

Research Article

Effects of 5-Hydroxymethylcytosine Epigenetic Modification on the Stability and Molecular Recognition of VEGF i-Motif and G-Quadruplex Structures

Rhianna K. Morgan ¹, Michael M. Molnar,² Harshul Batra,¹ Bethany Summerford ², Randy M. Wadkins ² and Tracy A. Brooks ^{1,3}

¹School of Pharmacy, Department of BioMolecular Sciences, Division of Pharmacology, University of Mississippi, University, MS 38677, USA

²Department of Chemistry and Biochemistry, University of Mississippi, University, MS 38677, USA

³School of Pharmacy and Pharmaceutical Sciences, Department of Pharmaceutical Sciences, Binghamton University, Binghamton, NY 13902, USA

Correspondence should be addressed to Tracy A. Brooks; tbrooks@binghamton.edu

Received 21 February 2018; Revised 31 March 2018; Accepted 3 April 2018; Published 16 May 2018

Academic Editor: Ramon Eritja

Copyright © 2018 Rhianna K. Morgan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Promoters often contain asymmetric G- and C-rich strands, in which the cytosines are prone to epigenetic modification via methylation (5-mC) and 5-hydroxymethylation (5-hmC). These sequences can also form four-stranded G-quadruplex (G4) or i-motif (iM) secondary structures. Although the requisite sequences for epigenetic modulation and iM/G4 formation are similar and can overlap, they are unlikely to coexist. Despite 5-hmC being an oxidation product of 5-mC, the two modified bases cluster at distinct loci. This study focuses on the intersection of G4/iM formation and 5-hmC modification using the vascular endothelial growth factor (VEGF) gene promoter's CpG sites and examines whether incorporation of 5-hmC into iM/G4 structures had any physicochemical effect on formation, stability, or recognition by nucleolin or the cationic porphyrin, TMPyP4. No marked changes were found in the formation or stability of iM and G4 structures; however, changes in recognition by nucleolin or TMPyP4 occurred with 5-hmC modification wherein protein and compound binding to 5-hmC modified G4s was notably reduced. G4/iM structures in the VEGF promoter are promising therapeutic targets for antiangiogenic therapy, and this work contributes to a comprehensive understanding of their governing principles related to potential transcriptional control and targeting.

1. Introduction

Guanine- and cytosine-rich DNAs are capable of forming higher order, noncanonical, intramolecular structures known as G-quadruplexes (G4s) and i-motifs (iMs), respectively. iM/G4 forming sequences are common throughout the genome with high abundances in promoters and 5' untranslated regions (UTRs) of both oncogenes and tumor suppressor genes [1, 2]. These structures are associated with an array of human diseases, including cancer and neurodegeneration [3, 4]. They are also highly polymorphic, both across regions of DNA and within one sequence, and vary in the number of intercalated-cytosines and stacked G-tetrads, as well as in loop length and directionality [5]. It has been suggested that

short loop lengths are favorable for G4s, while iMs are stable with long loops, but there are many more factors playing a role in these formations [5]. Noncanonical DNA structures have been characterized as transcriptional regulators, acting as either activators or suppressors depending on the gene involved. Most iMs and G4s identified to date act as transcriptional suppressors by hindering RNA polymerase activity and/or transcription factor binding. Such structures span the hallmarks of cancer; thus, most research focuses on utilizing small iM/G4-stabilizing molecules in order to decrease tumor activity [3–7].

During replication or transcription, the complementary strands of gene promoters are separated. Separation of the double-stranded DNA (dsDNA) helix to single-stranded

DNA (ssDNA) has the potential to form iM/G4 structures [8–11]. At physiological pH, both iMs and G4s have previously been described to form with dehydrating and crowding reagents or torsional stress aiding in the stability of structure formation [12–14]. Overall, the topology of iMs and G4s depends on the environment, specifically pH, dehydrating conditions, and epigenetics [13, 15, 16]. Noncanonical DNA structures can form in an equilibrium with B-DNA duplexes, depending on the sequence and surrounding nuclear conditions. As the interest in iM/G4 function in biological systems grows, understanding the structural dynamics in the presence of epigenetic modification will help to clarify what occurs in the cellular milieu.

Modifications, such as methylation or hydroxymethylation, to bases in a DNA sequence are epigenetic factors regulating gene expression. Unique regions in a genome prone to these types of modifications are clustered cytosine-guanine-rich sequences, known as CpG islands. These islands consist of repeats of a cytosine being placed next to a guanine in the same sequence with only a phosphate separating the two nucleotides [17, 18]. The occurrence of CpG islands in mammals is quite low, less than 1% in humans [19]. DNA methyltransferase (DNMT) enzymes, usually with the methyl group transferred from S-adenosylmethionine (SAM), generate the methylated cytosine [20]. The first step of cytosine demethylation is conversion of the methyl (5-mC) to a hydroxymethyl (5-hmC) by the ten-eleven-translocation (TET) proteins. From there, the cytosine is further modified until the base is removed and DNA repair proteins fill the basic site with an unmodified cytosine [21, 22]. When CpG islands occur, modification of a cytosine increases the probability of altering gene expression [23]. Transcriptional activity is positively correlated with 5-hmC density and negatively correlated with 5-mC density [24, 25].

In spite of 5-hmC being a modification of 5-mC, these two epigenetically modified bases have been described as having differential genomic density and to be nonoverlapping. Often, but not always, CpG islands and iM/G4 formation are capable of occurring in the same sequence; however, 5-mC modification is generally not found to be overlapping with iM/G4-forming sequences [20, 26]. A study of the effects of modification on a G4 found within the Bcl-2 promoter demonstrated a stabilization of the noncanonical structure, highlighting the need to study effects on individual structures within promoters where modification can potentially occur [27].

Vascular endothelial growth factor (VEGF) produces a protein that stimulates angiogenesis allowing oxygen restoration to tissue [28]. When overexpressed, VEGF is responsible for the rapid development of new blood vessels to cells, specifically tumor cells. This can cause vascular disease and can help tumors metastasize. VEGF is a key player in tumorigenesis for many cancer cell types making it a promising target for drug therapy. The promoter within this oncogene contains a main cis-element located –85 to –50 bp from the transcriptional start site (TSS). This element is an iM/G4-forming region consisting of five contiguous runs of cytosines/guanines wherein iM and G4 structures form from four contiguous runs, as described in ssDNA and supercoiled plasmids [12, 14, 29–31]; the fifth run has been postulated to

have a role in base-excision repair, as the first four runs are resistant [32]. Besides enabling higher order formation, this G/C-rich region close to the TSS contains CpG sequences favorable for cytosine hydroxymethylation. Given the generalized mutual exclusivity of 5-mC and 5-hmC, and also of 5-mC and iM/G4 formation (Figure 1(a)), this study sought to understand the potential effects of 5-hmC modification on the iM and G4 formations of the VEGF promoter. We utilized G- or C-rich G4- or iM-forming oligonucleotides from this sequence containing 5-hmC modifications, as compared to wild-type (iM_{WT} or G4₁₋₄), on different relevant cytosines involved in hemiprotonated C:C base-pairing for the iM, or in loop formation within the G4 (Figure 1(b)). Using this model, we sought to elucidate the effects of molecular crowding and dehydration [13, 33, 34], in combination with cationic strength and varying pH, on iM and G4 structure formation, stability, and recognition by proteins and compounds.

2. Materials and Methods

2.1. Materials. All oligonucleotides were synthesized and purchased from Midland Certified Reagent Co., Inc. (Midland, TX) (Table 1). iM oligonucleotides were stored in 10 mM Tris, 1 mM EDTA buffer at pH 8.0 at –20°C. G4s were stored in autoclaved dH₂O at 4°C. Acrylamide/bisacrylamide (29:1) solution and ammonium persulfate were purchased from Bio-Rad laboratories (Hercules, CA), and N,N,N',N'-tetramethylethylenediamine was purchased through Fisher Scientific (Pittsburgh, PA). The GelGreen nucleic acid stain was acquired from Phenix Research Products (Candler, NC). All other chemicals were purchased from Sigma-Aldrich (St. Louis, MO).

