BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

G-Quadruplex-Enabling Sequence within the Human Tyrosine Hydroxylase Promoter Differentially Regulates Transcription

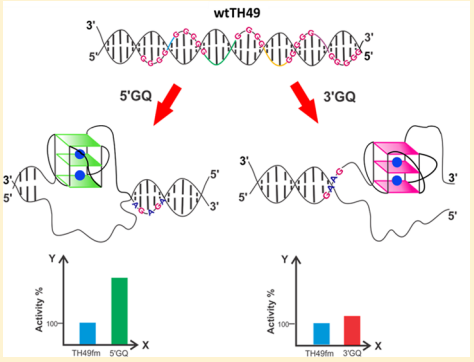
Mohamed M. Farhath,[†] Matthew Thompson,[†,§] Sujay Ray,[‡] Abby Sewell,[†] Hamza Balci,[‡] and Soumitra Basu*[,†]

[†]Department of Chemistry and Biochemistry, Kent State University, Kent, Ohio 44242, United States

[‡]Department of Physics, Kent State University, Kent, Ohio 44242, United States

Supporting Information

ABSTRACT: G-Quadruplexes (GQs) found within the promoter regions of genes are known to mostly act as repressors of transcription. Here we report a guanosine (G)-rich segment in the 3′-proximal promoter region of human tyrosine hydroxylase (*TH*), which acts as a necessary element for transcription. Tyrosine hydroxylase catalyzes the rate-limiting step in the catecholamine biosynthesis and is linked to several common neurological disorders such as Parkinson's and schizophrenia. A 45 nucleotide (nt) sequence (wtTH49) within the human *TH* promoter contains multiple G-stretches that are extremely well conserved among the primates but deviate in rodents, which raises the possibility of variation in the GQ structures formed in the two orders with the potential for a distinctive functional outcome. Biochemical and biophysical studies, including single-molecule Förster resonance energy transfer, indicate that the wtTH49 sequence can adopt multiple GQ structures by using different combinations of G-stretches. A functional assay performed with 2.8 kb of the 3′-proximal end of the *TH* promoter and a mutated version (TH49fm; mutated wtTH49) that is unable to form any GQ structure indicates that overall the GQ-enabling wtTH49 sequence is functionally necessary and enhances human *TH* promoter activity by 5-fold compared to that of the mutant. Two additional mutants, each of which was designed to form distinct GQs, differentially affected reporter gene transcription. A cationic porphyrin TMPyP4 destabilizes the wtTH49 GQ and lowers the level of reporter gene expression, although its analogue, TMPyP2, fails to elicit any response. The 45 nt G-rich sequence within the human *TH* promoter can form multiple GQ structures, is a necessary element in transcription, and depending on the utilized combination of G-stretches affects transcription in different ways.



G-Quadruplex structures present within the promoter of genes have been shown to mostly repress transcription, and a majority of the guanine-rich (G-rich) sequences reported in the literature contain four G-stretches indicating formation of one GQ or a very small set of GQs.[1−3] The major exceptions are bcl-2 and HIV,[4,5] which can utilize various combinations of the multiple G-stretches present within the G-rich segment to potentially adopt many GQs. Human tyrosine hydroxylase (*TH*) gene promoter contains a 45 nucleotide (nt) segment encompassing seven G-stretches that logically can support many GQ formations, and thus, it is important not only to deconvolute the functional consequence of such a complex sequence to understand their regulatory role vis-à-vis simpler sequences but also to delineate its effect on regulation of human *TH* transcription. Tyrosine hydroxylase is the enzyme that catalyzes the rate-limiting step in the catecholamine biosynthesis.[6,7] It is present in all types of catecholaminergic cells and is generally used as a marker for catecholaminergic neurons.[8−10] Dysregulation of *TH* expression has been linked to various neurological disorders, including bipolar disorder[11] and schizophrenia.[12,13] The loss of dopaminergic neurons that decreases the amount of *TH* in the substantia nigra is the primary reason for the motor symptoms in Parkinson's disease.[14] Therefore, a detailed analysis of human *TH* gene expression regulation will be extremely helpful in understanding these diseases and developing future strategies for treating them. To investigate the factors that regulate *TH* gene expression, a broad array of studies have been conducted in human[15,16] and rodent[17−19] models.

The regulation of *TH* transcription is a complex process and requires interaction of regulatory DNA elements with DNA-binding proteins.[10,20,21] The human *TH* promoter was determined to be extremely GC-rich,[10,11] and guanine-rich (G-rich) regions are known to adopt G-quadruplex (GQ) structures.[22−24] Such structures have been identified in promoter regions of many important growth-related genes and proto-oncogenes.[2−4,25−30] Almost all previous studies of promoter GQs showed that they repress transcription. We have identified a 45 nt GQ-forming sequence (wtTH49) within a 272 nt 3′-proximal region of the *TH* promoter that is known to

