between the 4MVF and the G777,781U mutants were responsible for the complete loss of activity of the IRES-A quadruple mutant. These mutations had little effect on the IRES function and were not solely responsible for the complete loss of function of the quadruple mutant but in conjunction with two other mutations were enough to fully abrogate the IRES activity. Thus, it can be argued that the lack of quadruplex formation is the primary cause for the loss of function of the quadruple (4MVF) mutant. Two other mutations were tested to inquire about the 'switchable' nature of the structure. Compared to the wild type, the G783,787U mutant was ~44% active while the G774,789U mutant was only ~17% active. These results demonstrate that the mutations that were made on the RNA still retained a viable IRES, however, with various degrees of reduction in activity. These observations raise an intriguing possibility that by utilizing different sub sets of G-stretches, the IRES can fine-tune its activity.

Engineered G-Quadruplex Can Rescue Activity of a Functionally Deficient Mutant. Since a functional G-quadruplex is required for activity of IRES-A, we rationalized that an inactive or barely active sequence, if engineered, can result in a rescue of the function. We mutated the G774,789U mutant (17% active) to design a new mutant sequence (A788G, Rescue Quad) that contains an adequate number of G-stretches to allow more flexibility in adoption of a G-quadruplex structure. The reason A788 was chosen to be mutated was because most of the other available A's if mutated would have resulted in unusually long G-stretches significantly deviating from the makeup of the wild-type sequence. The dual-luciferase activities were measured and plotted in a histogram showing percent activity of the dual-luciferase rescue mutant construct (Rescue Quad) normalized to its parent construct (G774,789U). This new mutant showed about 55% enhancement of IRES-A activity compared to the G774,789U mutant, indicating a partial rescue of the function. RNase T1 footprinting of the rescue mutant showed a protection pattern in the presence of K^+ (Figure 4) that is commensurate with the formation of a G-quadruplex structure. However, negligible if any protection was observed in the presence of an equimolar amount of Li^+. The data suggest that the engineered sequence is able to adopt a conformation that is more conducive to cap-independent initiation of translation by hVEGF IRES-A. This is the first example to our knowledge where an engineered sequence supporting and adopting a G-quadruplex structure was able to rescue translation initiation activity.

Discussion

Sequence Redundancy Guarantees the Formation of a 'Switchable' RNA G-Quadruplex Critical for IRES-A-Mediated Translation Initiation. G-quadruplex forming sequences located in 5′-UTR of mRNA have been almost exclusively found to be translation inhibitors.[30−33] This report provides a set of evidence to the contrary. The presented data indicate that G-quadruplex formation is most likely a prerequisite for successful IRES-A-mediated cap-independent translation initiation. The formation of such a structure may be facilitated by redundancy in the closely located G-stretches embedded within the IRES-A sequence. By employing systematic footprinting analyses, the 293 nt IRES-A was first narrowed down to a 17 nt (774−790) sequence that adopted a G-quadruplex structure (Figure 1A and 1B). However, it may encode for many putative G-quadruplexes, only a subset of which are depicted in Figure 1C. Analysis of the entire protected segment and at least one of the minimal G-quadruplex-forming sequences in isolation revealed that they also adopt a G-quadruplex conformation (Supporting Information Figure 2), suggesting that a quadruplex structure can be formed both in the context of the full IRES-A sequence and when analyzed in isolation. When one of the minimal G-quadruplex-forming segments was mutated to disrupt the quadruplex formation in the context of the entire IRES-A sequence (G777,781U, Figure 2A), surprisingly the mutation did not result in any change in activity compared to the wild type (Figure 2B). However, a quadruple mutant (4MVF), which lacks a sufficient number of G-stretches to adopt an intramo-

lecular G-quadruplex conformation, was completely inactive. To explain these results we propose that there is an obligatory requirement that a G-quadruplex structure must be formed for maintaining the IRES function. This emphasizes the importance of the G-quadruplex, enabling sequence on the function of IRES-A. Previously Bonal et al. reported that a G-quadruplex structure in conjunction with two stem-loop structures located within the FGF-2 IRES was needed for translation initiation.[52] However, the authors concluded that the main regulator of the IRES function was a 17-nt non-G-quadruplex region located within the domain. The putative G-quadruplex alone was insufficient to support FGF-2 IRES function.[52] Also, according to the authors, the quadruplex structure presumably was detrimental to initiation of translation at one of the internal codons, which is in consonance with the more conventional effect of the RNA G-quadruplex on translation. Thus, the G-quadruplex in hVEGF IRES-A is not only a unique example of a quadruplex exclusively required for translation initiation but may also represent a more general mechanism of the use of a 'switchable' G-quadruplex structure to ensure function by maintaining sufficient redundancy.