2.2. UV Visual Spectroscopy. The ECD thermal melt profiles of the iM oligonucleotides (2–5 μM) in 30 mM cacodylate buffers at varying pHs (5.4–8.0) with and without various amounts of KCl and acetonitrile (ACN) were validated using absorbance readings at 260 nm collected on Cary 100 UV-Visible Spectrometer (Agilent Technologies, Santa Clara, CA); the pH of each sample was measured in a pH meter immediately before placement in the spectrophotometer. Each solution was first heated to 80°C for 5 minutes and then cooled to 4°C for 2 minutes for two cycles prior to experimental data collection to remove mismatched dimers. The iM thermal denaturation was also monitored by UV-Vis spectroscopy on a Cary 100 UV-Visible Spectrometer (Agilent Technologies, Santa Clara, CA). Recordings were made over the wavelength range 225–350 nm. For thermal denaturation, increasing temperatures of 20–100°C with a ramp rate of 1°C per minute with a 1 minute hold per degree were used [35, 36]. Thermal melt values were determined by nonlinear regression fitting on KaleidaGraph software (Synergy Software, Reading, PA) of the spectral data to determine the midpoint of the transition (T_M) and ΔH for the folded to unfolded transition.

2.3. Electronic Circular Dichroism (ECD). The G4 oligonucleotides (5 μM) were fixed in 50 mM Tris-HCl (pH 7.4) and iM solutions in 30 mM cacodylate buffers ranging in

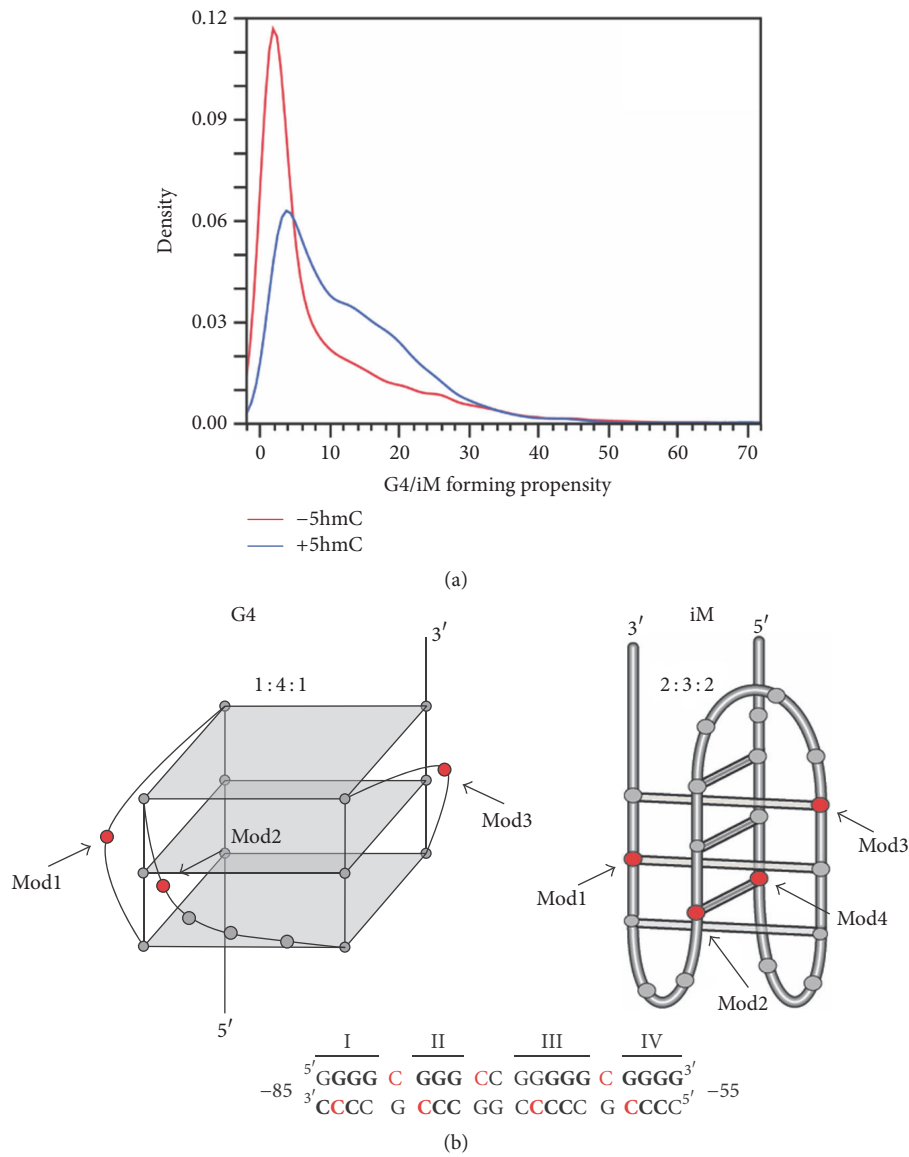


FIGURE 1

TABLE 1: The promoter sequence (24 mer) of the VEGF gene and its modifications for both G4/iM-forming regions. X denotes the 5-hydroxymethylated cytosine of each oligonucleotide sequence.

iM_{WT}	5' GACCCCGCCCCGGCCCCGCCCCGG3'
iM_{mod1}	5' GACCCXGCCCCGGCCCCGCCCCGG3'
iM_{mod2}	5' GACCCCGCCCCXCGGCCCCGCCCCGG3'
iM_{mod3}	5' GACCCCGCCCCGGCCXGCCCCGG3'
iM_{mod4}	5' GACCCCGCCCCGGCCCCGCCXCGG3'
$G4_{MT}$	5' CCGTTGCGTGCCGTTTGCGTTGTC3'
$G4_{1-4}$	5' CCGGGGCGGGCCGGGGGCGGGGTC3'
$G4_{mod1}$	5' CCGGGGXGGGCCGGGGGCGGGGTC3'
$G4_{mod2}$	5' CCGGGGCGGGXCXGGGGGCGGGGTC3'
$G4_{mod3}$	5' CCGGGGCGGGCCGGGGGXGGGGTC3'

pH from 5.4 to 8.0, both with varying amounts of KCl and ACN, and 10 μM TMPyP4. Additionally, TMPyP4 effects on G4 stability were examined. Spectra were collected with an Olis DSM-20 spectropolarimeter equipped with a CD 250 Peltier cell holder (Bogart, GA). Recordings were made over the wavelength range 225–350 nm and at increasing temperatures (20–100°C, every 10°C, with a 1 minute hold at temperature before spectra were recorded) in 1 mm quartz cuvette. The ordinate is reported as molar ellipticity (millidegrees); T_M s were determined by performing a singular value decomposition (SVD) [37] analysis available with the Olis GlobalWorks software, followed by nonlinear regression fitting to determine T_M using GraphPad Prism software (La Jolla, CA). The iM pK_a and cooperativity of the pH driven folding (Hill coefficient) were similarly determined from fitting of the Hill equation [38] using the ECD signal obtained at 298 nm plotted versus pH.

2.4. Isolation of Nucleolin Nuclear Extract. HEK-293 cells were plated at a density of 10×10^6 cells per well in a 60 mm dish in 6 mL of Dulbecco complete growth medium (DMEM) with 1x penicillin/streptomycin solution and 10% fetal bovine serum. Cells were transfected with 20 μg of nucleolin expressing plasmid upon reaching approximately 70–80% confluency. After 48 hrs, nuclear extraction was performed using the Thermo Scientific NE-PER Nuclear and Cytoplasmic Extraction kit. Briefly, cells were harvested with trypsin-EDTA and centrifuged at 4,000 rpm. The pellet was resuspended in 1x PBS and transferred to a 1.5 mL microcentrifuge tube before pelleting by centrifugation at $500 \times g$ for 2–3 mins. Ice-cold CER I was added and vortexed vigorously for 15 sec before being incubated on ice for 10 mins. Again, ice-cold CER II was added and vortexed for 5 seconds before a second incubation on ice for 1 min. Next, the sample was centrifuged for 5 mins at $16,000 \times g$. The cytoplasmic extract was discarded and NER was added to the pellet. The sample was vortexed for 15 sec every 10 mins, for a total of 40 mins. Lastly, the sample was centrifuged at $16,000 \times g$ for 10 mins, and the supernatant (nuclear extract) was collected and stored at -80°C until use.