Received: February 27, 2015
Revised: July 24, 2015
Published: August 18, 2015

DOI: 10.1021/acs.biochem.5b00209
Biochemistry 2015, 54, 5533−5545

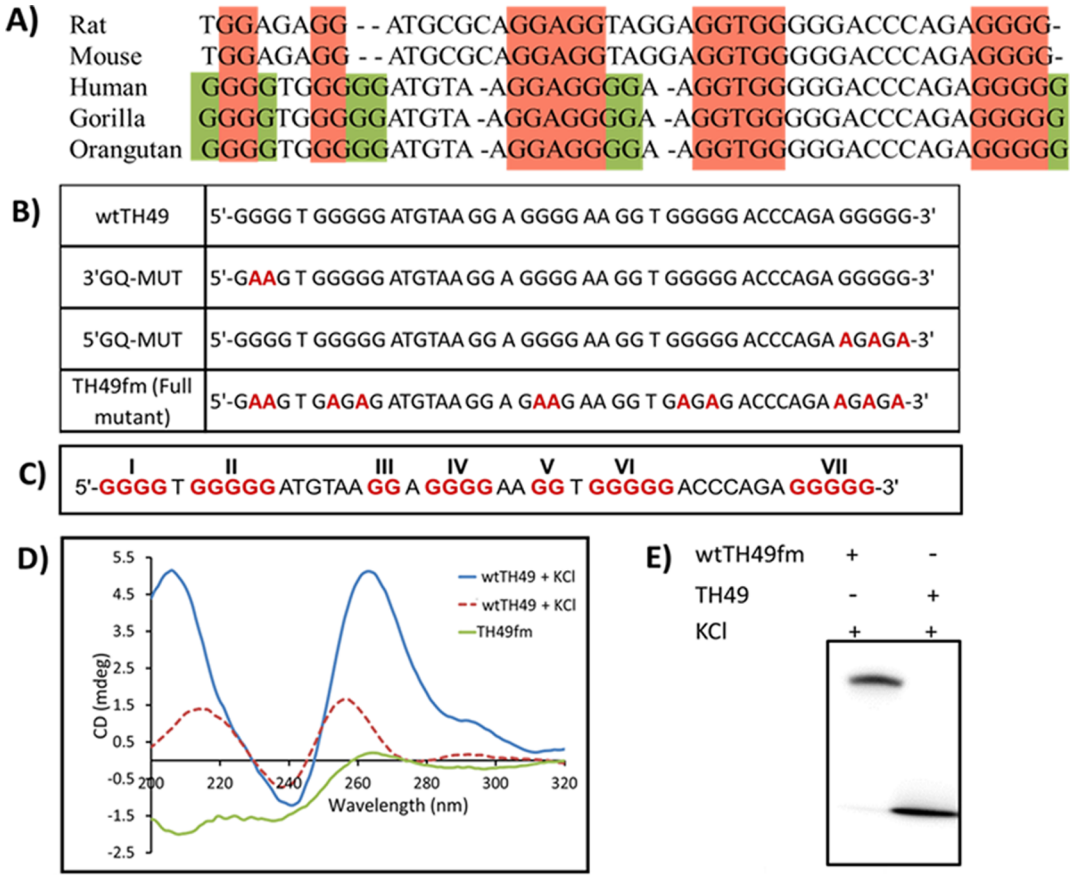


Figure 1. G-Rich wtTH49 sequence with seven G-stretches adopts a parallel G-quadruplex structure. (A) G:C-rich conserved region 1 (CR1) in the *TH*-proximal promoters that is located upstream of the well-characterized CRE and TATA box sequence. Orange shaded regions are highly conserved short nucleotide motifs among mammals located within the CRI and G residues that are not conserved in rodents but among hominoids are highlighted in green.[34] (B) Sequences of the wild-type G-rich element (wtTH49) and its mutants 5′GQ-MUT, 3′GQ-MUT, and TH49fm. (C) The seven G-stretches of wtTH49 are colored red, and their numbers are denoted with Roman numerals. (D) CD spectra of the wild type (wtTH49) and the full mutant (TH49fm) in the presence and absence of 100 mM KCl. Spectra were recorded in 10 mM Tris (pH 7.5) and 100 mM KCl at an oligonucleotide concentration of 5 μM. (E) Electrophoretic mobilities of folded wtTH49 and TH49fm single-stranded DNA in the presence of KCl in running buffer and nondenaturing gel.

have a responsive element for cyclic AMP.[31] wtTH49 is located −49 to −93 nt upstream of the transcription start site (TSS) and has robust sequence homology among human, gorilla, orangutan, rat, and mouse forms (Figure 1A).[15,19,21,32] However, there are small but significant differences between the common rodents and some of the primate sequences (Figure 1A), particularly in the composition of segments with contiguous guanosine stretches (G-stretches).[32] This variance may profoundly affect the identity and stability of the GQs formed in the primates and rodents, which can cause potential differences in the mechanism of regulation in the two orders. For example, the rodent sequences can only support two-tier GQs while the primate sequences can adopt three-tier or four-tier GQs, and it is well established that three- or four-tier GQs are significantly more stable than the two-tier ones.[33,34] Additionally, simply because of the composition of the primary structure, the primate sequence potentially can adopt a more complex set of GQ structures.

The wtTH49 sequence is in the proximity of several important transcription factor-binding sites, including cyclic AMP response element (CRE), and also the TATA box, suggesting that it might have a regulatory role.[19,21,32] The sequence contains seven G-stretches, each having two to five consecutive guanosines (Gs) separated by one to seven nucleotides, which can potentially form the loops of the GQ. The presence of seven G-stretches makes the wtTH49 sequence very heterogeneous as it allows the sequence to adopt more than one GQ by utilizing different combinations of G-stretches. Therefore, the regulatory roles played by these GQs may be more complex compared to those of most of the GQ-forming promoter sequences characterized thus far. We determined that the sequence adopts multiple GQs that differentially regulate the transcription of a reporter gene, and a subset of G-rich sections is necessary for *TH* expression.

MATERIALS AND METHODS

Preparation of Oligonucleotide Sequences. DNA oligonucleotides were purchased from Integrated DNA Technologies and purified using 17% denaturing polyacrylamide gel electrophoresis (PAGE). DNA oligonucleotides were harvested via a crush and soak method by tumbling the gel slices 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 ethanol precipitated with 3 volumes of ice-cold 100% ethanol. The salt was removed by washing the DNA pellets with 70% ethanol.

The 5′-end-radiolabeled single-stranded oligonucleotides were prepared by treating the DNA with T4 polynucleotide

DOI: 10.1021/acs.biochem.5b00209
Biochemistry 2015, 54, 5533−5545

kinase (NEB) and [γ-^{32}P]ATP (PerkinElmer) and incubated for 45 min at 37 °C. The radiolabeled DNA oligonucleotides were purified by 17% denaturing PAGE and extracted from the gel via the crush and soak method as described above.[35]

Circular Dichroism (CD) Studies. All DNA samples were prepared in a buffer containing 100 mM KCl and 10 mM Tris-HCl (pH 7.5) at a final DNA concentration of 5 μM. For the no KCl control, DNA samples were prepared in 10 mM Tris-HCl (pH 7.5) buffer. DNA structures were folded by heating the samples at 95 °C for 10 min followed by slow cooling to room temperature over a 90 min period. A Jasco J-810 spectrophotometer and quartz cell with a 0.1 cm path length were used to record CD spectra at a scan speed of 50 nm/min, with a response time of 1 s over a wavelength range of 200−320 nm. All spectra shown in this work were averaged over three scans taken at room temperature. All spectra were subjected to baseline corrections to eliminate any signal contribution from the buffer. After the collection of the initial spectrum, the samples were titrated with an increasing amount of the cationic porphyrins TMPyP4 and TMPyP2. In each step, the reaction solution was mixed well after addition of the cationic porphyrins and incubated for 5 min at room temperature before a new spectrum was recorded. Dilution correction was performed for each concentration of TMPyP4 and TMPyP2.[35]

Dimethyl Sulfate (DMS) Structure Mapping. Samples for DMS structure mapping were prepared with 1 μM 75 nt unlabeled DNA (sequences listed in Supplementary Table 1), 10 mM Tris-HCl buffer (pH 7.4), 100000 cpm 5′-end-radiolabeled oligonucleotides, and 100 mM KCl or no KCl in a final volume of 30 μL. DNA structures were folded as described above. Samples were then treated with 1% DMS for 2 min at room temperature before the methylation reactions were stopped by adding 300 μL of stop buffer (300 mM sodium acetate, 250 mg/mL sheared salmon sperm DNA, and 2 M mercaptoethanol). DNA samples were ethanol precipitated with 3 volumes of ice-cold 100% ethanol, and the DNA pellets were washed with 70% ethanol. The pellets were then dried in a vacuum centrifuge and then treated with 70 μL of freshly prepared 10% piperidine for 30 min at 95 °C. The cleaved products were resolved on a 12% denaturing polyacrylamide gel, and the dried gel was exposed to a phosphorimager screen and visualized on a Typhoon FLA 9500 Phosphorimager (GE Life Sciences).

Competition between GQ and Double-Stranded DNA. Three samples were prepared for the competition assay. For the first sample, GQ was folded without a complementary sequence (C). In the second sample, the 75 nt complementary sequence (C) was added (1:1 ratio) to the reaction mixture before GQ was folded (sample wtTH49&C) as described above. In the case of the third sample, the complementary sequence (C) was added to the reaction mixture after formation of the GQ (sample TH49+C). GQs were folded, and DMS structure mapping was performed as described above.

Electrophoretic Mobility Shift Assay (EMSA). Wild-type and mutant samples for the EMSA were prepared using 5′-end-radiolabeled oligonucleotides in a buffer containing 100 mM KCl and 10 mM Tris-HCl buffer (pH 7.4). The mobility of the folded structures was assayed on a 10% native polyacrylamide gel in the presence or absence of 100 mM KCl in both 1× TBE buffer and a native gel.