There may be two major reasons for the redundancy in the G-stretches within the protected sequence: (i) ensure availability of more than an adequate number of G-stretches to guarantee the formation of a G-quadruplex and (ii) form G-quadruplexes with differential ability of initiating translation, in effect regulating the translation efficiency.

In the case of the wild-type sequence, the redundancy in the G-stretches allows utilization of adjacent or nonadjacent G-stretches for G-quadruplex formation (Figure 1C). However, the quadruplex formed by the double mutant (G777,781U) must utilize nonadjacent G-stretches (Figure 1C). The formation of the quadruplex was supported by the RNase T1 footprinting and DMS modification data that showed that the cleavage pattern in the double mutant (2 Mut.) is commensurate with a G-quadruplex structure that used the available but noncontiguous stretches of Gs. These observations fit well if it is assumed that the formed G-quadruplex structure is inherently redundant in nature because to retain a quadruplex structure in the context of the doubly mutated IRES-A the original quadruplex had to 'switch' to successfully accommodate an alternate G-quadruplex structure (Figure 2C). It is a possibility that the G-quadruplexes formed by the mutant are a subset of the ensemble of structures with comparable energy formed by the wild-type sequence and thus are able to maintain the level of activity on par with the wild-type. This alternate quadruplex structure uses some of the original nucleotides and recruits new ones that impart the 'switchable' nature to the structure. The proposal assumes that not all of the G-stretches used by the double mutant (G777,781U) are used by the wild-type structure, because a G-quadruplex with shorter loops generally are preferred than a longer loop (in the case of G777,781U). The presence of proposed 'switchable' G-quadruplex structures is reinforced by the studies on two single mutants (G774U and G789U), in each of which one of the G-stretches was eliminated due to mutation, however, showed translation activity comparable to wild type (data not shown). It should be noted that these single mutations are at the opposite ends of the protected segment, and thus, the putative quadruplex structure will be formed by utilizing the G-stretches in the middle of the segment. In the case of G777,781U the first G-stretch had to be utilized; however, in the single mutant G774U that stretch is mutated but both mutants showed identical activity. This further attests to the 'switchable' nature of the quadruplex formed within 774−790. To support the 'switchable' quadruplex structures, different combinations of G-stretches and variable loops sizes need to be accommodated. Although shorter loops are preferred, G-quadruplex structures with up to a 7 nt loop has been reported.[57,58] In the double mutant, G777,781U, presumably a 6 nt loop containing a quadruplex is formed, which is supported by the footprinting data (Figure 3 and Supporting Information Figure 4). Additionally, single mutants such as G774U and G789U showed a similar level of activity as the wild type (data not shown). Although in each case one of the G-stretches was unavailable due to the mutation, presumably they were able to maintain activity by recruiting other G-stretches to form a functional G-quadruplex, further attesting to the switchable nature of the G-quadruplex.

Upon the basis of the primary sequence, the G-quadruplex is likely to adopt a two-tiered structure and may not be as stable compared to some of the reported RNA quadruplexes.[30,32] However, the very stable RNA G-quadruplexes reported previously are detrimental to translation, including a two-tiered non-natural quadruplex.[36] The putative two-tiered quadruplex described here is seemingly less stable than some of the reported naturally occurring RNA quadruplexes, and it is also unique because of the fact that it is a functionally essential motif for translation initiation activity. The lower stability may be functionally significant, as it has been demonstrated recently that stronger IRESs have weaker secondary structures.[59] The IRES structure perhaps needs to be malleable enough to adjust to conformational change as it accommodates binding to putative factors necessary for translation initiation. The redundancy in the G-rich stretches in the IRES-A sequence provides a certain degree of flexibility as well as robustness to the system because within this segment the minimum requirement of a quadruplex formation can be met in multiple ways, although it may result in different degrees of functional consequence, as will be discussed below.