2.5. Electrophoretic Mobility Shift Assay (EMSA). Non-FAM-labeled oligonucleotides were prepared with 140 mM KCl, 10 mM Tris-HCl (pH 5.4, 6, or 7.4) and 30–40% (v/v) ACN. The solutions were denatured by heating to 95°C for 5 minutes and then rapidly cooled for 5 minutes and repeated for five cycles to induce G4/iM formation. Upon addition of GelGreen nucleic acid stain and nondenaturing loading dye, the samples were loaded on a 10% native polyacrylamide gel. After running at 100 V, the gel was visualized under blue light LED using a FOTO/Analyst® Investigator FX Imager (Hartland, WI). A G4 knockout oligonucleotide was utilized as a linear control for G4 and iM migration. Nucleolin (0–30 μg) protein/G4 binding was examined using the same non-FAM-labeled oligonucleotides (4 μM) prepared in 2x binding buffer (40 mM Tris-HCl, pH 7.5, 40 mM NaCl, 10 mM KCl, 10% (w/v) glycerol, 2 mM EDTA, and 2 mM DTT), 1 $\mu\text{g}/\mu\text{L}$ BSA, and 0.5 $\mu\text{g}/\mu\text{L}$ calf thymus DNA. The samples were incubated on ice for 30 min prior to the addition of oligonucleotides.

Following the addition, samples were incubated on ice again for 20 min. Eva green nucleic acid stain and nondenaturing loading dye were added to each sample and ran on a 0.5% agarose gel at 100 V for 40 min at 4°C . The gel was visualized using the same blue light LED of the imager aforementioned.

3. Results

3.1. 5-hmC Modification Effects on the pK_a of the VEGF Promoter iM. Previously, we noted that the inclusion of a single 5-hmC residue in a mutant iM-forming sequence from the c-MYC gene promoter greatly enhanced the cooperativity of folding of the sequence in response to pH [18]. Although the cytosine queried previously was not in a CpG island and thus is unlikely to be modified under normal intracellular conditions, this cooperative folding was found with 5-hmC, but not 5-mC. This observation led to us to expand the 5-hmC work in a more physiologically relevant sequence and to study both the C- and G-rich strand formations. Here, we examined four different CpG positions in the C-rich sequence of DNA from the VEGF promoter that is known to fold into an iM [10]. The ssDNA oligonucleotides containing modifications were identified based on CpG mono- or di-nucleotide motifs. Cytosines modified for iM_{mod2} and iM_{mod4} (and G4_{mod2} described below) were chosen because cytosine modifications occur on the most distal nucleotide in a CCG₍₁₋₂₎ motif [25].

The DNA's ability to fold into its iM when a 5-hmC modification is present on the oligonucleotide was examined for pH dependence over the range 5.0 to 8.1 using electronic circular dichroism (ECD). The maximum circular dichroism signal was seen at 290 nm and minimum at 260 nm indicating iM formation occurs when 5-hmC modification is present at each CpG site (Figure S1). Using the maximum signal at 290 nm, the change in ellipticity as pH increased was used to determine the pK_a and Hill coefficient. The pK_a of an iM is generally accepted to be defined as the pH where 50% of the DNA strands in the sample are folded. Each 5-hmC modified and control sequences had a pK_a at 6.05 ± 0.06 . Also, in contrast to our observations with the iM from c-MYC [20], we did not observe a significant shift in the cooperativity of the pH-induced denaturation; all strands had a fitted Hill coefficient of -1.54 to -2.17 , with no discernable trend between Hill coefficient value and 5-hmC location (Table 2).

3.2. 5-hmC Modification Effects on Thermal Stability and Equilibrium. Thermal melting (T_M) experiments were done at pH 5.4 on modified oligonucleotides in the absence or presence of KCl alone or with the dehydrating agent, acetonitrile (ACN) (Table 3). No notable changes in thermal stability, as defined by T_M , occurred with epigenetic modification or cosolvent condition variation. As a general trend, ΔH decreased with increasing KCl concentrations by 10–22%, although there were exceptions with modifications at the second and third positions whose ΔH increased by up to 32%. The addition of ACN generally increased the ΔH by 20–30%, with the exception of a 14% decrease with iM_{mod2} in the presence of 100 mM KCl. From this experiment, it can be concluded that there are no global effects on iM

TABLE 2: Change in molar ellipticity (290 nm) of the electronic circular dichroism (ECD) spectra over the pH range of 5.0 to 8.1. ECD data was used to determine pK_a using singular value decomposition (SVD) analysis of pH-dependent folding curves for each 5-hmC modified sample and compared to wild-type control. The Hill coefficient is the fitting parameter that characterizes the cooperativity of folding.

	iM_{WT}	iM_{mod1}	iM_{mod2}	iM_{mod3}	iM_{mod4}
pKa	6.10 ± 0.04	6.00 ± 0.15	6.10 ± 0.11	6.10 ± 0.05	6.10 ± 0.04
Hill Coefficient	-2.06	-1.60	-1.54	-1.91	-2.17

thermal stability induced by either 5-hmC modification or cosolvent condition; few notable changes were noted in folding dynamics, particularly with 5-hmC incorporation into positions two or three in the presence of KCl.

Similar thermal melting studies of iMs were done at pH 6.1 (the pK_a). The ensemble of wild-type and modified iM oligonucleotides require less thermal energy to unfold (Table 4) compared to pH 5.4, since half of the ensemble was unfolded due to pH. At pH 6.1, there is also a general trend of lower thermal stability in conditions containing KCl across all oligonucleotides. The addition of KCl also decreased ΔH of iM_{mod1} , iM_{mod2} , and iM_{mod4} by more than 50%. Moreover, the combination of KCl and ACN in iM samples decreased both T_M and ΔH , as compared to either milieu change alone or versus control (Table 4 and Figures S2–S11).

3.3. 5-hmC Modification Effects on G4 Formation and Stability. Overall G4 formation and stability, as affected by 5-hmC modifications, was examined by ECD. The wild-type form, $G4_{1-4}$, forms a prominent parallel structure as evidenced by the positive cotton effect at 265 nm (Figure S12(A)) [12, 30, 31, 39]. G4 formation was induced in the presence of cations and osmolytes, using 100 mM KCl and 40% ACN to induce very stable structures as a measure of maximal formation (Table 5). $G4_{1-4}$ thermal stability increased from 62 to 84°C in 100 mM KCl and exceeded 100°C with the addition of 40% ACN. Similarly, $G4_{mod1}$ increased from 50 to 82°C with 100 mM KCl, and $G4_{mod3}$ increased from 73 to 84°C. Both of these latter modifications have increased thermal stability in the presence of 40% ACN to 95°C or more. $G4_{mod2}$, however, does not form a strong structure in the absence of any cosolvents; a parallel orientation is only apparent when KCl and ACN are added (Figure S12(A)). In the presence of 100 mM KCl, $G4_{mod2}$ has a melting temperature of 76°C, which increased to greater than 100°C with 40% ACN. A knockout oligonucleotide, $G4_{MT}$, with G-to-T mutations in each guanine run within the wild-type sequence was used as a negative control; disruption of G4 formation was confirmed by ECD in the presence of 100 mM KCl with and without 40% ACN (Figure S12(B)).