DNA Polymerase Stop Assay. To anneal the P20 stop primer with the template DNA, an equal volume of 5′-end-radiolabeled P20 stop primer (24 nM) was mixed with *TH* stop template DNA (24 nM) in a Tris-HCl buffer (10 mM Tris, pH 8.0) containing various concentration of K^+ ions, and the reaction mixture was heated at 95 °C for 10 min and cooled slowly to room temperature. Primer extension reactions were initiated by mixing dNTP (100 μM), DNA polymerase buffer (1×, no K^+), and *Taq* DNA polymerase (2 units/reaction; New England Biolab) in a reaction mixture volume of 20 μL and incubated at 37 °C for different time periods. The reactions were stopped by adding an equal volume of stop buffer (7 M urea, 10 mM EDTA, 10 mM NaOH, 0.025% xylene cyanol, and 0.025% bromophenol blue). The reaction products were resolved in a 12% denaturing polyacrylamide gel:[36] *TH* stop template, 5′-ACAGCAGGCAGGGGTGGGGGATGTAAGG-AGGGGAAGGTGGGGGACCCAGAGGGGGCTTTGACG-TCAGCTCAGCTTATAAGAGGCTG-3′; P20 stop primer, 5′-CAGCCTCTTATAAGCTGAGC-3′.

Single-Molecule Förster Resonance Energy Transfer (smFRET) Assay. Cleaned quartz slides and glass coverslips, coated with a mixture of 99% *m*-polyethylene glycol and 1% biotinylated polyethylene glycol (m-PEG-SVA-5000 and biotin-PEG-SVA-5000, respectively; Laysan Bio, Arab, AL), were used to prepare the sample chamber. This chamber was then treated with 0.05 mg/mL neutravidin. Biotinylated partial duplex DNA (pdDNA) constructs at a concentration of 15 pM were then incubated in the chamber, which were immobilized on the surface via neutravidin−biotin linkers. For steady state measurements, the DNA constructs were incubated in an imaging buffer [50 mM Tris (pH 7.5), 0.8 mg/mL glucose, 0.1 mg/mL bovine serum albumin, 140 mM β-mercaptoethanol, 0.1 mg/mL glucose oxidase, 0.02 mg/mL catalase, a saturating concentration of Trolox (∼2 mM), and 100 mM KCl] for 15 min before data were acquired.

A prism-type total internal reflection fluorescence setup, which was built around an Olympus IX-71 microscope and equipped with a 60× 1.20 N.A. water objective, was used for the smFRET measurements. Data were collected at an integration time of 100 ms using an Andor Ixon electron-multiplying CCD camera (iXon DV 887-BI, Andor Technology, South Windsor, CT). smFRET time traces for individual molecules were analyzed to generate FRET histograms. In this analysis, smFRET time traces are individually inspected for anticorrelation between donor−acceptor fluorophores and single-step photobleaching, which are required for their qualification as a single molecule. The background subtraction was performed in three steps. In the first step, the minimal intensity in an image is subtracted from the intensity of each pixel. In the second step, the image of the donor and acceptor signals is divided into multiple subsections each, to account for variations upon illumination of different parts of the image. Each of these subsections has ∼10−20 molecules in it. An average background signal is subtracted from each subsection based on the median intensity, which should be a representative of the remaining background because the molecule density is kept low enough that most pixels do not have a fluorescent molecule. The residual background for each molecule was then determined after the donor molecule is photobleached, at which point all the remaining intensity should by definition be the background. This subtraction is performed independently for the donor and acceptor signals to take into account the variations between the two. More than 180 molecules that satisfy these criteria and undergo these background subtraction processes are used for each FRET histogram.

DOI: 10.1021/acs.biochem.5b00209
Biochemistry 2015, 54, 5533−5545

In addition to these molecules, adequate numbers of molecules that have a donor fluorophore but not an acceptor fluorophore are included in the initial histogram. These molecules form the donor-only (DO) peak, which is a measure of the leakage from the donor emission into the acceptor channel. The E_{FRET} value corresponding to this peak (E_{DO_FRET}) is determined by fitting the histogram with multiple Gaussian peaks. The Gaussian peak with the lowest E_{FRET} is identified as the DO peak. This peak is then subtracted from the histogram. The remaining histogram, which represents the molecules that have both donor and acceptor fluorophores, is rescaled such that the E_{DO_FRET} value is moved to $E'_{FRET} = 0$ by the following scaling:

$$E'_{FRET} = \frac{E_{FRET} - E_{DO\_FRET}}{1 - E_{DO\_FRET}}$$

where E_{FRET} is the FRET efficiency before DO correction is made and E'_{FRET} is the FRET efficiency after DO correction. An example of DO correction and scaling of the smFRET is shown in Supplementary Figure S3 of ref 37. While the histograms were being constructed, the contribution of each molecule is scaled such that they equally contribute to the histogram to avoid having a greater contribution from molecules that fluoresce for a longer period of time compared to those that photobleach quicker. At the final step, the histogram is normalized to percent scale, i.e., such that the total area under the histogram equals 100%. Finally, the rescaled and normalized histogram is fit by multiple Gaussian peaks again to determine the peak positions in the corrected (E'_{FRET}) scale.

The following sequences were used for the smFRET assays: TH49-Long Strand, 5′-Cy3-GGG GTG GGG GAT GTA AGG AGG GGA AGG TGG GGG ACC CAG AGG GGG TTTTTG GCG ACG GCA GCG AGG C-3′; pdT50-Long Strand, 5′-Cy3-poly(dT)$_{50}$ GGC GAC GGC AGC GAG GC-3′; Short Strand, 5′-Biotin-GCC TCG CTG CCG TCG CCA Cy5-3′.

Preparation of Plasmids for Functional Analysis of the *TH* Promoter. The reporter constructs were generated using the pGL3-Basic plasmid reporter vector (Promega) for analysis of the active promoter region of the human *TH* gene. A *Kpn*I−*Bgl*II fragment (nucleotides −2804 to +15), containing 2819 bp of the *TH* upstream promoter region, was cloned into the pGL3-basic vector (Promega), directly upstream of the firefly luciferase gene, to generate the 2.8TH-pGL3 plasmid. The sequence of the construct was verified by The Ohio State University's Plant Microbe Genomic sequencing facility. A plasmid expressing Renilla (pRL-TK) was used as a transfection control.

Cell Culture. C6 cells were maintained in Dulbecco's modified Eagle's medium (Corning) supplemented with 10% fetal bovine serum and the 1% antibiotics streptomycin, penicillin, and amphotericin B at 37 °C in 5% CO_2 in a humidified incubator. SHSY5Y cells were maintained in a 1:1 mixture of Ham's F12 medium and Eagle's Minimum Essential Medium (EMEM) medium (Corning) supplemented with 10% FBS and 1% antibiotics.

Luciferase Assay. Transfection to C6 cells was performed with 188 ng of firefly luciferase construct 2.8TH-pGL3 and 63 ng of Renilla reporter plasmid (pRL-TK) in 96-well plates using Lipofectamine LTX and PLUS reagent, based on the manufacturer's protocol. Briefly, the firefly luciferase construct 2.8TH-pGL3 and Renilla reporter plasmid (pRL-TK) were mixed together in a 3:1 ratio in a total of 250 ng of plasmid and incubated with Lipofectamine LTX in a 1:4 ratio for 30 min. After transfection for 6 h, 200 μL of medium containing 10% fetal bovine serum and 1% antibiotics were added to the cells. At 48 h post-transfection, Renilla and firefly luciferase activities were measured using the Dual-Glo Luciferase assay system (Promega) by following the manufacturer's protocol on a Synergy 2 microplate reader (Bio-Tek Instruments). In the case of the small molecule treatment, plates were supplemented with a final concentration of 0, 10, 30, or 50 μM TMPyP4 or TMPyP2 following the 6 h transfection, and luciferase expression was measured after 48 h.

Both firefly luciferase construct 2.8TH-pGL3 and Renilla reporter plasmid (pRL-TK) were transfected to SHSY5Y cells using the *Trans*IT-X2 dynamic delivery system based on the manufacturer's protocol. Briefly, the firefly luciferase construct 2.8TH-pGL3 and Renilla reporter plasmid (pRL-TK) were mixed together in 3:1 ratio in a total of 250 ng of plasmid and incubated with *Trans*IT-X2 in a 1:2 ratio for 20 min. After transfection for 6 h, 200 μL of medium containing 10% fetal bovine serum and 1% antibiotics were added to the cells. At 48 h post-transfection, Renilla and firefly luciferase activities were measured using the Dual-Glo Luciferase assay system (Promega).