G-Quadruplex Structures Formed by Strategic Choice of G-Stretches May Fine-Tune IRES-A-Mediated Translation Initiation. Upon the basis of the differential functionality of a designed set of mutants, a G-quadruplex structure that is essential for translation initiation is proposed. However, some of the mutants used in this study did not retain wild-type level of activity. For example, the double mutants G774,783U, G783,787U, and G774,789U still have the possibility to form quadruplex structures based upon analysis of their primary sequence (Figure 2C), and footprinting results support the presence of such structures (data not shown). All three mutants utilize a different combination of G-tracts to form a quadruplex, and all show translation initiation activity. The activity, however, is reduced by various degrees in comparison to the wild-type IRES activity. We propose that due to the 'switchable' nature of the quadruplex structure, different mutants may utilize different G-stretches to presumably adopt a G-quadruplex conformation but with varying degrees of impact on the translation initiation function of the IRES. Upon the basis of our data we suggest that by utilizing different subsets of G-stretches, the IRES could potentially alter its conformation within nucleotides 774−790 so as to form specific G-quadruplexes, which can potentially be a mechanism for a conforma-

(57) Dai, J.; Dexheimer, T. S.; Chen, D.; Carver, M.; Ambrus, A.; Jones, R. A.; Yang, D. *J. Am. Chem. Soc.* **2006**, *128*, 1096–1098.
(58) Phan, A. T.; Modi, Y. S.; Patel, D. J. *J. Am. Chem. Soc.* **2004**, *126*, 8710–8716.
(59) Xia, X.; Holcik, M. *PLoS ONE* **2009**, *4*, e4136.

tional switch for regulating translation initiation. Thus, at least one of the possible quadruplexes can fully support translation initiation, while a subset of them support translation to various degrees. Utilization of the different G-stretches to adopt G-quadruplex conformation, which results in modulation of the translation initiation activity to a varying level, may be a direct or an indirect result of interaction with yet to be identified factors. The mutants in that sense serve as surrogates to structures that may result as the outcome of protein factor binding. Riboswitches that essentially are stem-loop structures, which change conformation upon specific metabolite or protein binding to attenuate transcription or initiate translation, have been described.[16,20] To our knowledge, however, there are no reports of conformational switches involving G-quadruplex structures resulting from a naturally occurring sequence.

Conclusion

Despite the growing number of RNA G-quadruplexes being discovered, most of the reports show that these quadruplex structures are inhibitory to translation.[30−33] Herein, we provide evidence that a quadruplex-forming sequence is essential for IRES-A-mediated translation of the hVEGF mRNA. When mutated, the IRES-A utilizes neighboring tracts of G-stretches to maintain its activity in eukaryotic cells. The 'switchable' character makes the RNA G-quadruplex described in this report unique. Additionally, the segment may also act as a G-quadruplex-dependent 'switch' that can regulate IRES-A function, as different mutants showed a variety of levels of activities. Disruption of the quadruplex formation interferes with IRES-A function, and because of its importance in translation initiation under hypoxic stress, a condition encountered in tumor angiogenesis, the quadruplex motif can be a tumor-specific therapeutic target.

Acknowledgment. We thank Professor Scott Strobel for his comments on the manuscript. We also thank E.F.M. for editorial comments. This work was supported by a start-up grant from KSU to S.B.

Supporting Information Available: Primary nucleotide sequence of the human VEGF IRES-A, circular dichroism (CD) spectra of an oligoribonucleotide encompassing the protected region in hVEGF IRES A, histograms representing quantitation of the gel shown in Figure 3A, and histogram representing the ratio of *Renilla* to firefly luciferase activities in HeLa cells along with the ratio of CT values for *Renilla* to firefly determined by qRT-PCR. This material is available free of charge via the Internet at http://pubs.acs.org.

JA106287X