Further studies were performed to examine G4 formation in the presence of cosolvents to mimic more physiologically relevant conditions, such as dehydration and structures existing in an equilibrium with unfolded DNA. Appropriate cosolvent concentrations were determined by evaluating the minimal concentration necessary (0–40% (v/v) ACN) to induce marked G4 formation, as determined by ECD spectra of $G4_{1-4}$ (Figure S13). The concentrations of 10 mM KCl and 30% ACN were chosen to investigate modified G4 formation as moderately stable conditions (analogous to pK_a for iM

structures). We monitored the effect of these cosolvents on G4 stability with epigenetic modifications present compared to wild-type (Figure S14). $G4_{1-4}$ thermal stability increased from 67 to 83°C upon addition of 30% ACN. Similar to wild-type, the presence of KCl and/or ACN increased the stability of G4s containing 5-hmC modifications as well (Table 6). $G4_{mod1}$ was enhanced from 68°C with 10 mM KCl to 80°C with 30% ACN. Melting temperatures of $G4_{mod2}$ increased from 68 to 84°C upon ACN addition, as compared to KCl control. $G4_{mod3}$, like $G4_{mod1}$, increased stability from 67 to 80°C with ACN. Cumulatively, a change in G4 stability in the presence of an epigenetic modification was only noted with $G4_{mod2}$ in 100 mM KCl (T_M of 76°C) as compared to wild-type $G4_{1-4}$ (T_M of 84°C). Overall, 5-hmC modification did not markedly affect G4 formation or stability.

3.4. Inter- and Intramolecular Higher Order DNA Formations. Electrophoretic mobility shift assays (EMSA) were then used to determine inter- versus intramolecular iM and G4 structure formation (Figure 2). The knockout oligonucleotide was used as a linear reference to align the iM EMSAs and to indicate G4 migration patterns. Each iM sample was examined at distinct pHs in order to determine the dependence of C-rich structure formation on molecular interactions with surrounding protons. Each G4 sample was examined in the absence and presence of monovalent cation to stabilize the structure. Both iM and G4 structures were also examined in the presence of dehydrating cosolvent (40% ACN).

Generally, iM species at pH 5.4 migrated faster than the same sequences at pH 8.0, where they have been determined to be linear ssDNA by ECD (Figure S1). All species at pH 6.0 (the pK_a) span the distance between the pH 5.4 and the pH 8.0 bands, in agreement with a mixed structured and unstructured population. None of the 5-hmC modifications markedly changed migration patterns, nor did the addition of 40% ACN (Figure 2(a)).

$G4_{1-4}$ and the 5-hmC-modified G4s migrated further than the knockout reference, MT, suggesting the existence of intramolecularly folded species. Further migration is seen when 100 mM KCl is added; however, the shift is not marked. When 40% ACN is added, slower migration of each sequence displays banding above the corresponding control and 100 mM KCl samples. Notably, $G4_{mod1}$, with 100 mM KCl alone and in the presence of 40% CAN, demonstrates the formation of intermolecular structures, which may be correlated to the change in enthalpy noted with $G4_{mod1}$ (Table 6). This decrease in enthalpy likely signifies an increase in the number of isoforms existing, which may include a cohort of inter- and intramolecular G4 species. No other sequence mediated the formation of intermolecular species; $G4_{mod2}$

TABLE 3: Comparison of T_M and ΔH° of transition of iM containing 5-hmC \pm KCl (100 mM) and \pm 40% acetonitrile (ACN) at pH 5.4. The unmodified control iM VEGF sequence (iM_{WT}) is also shown for comparison. Error measurements are omitted for clarity. Errors in ΔH° were less than 12% while errors in T_m values were less than 3%.

	iM _{WT}		iM _{mod1}		iM _{mod2}		iM _{mod3}		iM _{mod4}	
	ΔT_M (°C)	ΔH° (kcal/mol)	ΔT_M (°C)	ΔH° (kcal/mol)	ΔT_M (°C)	ΔH° (kcal/mol)	ΔT_M (°C)	ΔH° (kcal/mol)	ΔT_M (°C)	ΔH° (kcal/mol)
Control	61	34	62	34	62	31	61	32	62	33
100 mM KCl Alone	62	27	61	29	61	41	61	35	61	28
40% ACN Alone	63	45	63	41	65	40	63	40	65	39
100 mM KCl + 40% ACN	59	36	62	32	62	35	57	44	62	35

TABLE 4: Comparison of T_M and ΔH° of transition of iM containing 5-hmC \pm KCl (100 mM) and \pm 40% acetonitrile (ACN) at pH 6.1. The unmodified control iM VEGF sequence (iM_{WT}) is also shown for comparison. Error bars are omitted for clarity and are <7% or 50% of the values shown for T_M and ΔH° , respectively.

	iM _{WT}		iM _{mod1}		iM _{mod2}		iM _{mod3}		iM _{mod4}	
	ΔT_M ($^\circ$ C)	ΔH° (kcal/mol)	ΔT_M ($^\circ$ C)	ΔH° (kcal/mol)	ΔT_M ($^\circ$ C)	ΔH° (kcal/mol)	ΔT_M ($^\circ$ C)	ΔH° (kcal/mol)	ΔT_M ($^\circ$ C)	ΔH° (kcal/mol)
Control	52	16	52	17	53	15	47	12	51	16
100 mM KCl Alone	48	18	49	8	48	7	44	13	49	7
40% ACN Alone	50	20	48	17	50	18	46	12	49	16
100 mM KCl + 40% ACN	40	10	40	9	41	10	40	5	42	12

TABLE 5: Thermodynamic properties of VEGF G4 modification sequences in the presence and absence of 100 mM KCl and 40% ACN alone or combined. Marked increases in thermal stability were noted upon addition of 100 mM KCl alone and with 40% ACN as compared to their own controls. Thermal stability was not substantially different between 5-hmC modifications and wild-type. No pronounced changes in transition state dynamics were observed. Error bars are omitted for clarity and are <10% of the values shown; ND = not determinable.

	G ⁴ ₁₋₄		G ⁴ _{mod1}		G ⁴ _{mod2}		G ⁴ _{mod3}	
	ΔT_M (°C)	ΔH° (kcal/mol)	ΔT_M (°C)	ΔH° (kcal/mol)	ΔT_M (°C)	ΔH° (kcal/mol)	ΔT_M (°C)	ΔH° (kcal/mol)
Control	62	21	50	12	ND	ND	73	16
100 mM KCl	84	22	82	27	76	30	84	15
100 mM KCl + 40% ACN	>100	25	>100	19	>100	20	95	12

TABLE 6: Thermal stability of VEGF G4 modification sequences at lower concentrations of cosolvents. 30% ACN increased thermal profiles of each modification as compared to their respective controls in 10 mM KCl. No marked changes in enthalpy were observed. Error bars are omitted for clarity and are <10% of the values shown; ND = not determinable.

	G ⁴ ₁₋₄			G ⁴ _{mod1}			G ⁴ _{mod2}			G ⁴ _{mod3}		
	ΔT_M (°C)	ΔH° (kcal/mol)	ΔT_M (°C)	ΔH° (kcal/mol)	ΔT_M (°C)	ΔH° (kcal/mol)	ΔT_M (°C)	ΔH° (kcal/mol)	ΔT_M (°C)	ΔH° (kcal/mol)	ΔT_M (°C)	ΔH° (kcal/mol)
Control	62	21	50	12	ND	73	16					
10 mM KCl	67	29	68	29	68	67	25					
10 mM KCl + 30% ACN	83	33	80	24	84	80	20					