Quantitative Reverse Transcriptase Polymerase Chain Reaction (RT-PCR) Assays. Total RNA from transfected C6 cells was extracted using the E.Z.N.A MicroElute Total RNA kit (OMEGA Bio-Tek). Endogenous DNAs were eliminated from each RNA sample upon treatment with RQ1 RNase-Free DNase (Promega) prior to reverse transcription. Finally, 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 (Quanta Biosystems) on an Eppendorf Mastercycler RealPlex2 Sequence Detection System in the presence of the appropriate set of primers. Primers for Firefly m-RNA: FL-FP, 5′-TTCGCTAAGAGCACCCTGAT-3′; FL-RP, 5′-GCTGCA-GCAGGATAGACTCC-3′. Primers for Renilla m-RNA: RL-FP, 5′-GTAACGCTGCCTCCAGCTAC-3′; RL-RP, 5′-GTA-GGCAGCGAACTCCTCAG-3′.

RESULTS

A G-Rich 45 nt Segment within the Human Tyrosine Hydroxylase Promoter Adopts Multiple G-Quadruplex Structures. wtTH49 with seven G-stretches has strong potential to form GQ structure in the presence of K^+ ions. To search for such GQ formation, we first investigated the wtTH49 sequence using circular dichroism (CD) ellipticity at a physiologically relevant K^+ concentration (100 mM KCl).[38] CD measurements performed on wtTH49 and its mutant (TH49fm) are shown in Figure 1D. TH49fm is created by mutating five of the seven G-stretches where guanines are replaced with adenines, eliminating the possibility of GQ formation, while keeping the purine nature of the bases intact. The CD spectrum for wtTH49 showed a characteristic positive ellipticity at ∼263 nm and negative ellipticity at ∼240 nm in the presence of K^+, implying that it adopts parallel GQ structure (Figure 1D).[39−41] However, CD spectra of both wtTH49 in the absence of K^+ and TH49fm do not show any characteristic features of a GQ. This observation strongly suggests K^+-dependent GQ structure formation by the wtTH49 sequence.

To determine if the GQs adopted by wtTH49 are intra- or intermolecular, we analyzed the electrophoretic mobility of

DOI: 10.1021/acs.biochem.5b00209
Biochemistry 2015, 54, 5533−5545

wtTH49 under nondenaturing conditions in the absence (Supplementary Figure 1) and presence (Figure 1E) of KCl (100 mM), in the gel and running buffer. We observed that in the presence of 100 mM K^+, the secondary structure formed by wtTH49 migrated in a 10% native polyacrylamide gel faster than the corresponding unstructured single-stranded TH49fm, and no slower-migrating bands were detected (Figure 1E). As expected, there was no significant difference in mobility in the absence of K^+ in gel as well as running buffer (Supplementary Figure 1). Given that intramolecular GQ is more compact than unstructured DNA while intermolecular GQ is more extended, this observation suggests that wtTH49 most likely adopts an intramolecular GQ structure.[42]

These bulk studies were complemented by smFRET measurements that were used to identify possible formation of multiple GQs. For these studies, the DNA construct is immobilized on a polyethylene glycol-coated quartz surface via biotin−neutravidin binding. After incubation of the DNA with K^+ for ∼15 min, a steady state is reached and a FRET histogram is constructed. In the case of wtTH49 at 100 mM K^+, the FRET histograms corresponding to the folded GQ structures are several times broader than would be expected when a single GQ conformation is present (Figure 2A).[43] The

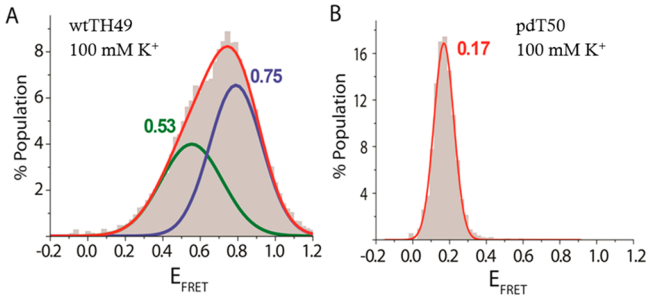


Figure 2. wtTH49 sequence that forms multiple intramolecular GQs. (A) FRET efficiency histogram for the wtTH49 construct at 100 mM K^+ and pH 7.5. A broad histogram suggests formation of multiple GQ structures. (B) FRET efficiency histogram for the pdT50 construct, which is an unstructured DNA of a length similar to that of wtTH49. A single peak at 0.17 determines the position of the unstructured random coiled DNA under the same ionic and pH conditions.

FRET histogram can be fitted by two Gaussian peaks (R^2 = 0.99), which should be considered a minimal number of GQ states given the width of the peak and the dynamics observed at the single-molecule traces. The higher peak is at FRET efficiency E_{FRET} = 0.75, and the lower peak is at E_{FRET} = 0.53. The almost perfect R^2, the coefficient of determination, for two Gaussian fit makes it difficult to justify fitting more than two Gaussian peaks. However, the dynamics of the system and the transitions observed between different FRET states (Supplementary Figure 2C) suggest that more than two FRET states exist in this system. Along these lines, we have demonstrated that five different segments of wtTH49 are capable of forming stable GQ folds under identical assay conditions (data not shown), which further support a dynamic picture in which transitions and interconversions between different GQ states are possible.

To distinguish the lower FRET peak from a random coiled DNA, control measurements were performed on another partial duplex DNA construct that has the same duplex stem and a 50 nt polythymine overhang (pdT50). The pdT50 sequence has a similar separation between the fluorophores as the wtTH49 construct (50 nt vs 49 nt overhang) but does not form any secondary structure. In the case of pdT50, a single peak at E_{FRET} = 0.17 was observed under identical experimental conditions (Figure 2B), which represents the FRET peak of an unstructured random coil DNA. Having a similar length, the wtTH49 sequence is also expected to have a similar peak when it is unstructured, in the form of a random coil. The E_{FRET} = 0.17 peak is significantly different from the two peaks observed for the wtTH49 construct under the same ionic and pH conditions. These measurements suggest that the peaks observed in the FRET histogram of the wtTH49 construct correspond to compact DNA secondary structures. However, significant folded GQ conformation was detected even at ∼0 mM K^+ in the chamber (Supplementary Figure 2A). This is due to the very stable nature of some of the GQ states formed by this construct that might be trapped in the folded state, possibly because of trace amounts of monovalent ions in the buffers. Nevertheless, there is a clear shift in the population toward more compact structures upon addition of K^+. To demonstrate this shift, we also performed FRET measurements at 50 mM K^+ (Supplementary Figure 2B), which demonstrates an intermediate between the 0 and 100 mM K^+ distributions. In combination with the CD and electrophoresis data, this assay confirms that the wtTH49 sequence forms multiple intramolecular GQs.

To further test if wtTH49 forms multiple GQs, we performed a DNA polymerase stop assay. G-Quadruplexes arrest primer extension by DNA polymerase in a K^+-dependent manner.[41,44] We rationalized that more than one stop in the assay would potentially provide evidence of the presence of multiple secondary structures. The template DNA-containing wtTH49 sequence was annealed with a radiolabeled primer (Figure 3A) and subsequently extended with *Taq* DNA polymerase in the presence of increasing concentrations of K^+. Figure 3B shows the resulting primer-extended product from the DNA polymerase stop assay at various concentrations of K^+ ions at 37 °C. The full length extension product was synthesized in the absence of K^+ (Figure 3B, lane 3) and at low K^+ concentrations with longer incubation times (Figure 3B, lanes 6 and 9). However, upon addition of K^+ to the reaction mixture, two distinctly premature primer arrest sites were detected (Figure 3B, lanes 4−12). The DNA polymerase arrest sites mainly occurred after extension of the primer by 11 and 23 nt (Figure 3A,B), which are just before G45 and G33 (see the numbering in Figure 3A) in the template strand, respectively. This result provides further evidence that the wtTH49 sequence has the capacity to form at least two separate intramolecular structures under physiologically relevant temperature and salt conditions.

To directly identify the G residues involved in the formation of the GQ adopted by the wtTH49 sequence, we probed the accessibility of guanine's N7 by dimethyl sulfate (DMS) modification. In duplex and single-stranded DNA, the N7 position of guanine is available for methylation with DMS but not in GQs in which it is involved in Hoogsteen base pairing. The methylation pattern produced by wtTH49 with and without K^+ is shown in Figure 3C. It is important to note that in the presence of 100 mM K^+, almost all G-stretches within the wtTH49 sequence were either partially or completely protected from DMS methylation. Because intramolecular GQ formation requires only four G-stretches,[45] this observation can be argued as evidence that wtTH49 can adopt more than one GQ utilizing different combinations of its G-stretches.

DOI: 10.1021/acs.biochem.5b00209
Biochemistry 2015, 54, 5533−5545

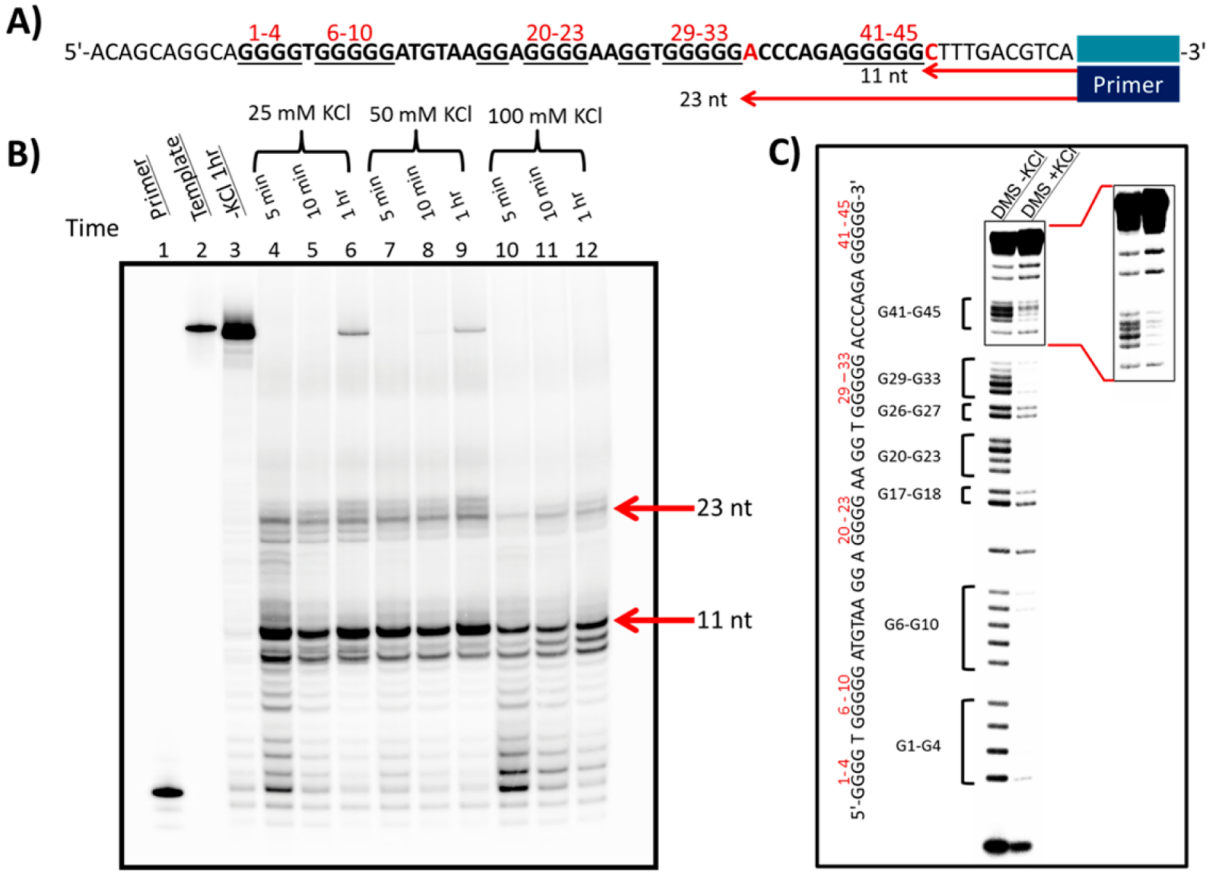


Figure 3. wtTH49 sequence forms multiple intramolecular GQs utilizing different combinations of G-stretches. (A) DNA polymerase stop assay to investigate the ability of wtTH49 to form multiple GQ structures. The data show two major K^+-dependent DNA polymerase arrest points that are designated by a red nucleotide in the sequence in panel A and with red arrows in the gel data in panel B. Lane designations are as follows: lane 1, radiolabeled free primer; lane 2, full length radiolabeled template; lane 3, *Taq* DNA polymerase extension without KCl (a positive control); lanes 4–6, with 25 mM KCl; lanes 7–9, with 50 mM KCl; lanes 10–12, with 100 mM KCl for different reaction times. (C) DMS structural mapping of wtTH49 in the absence (lane 1) and presence (lane 2) of 100 mM KCl.

To test wtTH49's ability to adopt GQ structures in the presence of the complementary strand, a competition assay was performed. A 75 nt sequence from the *TH* promoter encompassing wtTH49 (G-rich strand) and its complement sequence (C-rich strand) were used in this assay. Two different conditions were studied; in one case, DMS structure mapping was performed on the radiolabeled G-rich strand when its complementary sequence was introduced before GQ formation was triggered by addition of 100 mM K^+. In the second case, DMS treatment was performed when the C-rich complementary strand was added after GQ formation had been triggered. We observed that the methylation protection pattern of wtTH49 embedded within the 75 nt sequence was identical in both cases (Figure 4A), implying that the wtTH49 is capable of forming GQ structure even in the presence of its complementary sequence.

Results from smFRET, DMS structure mapping, and the polymerase stop assay provide strong evidence that wtTH49 forms a mixture of GQ structures in the presence of K^+. Additionally, the competition assay indicates that the GQ structures can form even when there is a direct competition from the complementary C-rich strand, which alludes to the cellular relevance of the GQ when the structures have to form in the context of duplex DNA.