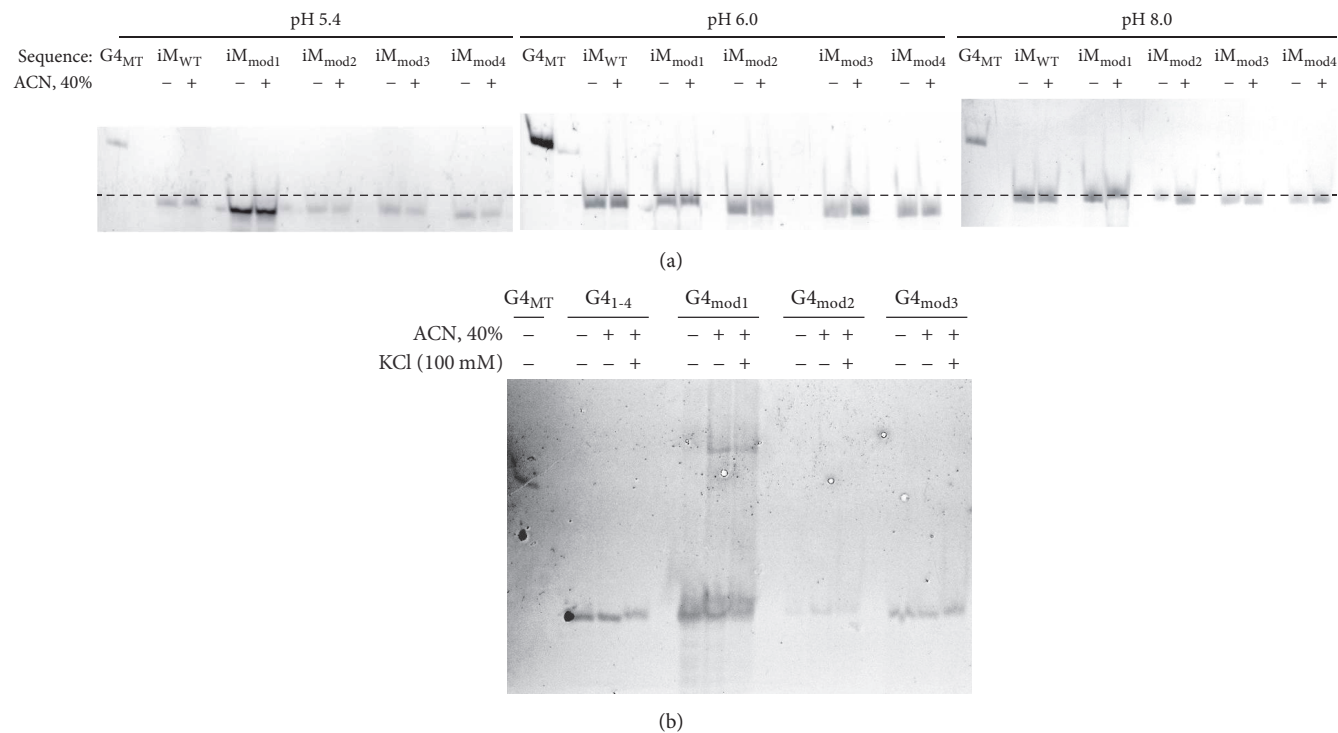


FIGURE 2

and G₄_{mod3} closely resembled wild-type G₄₁₋₄ (Figure 2(b)). Collectively, 5-hmC modifications on cytosines involved in iM hemiprotonated C:C bonds or in loop formation of G4s do not impair the stability or formation of these DNA structures.

3.5. 5-hmC Modification Modulates G4 Recognition. We further examined whether epigenetic modifications impact structure recognition by G4-stabilizing small molecules and regulatory binding proteins. The universal G4-stabilizing cationic porphyrin, TMPyP4, was used to monitor compound recognition (Figure 3(a)). Previous data demonstrated stabilization of the VEGF G₄₁₋₄ by TMPyP4, with a related decrease in transcription [27, 28]. TMPyP4 (10 μ M) was incubated with the VEGF G4 with and without 5-hmC modifications; the compound increased both G₄₁₋₄ and G₄_{mod1} stability by 12°C; in comparison, neither G₄_{mod2} nor G₄_{mod3} were stabilized by the porphyrin. No changes in ΔH were observed across the WT or modified sequences (Figure 3(b)).

To investigate the effect of 5-hmC modification on protein recognition, the binding of nucleolin was examined. Nucleolin has been reported to bind the VEGF promoter in both negative supercoiled plasmids and chromosomal DNA [14]. Nuclear extract from HEK-293 cells transfected with nucleolin was incubated with G4s. EMSA was performed to visualize protein: DNA binding. Nucleolin demonstrated concentration-dependent binding to the wild-type G4 structure, with 75% of DNA bound in the presence of 30 μ g of nucleolin extract. Decreased binding to G₄_{mod1} (~30%) was seen at the higher concentrations of nucleolin (25 and 30 μ g); however, neither G₄_{mod2} nor G₄_{mod3} were bound by nucleolin

at concentrations of extract up to 30 μ g (Figure 3(c)). Collectively, this suggests that epigenetic modifications can influence molecular recognition by both small molecules, as seen with TMPyP4, or by regulatory binding proteins, as observed with nucleolin. The propensity of these regions to exist in transient states of epigenetic modification, and its effects on higher order DNA formation and recognition, is a field that requires more study. From the presented data; however, it is clear that G4s, and likely iMs, with CpG islands demonstrate alternative recognition patterns that ought to be taken into consideration in drug discovery research and in mechanistic studies of developing compounds in vitro and in vivo.

4. Discussion

G/C-rich DNA is capable of forming higher order DNA structures and is often a template for epigenetic modification. It has been shown that methylated cytosines are generally nonoverlapping with G4-forming regions [26], but when they coincide, there is a stabilization of the G4 [27]. In addition, methylated cytosines and hydroxymethylated cytosines, despite being modifications of each other, are found in mutually exclusive regions of the genome and differentially correlate with transcriptional activity. Given the exclusivity of 5-mC with both iM/G4-forming regions and with 5-hmC, the overall purpose of this study was to examine the effects of 5-hmC modification on the formation, stability, and recognition of secondary DNA structures. As a physiologically relevant model, we focused on the structures in the promoter of the VEGF gene, which have

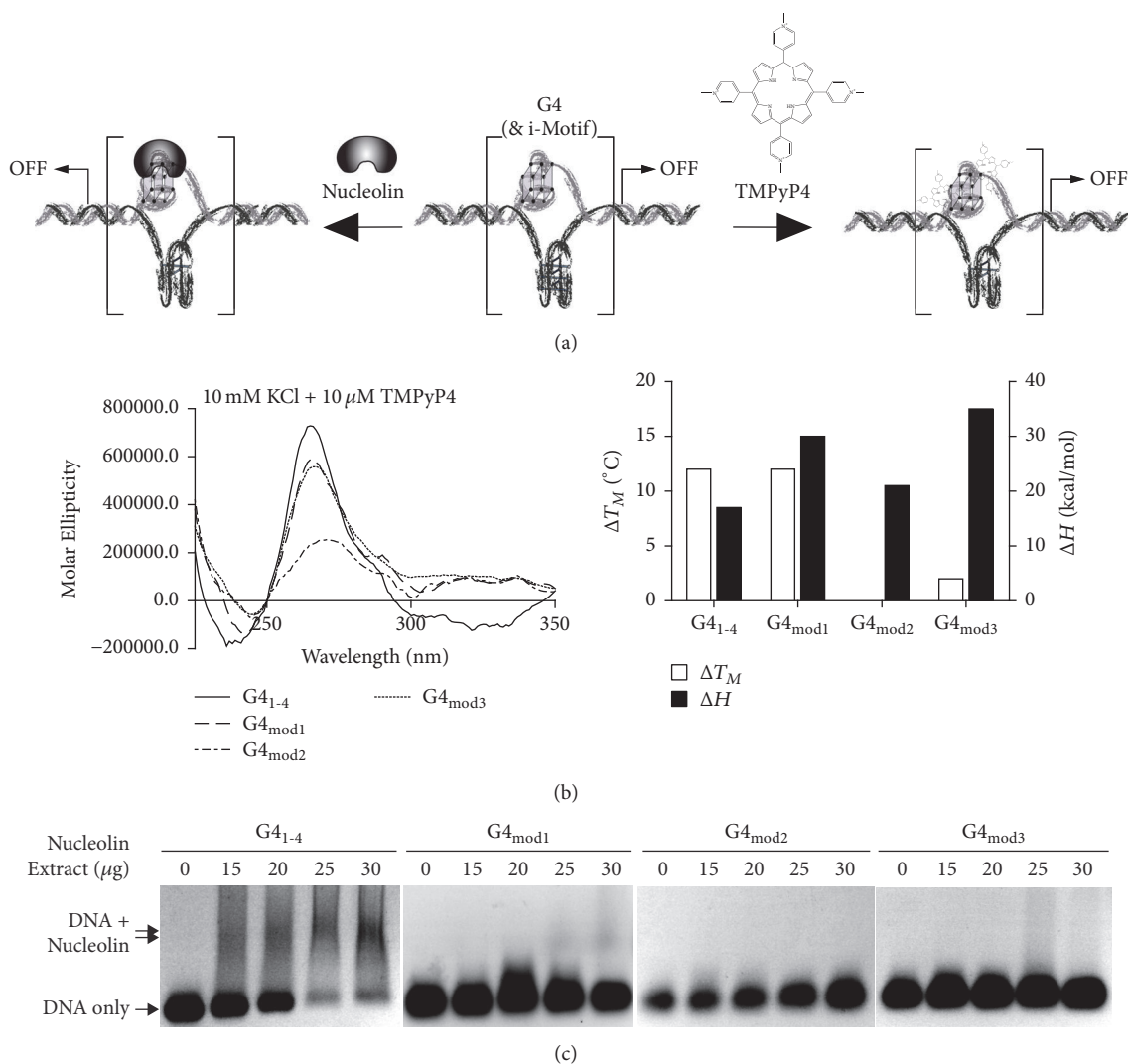


FIGURE 3

been well-described for both iM and G4 formation [12, 14, 29–31] and containing natural CpG islands [40]. Cytosine modification of this promoter has been shown to vary with pathophysiological states [41, 42]. Moreover, transcription is negatively regulated by methylation and positively regulated by hydroxymethylation, which have specifically been shown to be mutually exclusive in the VEGF promoter [41, 42].