G-Quadruplex Structures Formed by the wtTH49 Sequence Enhance the Human Tyrosine Hydroxylase Promoter Activity at the Transcriptional Level. The *in vitro* studies presented have established that the conserved and highly G-rich wtTH49 sequence located upstream of the *TH* promoter TSS is capable of forming multiple intramolecular GQ structures, which are parallel in conformation based upon CD data. However, the key question is whether the wtTH49 sequence is critical to regulation of transcriptional activity of the *TH* promoter. To address this question, a luciferase reporter plasmid was prepared by cloning an active 2.8 kb 3′-end fragment of the human *TH* promoter at a location upstream of the firefly luciferase gene in the promoterless pGL3 basic vector. A luciferase activity assay was performed using the cloned *TH* promoter and an empty pGL3 vector. The activity of the *TH* promoter was much higher than that of the empty vector (Supplementary Figure 3), conforming that the cloned *TH* promoter is significantly active in C6 cells. To assess the role of GQ-forming sequence, TH49fm, the full mutant of wtTH49 that does not form any GQ, was used as a control. As shown in Figure 4B, the activity of wtTH49 was ~5-fold higher than that of TH49fm, implying that *TH* promoter function is enhanced by wtTH49's ability to form GQ structures. The observation mentioned above suggests that the overall GQs formed by wtTH49 sequence enhance *TH* gene expression; however, the effect is most likely an ensemble average of contributions from all of the GQs formed within wtTH49. To confirm that regulation of *TH* gene expression occurs at the

DOI: 10.1021/acs.biochem.5b00209
Biochemistry 2015, 54, 5533–5545

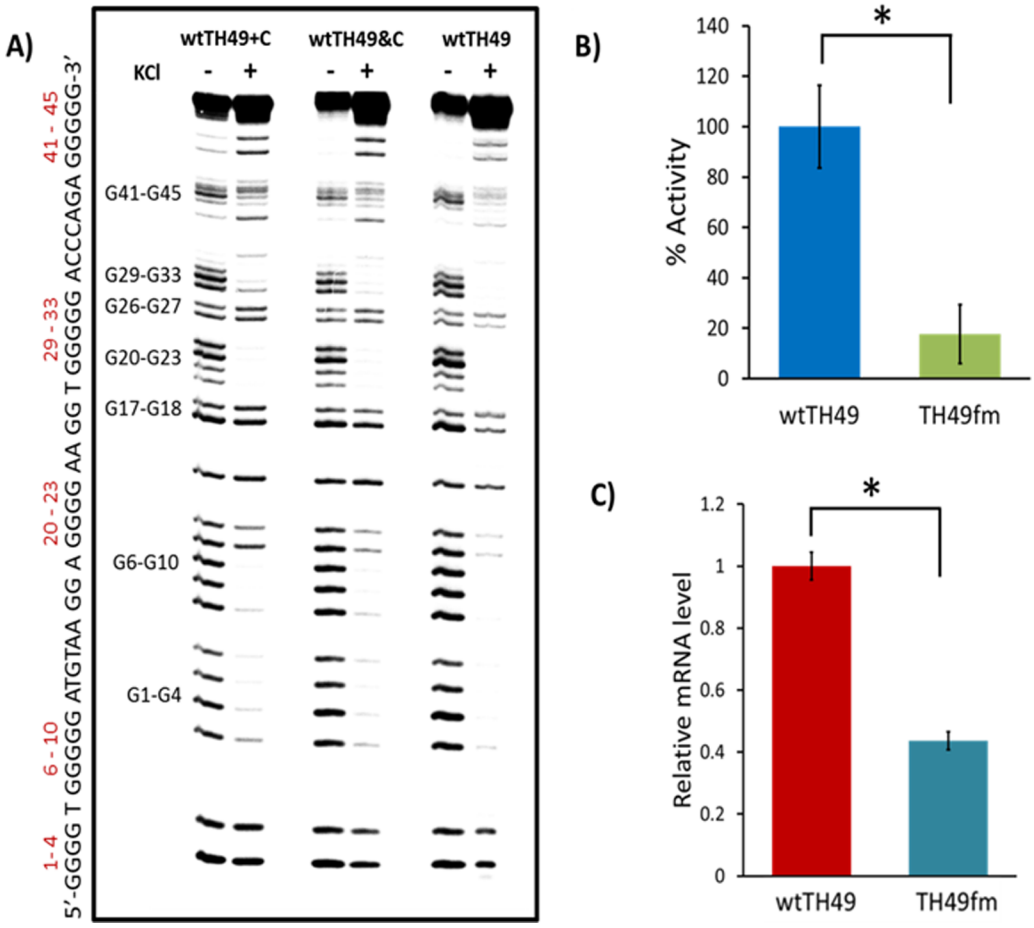


Figure 4. G-Quadruplex forming wtTH49 enhances the *TH* transcription by 5-fold. (A) The competition between GQ and double-stranded DNA was assessed using the complementary sequence of wtTH49. In the wtTH49+C lanes, the complementary sequence was added after the GQ structure had been folded. In the wtTH49&C lanes, the complementary sequence was added before the GQ structure had been folded. In the wtTH49 lanes, the complementary sequence was not added. (B) Percent luciferase activity of wtTH49 and the full mutant, TH49fm. The activity of the wild-type construct has been set to 100%, and the full mutant activity has been normalized accordingly (data represent mean values ± the standard error of the mean; $n = 9$; $^*P < 0.001$ from a Student's t test). (C) Histogram representing the ratio of firefly to Renilla luciferase mRNA levels in C6 cells. The ratio of CT values (obtained from qRT-PCR) for firefly to Renilla luciferase was determined using the Livak method.[38] The RNA level of the wild-type construct was set to 1, and the full mutant level was normalized accordingly (data represent mean values ± the standard error of the mean; $n = 9$; $^*P < 0.001$ from a Student's t test).

transcriptional level, we performed qRT-PCR on both the wtTH49 and TH49fm constructs. The qRT-PCR results indicate that the wtTH49 construct produced a mRNA level higher than that of mutant TH49fm (Figure 4C). This observation correlates the loss of activity of the mutant construct to an impediment at the transcriptional level. Although wtTH49 can adopt multiple GQs and the observed enhancement of activity by wtTH49 can be construed as an ensemble average, this finding is contrary to the general notion that GQs in the promoter regions act as transcription repressors.

Diverse G-Quadruplexes Formed by wtTH49 Sequence Lead to Differential Regulation of the Tyrosine Hydroxylase Promoter Activity. Next we investigated the role of specific GQ that can be formed using particular combinations of the G-stretches. To address this question, we designed two more mutant sequences (Figure 1B) based upon our DMS structure mapping results (Figure 3C) and the arrest sites observed in the DNA polymerase stop assay (Figure 3B). These mutants (5′GQ-MUT and 3′GQ-MUT) were designed to allow only a subset of GQs to be formed by each so that their roles can be deconvoluted. The 5′GQ-MUT construct eliminates the GQ-forming sequences at the 3′-end, while 3′GQ-MUT does not allow any GQ to be formed at the 5′-end of wtTH49. Initially, we established the ability of these two mutants to form GQs using CD spectroscopy and DMS structure mapping. CD studies performed on 5′GQ-MUT and 3′GQ-MUT showed spectral signatures similar to that of wtTH49 in the presence of 100 mM KCl, suggesting that they too most likely adopt a parallel GQ conformation (Supplementary Figure 4). However, 3′GQ-MUT does not show any characteristic feature of GQ in the absence of K^+. Although 5′GQ-MUT shows characteristic features for GQ even in the absence of K^+, its molar ellipticity significantly increased with the addition of K^+ ions, further confirming its ability to form K^+ ion-dependent GQ structure.

We then proceeded to identify the guanine residues involved in GQ formation in these mutants using DMS structure mapping. In the presence of K^+ ions, strong protection from methylation was observed in four G-stretches in both 5′GQ-MUT [G-stretches I, II, IV, and VI (see the numbering in Figure 1C)] and 3′GQ-MUT [G-stretches II, IV, VI, and VII

DOI: 10.1021/acs.biochem.5b00209
Biochemistry 2015, 54, 5533−5545

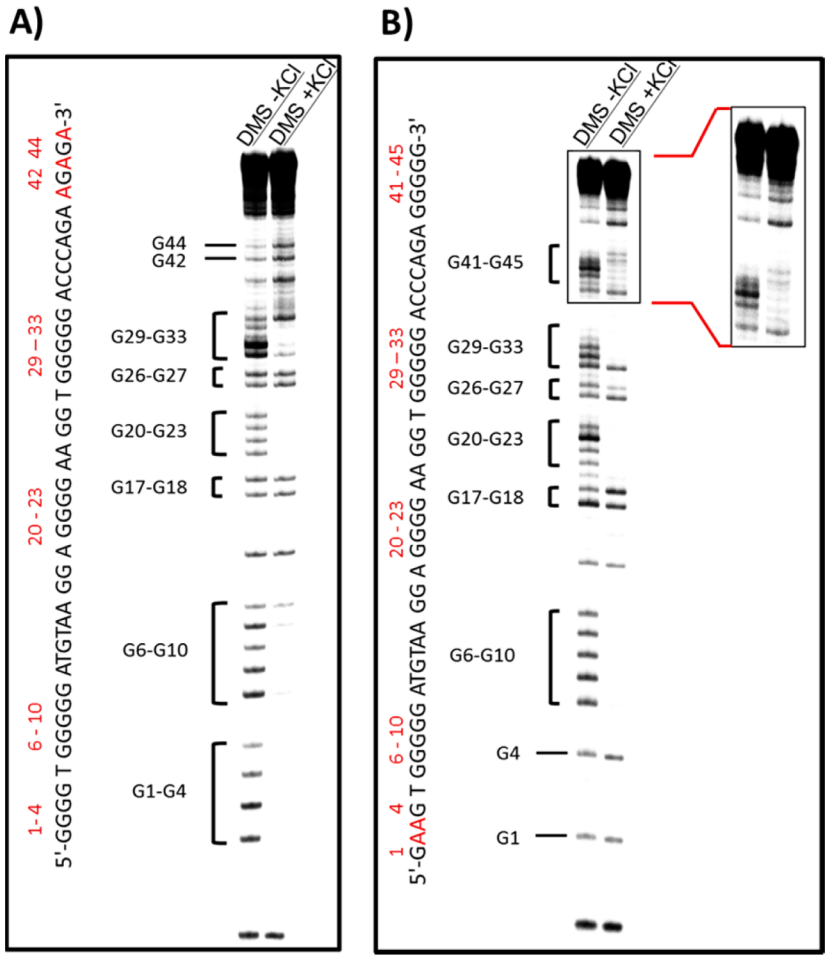


Figure 5. Both mutants adopt GQs using different combinations of G-stretches. (A) DMS structural mapping of 5′GQ-MUT in the absence of KCl (lane 1) and with 100 mM KCl (lane 2). (B) DMS structural mapping of 3′GQ-MUT in the absence of KCl (lane 1) and with 100 mM KCl (lane 2).

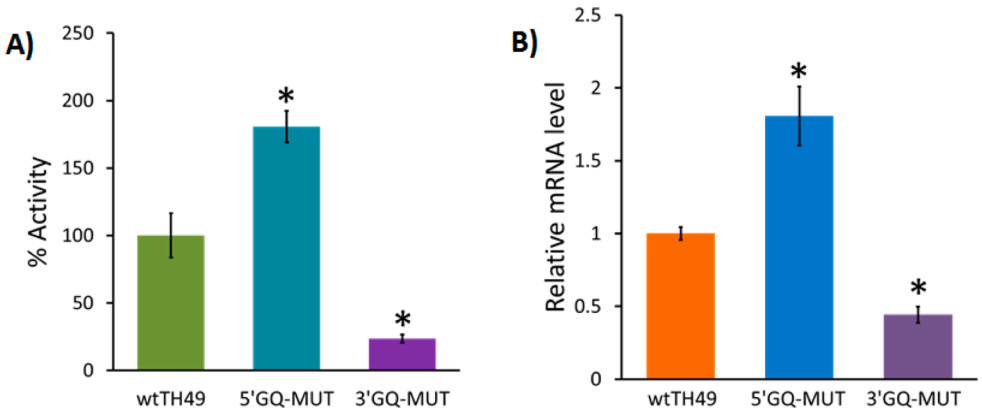


Figure 6. Multiple GQs formed by wtTH49 (5′GQ-MUT and 3′GQ-MUT) differentially enhance *TH* transcription. (A) Luciferase activities of wtTH49, 5′GQ-MUT, and 3′GQ-MUT. The 5′- and 3′-mutant activities have been normalized assuming 100% activity for wtTH49 (data represent mean values ± the standard error of the mean; $n = 9$; $^*P < 0.001$ from a Student's t test). (B) Histogram representing the ratio of firefly to Renilla luciferase mRNA levels in C6 cells. The ratio of CT values (obtained from qRT-PCR) for firefly to Renilla luciferase was determined by setting the value for the wtTH49 plasmid to 1, and all other values were normalized accordingly (data represent mean values ± the standard error of the mean; $n = 9$; $^*P < 0.001$ from a Student's t test).

(see the numbering in Figure 1C)], indicating formation of GQs (Figure 5A,B). Thus, DMS structure mapping and CD spectra support GQ formation by both 5′GQ-MUT and 3′GQ-MUT using different sets of G-stretches, presumably in accordance with the GQs formed within the wtTH49 sequence.

Two more luciferase constructs were prepared in which the 5′GQ-MUT or 3′GQ-MUT sequences were inserted in place of the wtTH49 sequence. 5′GQ-MUT showed ∼2-fold higher activity than wtTH49, while 3′GQ-MUT showed ∼4-fold lower activity than wtTH49 in C6 cells (Figure 6A). Using qRT-PCR

DOI: 10.1021/acs.biochem.5b00209
Biochemistry 2015, 54, 5533−5545

with the mutant constructs, we concluded that the modulation of *TH* promoter activity indeed occurs at the transcriptional level (Figure 6B). The data show that wtTH49 is involved in a range of activities in terms of modulating gene expression that depends upon the set of GQ structures that are utilized.

A luciferase assay in human neuroblastoma SHSY5Y cells was performed to validate the GQ-mediated regulation of human *TH* promoter activity observed in C6 cells. We detected ~6-fold higher activity for the wtTH49 promoter compared to that of the TH49fm the full mutant, whose activity is a direct result of complete lack of GQ formation (Figure 7A). Further, 5′GQ-MUT showed activity ~2-fold higher than that of wtTH49, and the 3′GQ-MUT showed activity ~2-fold lower than that of wtTH49 (Figure 7B). Thus, the human *TH* promoter activity in human neuroblastoma SHSY5Y cells is very similar to the activity observed in rat glioblastoma C6 cells. Because of the size of the plasmid, transfection in the SHSY5Y cells was extremely difficult, and also C6 cells have been used previously by several groups for human *TH* gene analysis that led us to initially choose C6 cells as a model.[19,45,46] However, our data directly prove that the human *TH* promoter activity in SHSY5Y cells is similar to that in C6 cells.

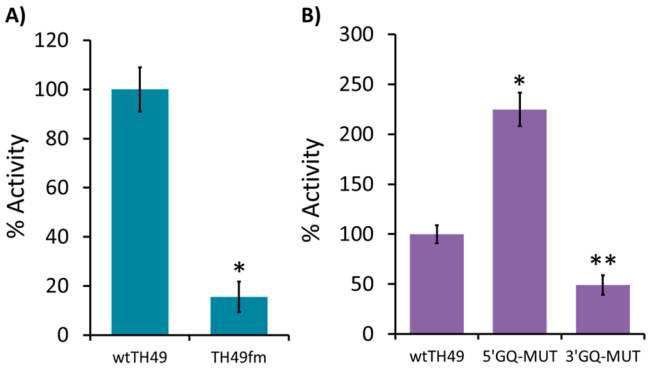


Figure 7. G-Quadruplex-mediated regulation of human *TH* promoter activity in human neuroblastoma cells, SHSY5y. (A) Percent luciferase activity of wtTH49 and the full mutant, TH49fm. The activity of the wild-type construct has been set to 100%, and the full mutant activity has been normalized accordingly (data represent mean values ± the standard error of the mean; $n = 5$; $^*P < 0.001$ from a Student's t test). (B) Luciferase activities of wtTH49, 5′GQ-MUT, and 3′GQ-MUT. The 5′- and 3′-mutant activities have been normalized assuming 100% activity for wtTH49 (data represent mean values ± the standard error of the mean; $n = 5$; $^*P < 0.001$ from a Student's t test; $^{**}P < 0.005$).