Overall, 5-hmC modifications did not markedly affect iM or G4 formation or stability. The most notable structural affect was the identification of intermolecular G4 formation with modification of the most 5' CpG island. The recognition of the G4 by a stabilizing compound, TMPyP4 [30, 31], and protein, nucleolin [14], was also examined. Remarkably, with the exception of the intermolecular forming G4_{mod1}, TMPyP4 stabilization was abolished by 5-hmC modification, as was nucleolin recognition; protein recognition of G4_{mod1} was markedly reduced. Similar studies were not performed with the VEGF iM, as stabilizing proteins and selective compounds are not fully established, as they are for the G4.

Our studies suggest that even in the absence of marked structural iM/G4 changes with epigenetic modification, overall biological effects related to secondary structure regulation of transcription may still occur due to a change in endogenous and exogenous recognition.

Although 5-mC and G4 formation have been shown to be nonconcurrent [26], a study was undertaken to examine the effect of methylation on G4 structure, formation, and function in the Bcl-2 promoter [27]. This region also contains naturally occurring CpGs, capable of being methylated. In that study, methylation of all four internal cytosines that are part of CpGs altered the G4 structure, converting it from a mixed species to a parallel formation, and increasing its thermal stability by 9°C. Previous works on a nonbiologically occurring intermolecular G4 demonstrated that trimethylation of loop cytosines formed a more stable structure than the unmethylated version [38]. iM studies with a modified sequence from the MYC promoter, which does not contain any native CpG sequences, demonstrated slight enhancement

in thermal stability and pK_a with a single 5-mC modified position, and slightly decreased stability (both thermal and pH-dependent) with a 5-hmC modification at the same location, both in the absence of cosolvents. In the presence of varying cosolvents, any effects on thermal stability mediated by epigenetic changes were lost [20]. It is important to note that G4-forming regions within promoters are more likely to contain 5-hmC than 5-mC modifications (Figure 1(a)) [20, 26]. In the study presented here, we focused only on the 5-hmC modification and examined the effect on both the C-rich (iM-forming) and G-rich (G4-forming) sequences with naturally occurring CpGs. Modifying one cytosine at a time, we monitored pH-dependency, thermal stability, and the effect of cosolvents. Overall, our findings demonstrate a lack of marked changes in G4 or iM formation. The effect of cytosine modification on iM/G4 formation is largely understudied. Effects may be sequence-specific, with variations in each promoter sequence; structure specific, with varying effects on intermolecular, mixed, or antiparallel-predominant structures; or epigenetic modification specific, varying with methylation and hydroxymethylation. To our knowledge, the present study is the first to examine the effects of epigenetic modification on compound and protein recognition, and more work is needed to be established if the lack of G4 recognition with 5-hmC modifications in the loop region is a sequence-specific effect, an epigenetic modification-specific effect, or a consistent phenomenon.

5. Conclusion

Previous literature described epigenetic modification of CpG islands in induced pluripotent stem cells and endothelial stem cells and their correlation to gene expression [25]. Endothelial stem cells play key roles in vasculogenesis and angiogenesis, as they form the lining of blood vessels. VEGF is overexpressed by tumors, secreted into the local milieu, and acts upon nearby endothelial cells to trigger angiogenesis in order to increase oxygen levels and provide essential nutrients, mediating tumor growth, invasion, and metastasis [43–45]. Pathological states, such as intracerebral hemorrhage and preterm preeclampsia, have been shown to modulate VEGF expression through epigenetic modulation of the promoter [41, 42]. In particular, methylation of the VEGF promoter has been shown to correlate with a decrease in transcription, and hydroxymethylation coincides with an increase in gene expression [42]. Here, we investigated the iM/G4-forming region within the cis-element on the VEGF promoter which has previously been identified as an antiangiogenic target for therapeutic development [14, 30, 39, 46, 47]. The iM forms a tri-intercalated structure with loops lengths of 2 : 3 : 2 and the G4 forms a tristacked structure with parallel loop configurations of lengths 1 : 4 : 1; stabilization of secondary DNA structures in the promoter has been shown to silence transcription [12, 31]. The iM- and G4-forming sequences contain several CpG sites, so for drug discovery purposes, it is necessary to explore physiological factors, like epigenetic modifications, that can occur in CpG islands that may affect iM/G4 formation and stability [40]. The current study examined the effects of 5-hmC modifications on the VEGF

iM/G4-forming region within the promoter and learned that such modifications do not impact overall G4/iM formation. However, these modifications do alter G4 structure recognition by a pan-stabilizing compound, TMPyP4, and the regulatory binding protein, nucleolin. This suggests that drug discovery efforts should consider the presence of epigenetic modifications on the G4, and likely the iM, of interest to consider all potential biological interactions. These studies add to the understanding of the mechanism behind gene regulation via epigenetic modifications.

Data Availability

Either all data referred to in our manuscript are presented or the references are implicated with the data available in the published figures/paper.

Disclosure

Preliminary data generated for this project has been previously published as an undergraduate Honors College thesis [48].

Conflicts of Interest

The authors have no conflicts of interest.

Acknowledgments

This research was supported by National Institute of Health/ National Cancer Institute Grant 1R15CA173667-01A1 (Tracy A. Brooks, PI; Randy M. Wadkins, co-PI).

Supplementary Materials

Figure S1: pH dependence of iM formation in the native (iM_{WT}) and modified (iM_{mod1-4}) sequences. ECD spectra of all C-rich sequences were obtained at a range of pHs (5.0–8.1) at 20°C. Figure S2: ECD spectra of VEGF iM_{WT} at pH 5.4 exposed to 0 KCl (mM)/0% acetonitrile, 100 mM KCl by itself, 40% acetonitrile by itself, and both 100 mM KCl/40% acetonitrile present. Each spectrum was collected at 20°C, the determined T_M , and 90°C. Figure S3: ECD spectra of VEGF iM_{mod1} at pH 5.4 exposed to 0 KCl (mM)/0% acetonitrile, 100 mM KCl by itself, 40% acetonitrile by itself, and both 100 mM KCl/40% acetonitrile present. Each spectrum was collected at 20°C, the determined T_M , and 90°C. Figure S4: ECD spectra of VEGF iM_{mod2} at pH 5.4 exposed to 0 KCl (mM)/0% acetonitrile, 100 mM KCl by itself, 40% acetonitrile by itself, and both 100 mM KCl/40% acetonitrile present. Each spectrum was collected at 20°C, the determined T_M , and 90°C. Figure S5: ECD spectra of VEGF iM_{mod3} at pH 5.4 exposed to 0 KCl (mM)/0% acetonitrile, 100 mM KCl by itself, 40% acetonitrile by itself, and both 100 mM KCl/40% acetonitrile present. Each spectrum was collected at 20°C, the determined T_M , and 90°C. Figure S6: ECD spectra of VEGF iM_{mod4} at pH 5.4 exposed to 0 KCl (mM)/0% acetonitrile, 100 mM KCl by itself, 40% acetonitrile by itself, and both 100 mM KCl/40% acetonitrile present. Each spectrum was collected at 20°C, the determined T_M , and 90°C. Figure S7:

ECD spectra of VEGF iM_{WT} at the pKa exposed to 0 KCl (mM)/0% acetonitrile, 100 mM KCl by itself, 40% acetonitrile by itself, and both 100 mM KCl/40% acetonitrile present. Each spectrum was collected at 20°C, the determined T_M , and 90°C. Figure S8: ECD spectra of VEGF iM_{mod1} at the pKa exposed to 0 KCl (mM)/0% acetonitrile, 100 mM KCl by itself, 40% acetonitrile by itself, and both 100 mM KCl/40% acetonitrile present. Each spectrum was collected at 20°C, the determined T_M , and 90°C. Figure S9: ECD spectra of VEGF iM_{mod2} at the pKa exposed to 0 KCl (mM)/0% acetonitrile, 100 mM KCl by itself, 40% acetonitrile by itself, and both 100 mM KCl/40% acetonitrile present. Each spectrum was collected at 20°C, the determined T_M , and 90°C. Figure S10: ECD spectra of VEGF iM_{mod3} at the pKa exposed to 0 KCl (mM)/0% acetonitrile, 100 mM KCl by itself, 40% acetonitrile by itself, and both 100 mM KCl/40% acetonitrile present. Each spectrum was collected at 20°C, the determined T_M , and 90°C. Figure S11: ECD spectra of VEGF iM_{mod4} at the pKa exposed to 0 KCl (mM)/0% acetonitrile, 100 mM KCl by itself, 40% acetonitrile by itself, and both 100 mM KCl/40% acetonitrile present. Each spectrum was collected at 20°C, the determined T_M , and 90°C. Figure S12: 5-hmC modifications within VEGF G4-forming region allows stable G4 formation with cosolvents. Upon addition of 100 mM KCl alone and combined with 40% ACN, maximal G4 formation occurred compared to control. Also, thermal stability markedly increased with each addition for all sequences. (A) Total G-to-T mutations prevent G4 formation alone and in the presence of 100 mM KCl and 40% ACN (B). Figure S13: determination of cosolvent concentrations needed for maximum VEGF₁₋₄ formation. 0–100 mM gradient of KCl was added to wild-type VEGF G4 to determine 10 mM KCl to be the minimal concentration affecting G4 formation markedly. Similarly, 0–40% of ACN, in 10% increments, was added to wild-type resulting in 30% ACN being identified for further study. Figure S14: low concentrations of cosolvents increase thermodynamic stability of modified VEGF structures. 30% ACN increase thermal profiles of each modification as compared to their own control (10 mM KCl). (Supplementary Materials)

References

- [1] V. S. Chambers, G. Marsico, J. M. Boutell, M. Di Antonio, G. P. Smith, and S. Balasubramanian, "High-throughput sequencing of DNA G-quadruplex structures in the human genome," *Nature Biotechnology*, vol. 33, no. 8, pp. 877–881, 2015.
- [2] A. Verma, K. Halder, R. Halder et al., "Genome-wide computational and expression analyses reveal G-quadruplex DNA motifs as conserved cis-regulatory elements in human and related species," *Journal of Medicinal Chemistry*, vol. 51, no. 18, pp. 5641–5649, 2008.
- [3] R. Hänsel-Hertsch, M. Di Antonio, and S. Balasubramanian, "DNA G-quadruplexes in the human genome: Detection, functions and therapeutic potential," *Nature Reviews Molecular Cell Biology*, vol. 18, no. 5, pp. 279–284, 2017.
- [4] N. Maizels, "G4-associated human diseases," *EMBO Reports*, vol. 16, no. 8, pp. 910–922, 2015.
- [5] T. A. Brooks, S. Kendrick, and L. Hurley, "Making sense of G-quadruplex and i-motif functions in oncogene promoters," *FEBS Journal*, vol. 277, no. 17, pp. 3459–3469, 2010.
- [6] S. Benabou, A. Aviñó, R. Eritja, C. González, and R. Gargallo, "Fundamental aspects of the nucleic acid i-motif structures," *RSC Advances*, vol. 4, no. 51, pp. 26956–26980, 2014.
- [7] H. A. Day, P. Pavlou, and Z. A. E. Waller, "I-Motif DNA: Structure, stability and targeting with ligands," *Bioorganic & Medicinal Chemistry*, vol. 22, no. 16, pp. 4407–4418, 2014.
- [8] S. Ahmed, A. Kintanar, and E. Henderson, "Human telomeric c-strand tetraplexes," *Nature Structural & Molecular Biology*, vol. 1, no. 2, pp. 83–88, 1994.
- [9] G. A. Michelotti, E. F. Michelotti, A. Pullner, R. C. Duncan, D. Eick, and D. Levens, "Multiple single-stranded cis elements are associated with activated chromatin of the human c-myc gene in vivo," *Molecular and Cellular Biology*, vol. 16, no. 6, pp. 2656–2669, 1996.
- [10] A. Rangan, O. Y. Fedoroff, and L. H. Hurley, "Induction of duplex to G-quadruplex transition in the c-myc promoter region by a small molecule," *The Journal of Biological Chemistry*, vol. 276, no. 7, pp. 4640–4646, 2001.
- [11] T. Simonsson, M. Pribylova, and M. Vorlickova, "A nuclease hypersensitive element in the human c-myc promoter adopts several distinct i-tetraplex structures," *Biochemical and Biophysical Research Communications*, vol. 278, no. 1, pp. 158–166, 2000.
- [12] K. Guo, V. Gokhale, L. H. Hurley, and D. Sun, "Intramolecularly folded G-quadruplex and i-motif structures in the proximal promoter of the vascular endothelial growth factor gene," *Nucleic Acids Research*, vol. 36, no. 14, pp. 4598–4608, 2008.
- [13] M. C. Miller, R. Buscaglia, J. B. Chaires, A. N. Lane, and J. O. Trent, "Hydration is a major determinant of the G-quadruplex stability and conformation of the human telomere 3' sequence of d(AG₃(TTAG₃)),³" *Journal of the American Chemical Society*, vol. 132, no. 48, pp. 17105–17107, 2010.
- [14] D. Sun, K. Guo, and Y.-J. Shin, "Evidence of the formation of G-quadruplex structures in the promoter region of the human vascular endothelial growth factor gene," *Nucleic Acids Research*, vol. 39, no. 4, pp. 1256–1265, 2011.
- [15] M. M. Dailey, M. Clarke Miller, P. J. Bates, A. N. Lane, and J. O. Trent, "Resolution and characterization of the structural polymorphism of a single quadruplex-forming sequence," *Nucleic Acids Research*, vol. 38, no. 14, Article ID gkq166, pp. 4877–4888, 2010.
- [16] D.-H. Zhang, T. Fujimoto, S. Saxena, H.-Q. Yu, D. Miyoshi, and N. Sugimoto, "Monomorphic RNA G-quadruplex and polymorphic DNA G-quadruplex structures responding to cellular environmental factors," *Biochemistry*, vol. 49, no. 21, pp. 4554–4563, 2010.
- [17] J. Josse, A. D. Kaiser, and A. Kornberg, "Enzymatic synthesis of deoxyribonucleic acid. VIII. Frequencies of nearest neighbor base sequences in deoxyribonucleic acid," *The Journal of Biological Chemistry*, vol. 236, pp. 864–875, 1961.
- [18] J. M. Morrison, H. M. Keir, H. Subak-Sharpe, and L. V. Crawford, "Nearest neighbour base sequence analysis of the deoxyribonucleic acids of a further three mammalian viruses: Simian virus 40, human papilloma virus and adenovirus type 2," *Journal of General Virology*, vol. 1, no. 1, pp. 101–108, 1967.
- [19] E. Scarano, M. Iaccarino, P. Grippo, and E. Parisi, "The heterogeneity of thymine methyl group origin in DNA pyrimidine isostichs of developing sea urchin embryos," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 57, no. 5, pp. 1394–1400, 1967.

- [20] Y. P. Bhavsar-Jog, E. Van Dornshuld, T. A. Brooks, G. S. Tschumper, and R. M. Wadkins, "Epigenetic modification, dehydration, and molecular crowding effects on the thermodynamics of i-motif structure formation from C-Rich DNA," *Biochemistry*, vol. 53, no. 10, pp. 1586–1594, 2014.
- [21] M. Kawahara and T. Kono, "Roles of genes regulated by two paternally methylated imprinted regions on chromosomes 7 and 12 in mouse ontogeny," *The Journal of Reproduction and Development*, vol. 58, no. 2, pp. 175–179, 2012.
- [22] A. K. Ludwig, P. Zhang, and M. C. Cardoso, "Modifiers and readers of DNA modifications and their impact on genome structure, expression, and stability in disease," *Frontiers in Genetics*, vol. 7, article no. 115, 2016.
- [23] K. Jabbari and G. Bernardi, "Cytosine methylation and CpG, TpG (CpA) and TpA frequencies," *Gene*, vol. 333, pp. 143–149, 2004.
- [24] W. A. Pastor, U. J. Pape, Y. Huang et al., "Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells," *Nature*, vol. 473, no. 7347, pp. 394–397, 2011.
- [25] H. Stroud, S. Feng, S. Morey Kinney, S. Pradhan, and S. E. Jacobsen, "5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells," *Genome Biology*, vol. 12, no. 6, article R54, 2011.
- [26] R. Halder, K. Halder, P. Sharma, G. Garg, S. Sengupta, and S. Chowdhury, "Guanine quadruplex DNA structure restricts methylation of CpG dinucleotides genome-wide," *Molecular BioSystems*, vol. 6, no. 12, pp. 2439–2447, 2010.
- [27] J. Lin, J.-Q. Hou, H.-D. Xiang et al., "Stabilization of G-quadruplex DNA by C-5-methyl-cytosine in bcl-2 promoter: implications for epigenetic regulation," *Biochemical and Biophysical Research Communications*, vol. 433, no. 4, pp. 368–373, 2013.
- [28] D. R. Senger, S. J. Galli, A. M. Dvorak, C. A. Perruzzi, V. S. Harvey, and H. F. Dvorak, "Tumor cells secrete a vascular permeability factor that promotes accumulation of ascites fluid," *Science*, vol. 219, no. 4587, pp. 983–985, 1983.
- [29] D. Sun, "In vitro footprinting of promoter regions within supercoiled plasmid DNA," in *Drug-DNA Interaction Protocols*, vol. 613 of *Methods in Molecular Biology*, pp. 223–233, Humana Press, Totowa, NJ, USA, 2010.
- [30] D. Sun, K. Guo, J. J. Rusche, and L. H. Hurley, "Facilitation of a structural transition in the polypurine/polypyrimidine tract within the proximal promoter region of the human VEGF gene by the presence of potassium and G-quadruplex-interactive agents," *Nucleic Acids Research*, vol. 33, no. 18, pp. 6070–6080, 2005.
- [31] D. Sun, W.-J. Liu, K. Guo et al., "The proximal promoter region of the human vascular endothelial growth factor gene has a G-quadruplex structure that can be targeted by G-quadruplex-interactive agents," *Molecular Cancer Therapeutics*, vol. 7, no. 4, pp. 880–889, 2008.
- [32] A. M. Fleming, J. Zhou, S. S. Wallace, and C. J. Burrows, "A role for the fifth G-track in G-quadruplex forming oncogene promoter sequences during oxidative stress: Do these "spare tires" have an evolved function?" *ACS Central Science*, vol. 1, no. 5, pp. 226–233, 2015.
- [33] R. Buscaglia, M. C. Miller, W. L. Dean et al., "Polyethylene glycol binding alters human telomere G-quadruplex structure by conformational selection," *Nucleic Acids Research*, vol. 41, no. 16, pp. 7934–7946, 2013.
- [34] S. Dhakal, Y. Cui, D. Koirala et al., "Structural and mechanical properties of individual human telomeric G-quadruplexes in molecularly crowded solutions," *Nucleic Acids Research*, vol. 41, no. 6, pp. 3915–3923, 2013.
- [35] S. M. Reilly, D. F. Lyons, S. E. Wingate et al., "Folding and hydrodynamics of a DNA i-motif from the c-MYC promoter determined by fluorescent cytidine analogs," *Biophysical Journal*, vol. 107, no. 7, pp. 1703–1711, 2014.
- [36] E. P. Wright, J. L. Huppert, and Z. A. E. Waller, "Identification of multiple genomic DNA sequences which form i-motif structures at neutral pH," *Nucleic Acids Research*, vol. 45, no. 6, pp. 2951–2959, 2016.
- [37] R. D. Gray, R. Buscaglia, and J. B. Chaires, "Populated intermediates in the thermal unfolding of the human telomeric quadruplex," *Journal of the American Chemical Society*, vol. 134, no. 40, pp. 16834–16844, 2012.
- [38] C. C. Hardin, M. Corregan, B. A. Brown, and L. N. Frederick, "Cytosine-cytosine+ base pairing stabilizes dna quadruplexes and cytosine methylation greatly enhances the effect," *Biochemistry*, vol. 32, no. 22, pp. 5870–5880, 1993.
- [39] P. Agrawal, E. Hatzakis, K. Guo, M. Carver, and D. Yang, "Solution structure of the major G-quadruplex formed in the human VEGF promoter in K+: Insights into loop interactions of the parallel G-quadruplexes," *Nucleic Acids Research*, vol. 41, no. 22, pp. 10584–10592, 2013.
- [40] W. Yoshida, H. Yoshioka, D. H. Bay et al., "Detection of DNA methylation of G-quadruplex and i-motif-forming sequences by measuring the initial elongation efficiency of polymerase chain reaction," *Analytical Chemistry*, vol. 88, no. 14, pp. 7101–7107, 2016.
- [41] D. P. Sundrani, U. S. Reddy, P. M. Chavan-Gautam, S. S. Mehendale, G. R. Chandak, and S. R. Joshi, "Altered methylation and expression patterns of genes regulating placental angiogenesis in preterm pregnancy," *Reproductive Sciences*, vol. 21, no. 12, pp. 1508–1517, 2014.
- [42] Y. Tang, S. Han, T. Asakawa et al., "Effects of intracerebral hemorrhage on 5-hydroxymethylcytosine modification in mouse brains," *Neuropsychiatric Disease and Treatment*, vol. 12, pp. 617–624, 2016.
- [43] N. Ferrara and T. Davis-Smyth, "The biology of vascular endothelial growth factor," *Endocrine Reviews*, vol. 18, no. 1, pp. 4–25, 1997.
- [44] J. Y. Kim, J. H. Hwang, W. Zhou et al., "The expression of VEGF receptor genes is concurrently influenced by epigenetic gene silencing of the genes and VEGF activation," *Epigenetics: Official Journal of the DNA Methylation Society*, vol. 4, no. 5, pp. 313–321, 2009.
- [45] M. P. Turunen and S. Ylä-Herttuala, "Epigenetic regulation of key vascular genes and growth factors," *Cardiovascular Research*, vol. 90, no. 3, pp. 441–446, 2011.
- [46] B. G. Kim, J. Long, D. N. Dubins, and T. V. Chalikian, "Ionic effects on VEGF G-quadruplex stability," *The Journal of Physical Chemistry B*, vol. 120, no. 22, pp. 4963–4971, 2016.
- [47] M. Marušič, R. N. Veedu, J. Wengel, and J. Plavec, "G-rich VEGF aptamer with locked and unlocked nucleic acid modifications exhibits a unique G-quadruplex fold," *Nucleic Acids Research*, vol. 41, no. 20, pp. 9524–9536, 2013.
- [48] M. M. Molnar, R. K. Morgan, T. A. Brooks, and R. M. Wadkins, "Effects of 5-hydroxymethylcytosine epigenetic modifications within the VEGF promoter region on G-quadruplex and I-motif DNA structure and stability," *Biophysical Journal*, vol. 112, no. 3, 2017, <http://thesis.honors.olemiss.edu/id/eprint/511>.