A Cationic Porphyrin Destabilizes the G-Quadruplex-Enabling Region within the Human Tyrosine Hydroxylase Promoter and Decreases the Level of Transcription. On the basis of our data from the functional assay, it was clear that the GQ-enabling sequence within the human *TH* promoter is a necessary element for transcription. We intended to interfere with the GQ structure to test its effect on transcription. We used the cationic porphyrin molecule TMPyP4 that is reported to act as both a stabilizer and a destabilizer of GQ structures depending upon the target sequence.[25,35] As a control small molecule, we used another cationic porphyrin molecule, TMPyP2, which is a positional isomer of TMPyP4. However, it has been reported that TMPyP2 has a low affinity for GQ structures compared to that of TMPyP4.[47] Consequently, genes that bear potential GQ-forming sequences such as c-myc and human telomerase reverse transcriptase are regulated by only TMPyP4 but not by TMPyP2 via stabilizing or destabilizing GQ structures.[25,35,47] Although TMPyP4 can interact with both GQ and duplex DNA, it has a higher selectivity for the GQ structure than for the duplex structure.[48]

Initially, we tested how TMPyP4 affects the GQ formed by the wtTH49 sequence using CD spectroscopy. A progressive decrease in the magnitude of CD signals was observed as more TMPyP4 was added, indicating the disappearance of the GQ (Supplementary Figure 5A). Given that the CD signals are not the most conclusive in terms of GQ characterization, DMS footprinting was performed, which is an independent and reliable probe of GQ formation. DMS footprinting of wtTH49 in the presence of TMPyP4 showed a decreased level of protection of G-tracts (Supplementary Figure 6A) indicating the loss of the GQ structure presumably due to destabilization. Thus, a change in the spectral intensity and the CD trace along with the DMS footprinting pattern show that TMPyP4 indeed destabilizes GQs in wtTH49.

Given that GQ formation by wtTH49 sequence is needed for the optimal transcription activity of the promoter and TMPyP4 destabilizes the GQ structure formed by the wtTH49 sequence, the small molecule should act as an inhibitor and lower the level of transcription imitating the TH49fm. We transfected a reporter plasmid under the control of the human *TH* promoter described earlier and treated the cells with different concentrations of TMPyP4. The results showed significant loss of transcriptional activity, which is consistent with destabilization of GQs formed by wtTH49 by TMPyP4 (Figure 8A). When similar experiments were performed using

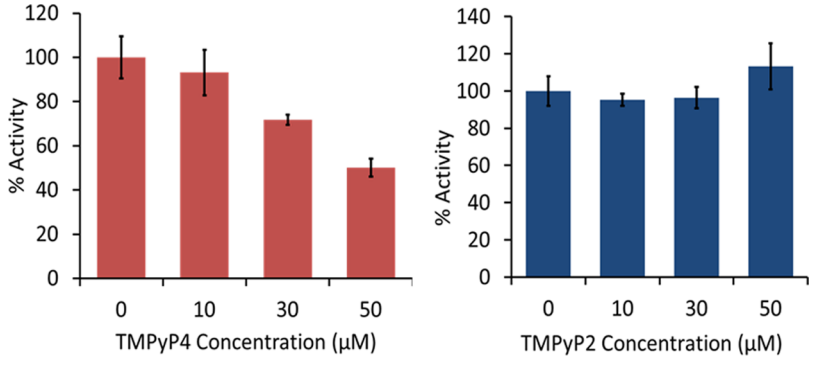


Figure 8. Small molecule-mediated regulation of *TH* transcription. (A) Percent luciferase activity of wtTH49 as a function of TMPyP4 concentration. (B) Percent luciferase activity of wtTH49 as a function of TMPyP2 concentration.

DOI: 10.1021/acs.biochem.5b00209
Biochemistry 2015, 54, 5533−5545

TMPyP2, which is a close analogue of TMPyP4, we did not observe a significant decrease in CD ellipticity (Supplementary Figure 5B) or deprotection of G-stretches in DMS measurements for the wtTH49 sequence (Supplementary Figure 6B), implying that TMPyP2 does not destabilize the GQ formed by wtTH49. These *in vitro* observations were further confirmed by performing a *TH* promoter-driven luciferase reporter assay in TMPyP2-treated cells. As shown in Figure 8B, no change in reporter activity was observed in the presence of various concentrations of TMPyP2, suggesting that it does not affect the stability of the GQ adopted by wtTH49. The data suggest that TMPyP4-mediated effects are specifically related to their influence of the GQ stability.

DISCUSSION

GQ-forming sequences located in promoter regions have been typically identified as transcription repressors,[25,28,49] but here we report a GQ that in fact is required for transcription. A majority of the G-rich sequences that are characterized so far are considered to adopt one GQ structure, possibly with different conformations.[1,2,50] wtTH49 with a series of G-stretches within its sequence creates an unprecedented complexity in the variety of stable GQ structural combinations that can form. This report provides evidence that the redundancy of the G-stretches embedded within wtTH49 enables formation of multiple GQ structures that could act as a tunable switch for regulating gene expression at the transcriptional level perhaps in the presence of trans elements. Two of the well-studied promoter sequences that contain more than four G-stretches are c-myc and Bcl-2.[1,4,25] Hurley and co-workers have reported formation of multiple GQ structures in G-rich c-myc and Bcl-2 promoters, which function as transcription repressors.[4,25] Recently, Richter and co-workers have reported that a 58 nt HIV 1 LTR promoter region forms multiple GQ structures. However, mutational studies of the sequence showed that GQ structures act as a repressor of transcription in this system.[5] It is clear that among the numerous potential combinations of G-stretches in wtTH49 that can take part in GQ formation, a subset will be thermodynamically more favorable based upon variables such as the loop length and the number of G-quartets.[51,52] Along these lines, the smFRET data indicate the presence of at least two distinct GQ structures. Interestingly, mutations 3′GQ-MUT and 5′GQ-MUT have differential impacts on promoter activity compared to that of wtTH49 (Figure 6A,B), which is unprecedented. 5′GQ-MUT, in which only 3′-end GQ formation is excluded, increases the promoter activity by ∼2-fold compared to that of wtTH49. On the other hand, 3′GQ-MUT, in which only 5′-end GQ formation is eliminated, decreases the level of transcription by ∼4-fold compared to that of wtTH49. On the basis of these observations, we propose that utilization of different combinations of G-stretches to form specific GQs can be a potential mechanism for the regulation of *TH* expression at the transcriptional level. However, further analysis is required to confirm the switching between possible GQ structures in the naturally occurring sequence and the factors that may potentially regulate such a process. A detailed study of the rat tyrosine hydroxylase promoter was recently published.[32] Although there are conserved regions in the human and rat promoter primary sequences, a key difference is the length of the G-stretches, which can have a significant impact on the GQs that can be formed. The primary sequence of the rat G-rich region can allow only two-tier GQs, while the human sequence forms three-tier or four-tier GQs, which are significantly more stable than the two-tier GQs.[33,34] Thus, the human G-rich region (wtTH49) and the rat promoter may possibly differ in terms of how they control transcription. The authors showed that heterogeneous nuclear ribonucleoprotein (hnRNP) K acts in trans to activate rat *Th* transcription. It binds to conserved G:C-rich regions of the rat *Th*-proximal promoter. hnRNP K is known to have a high affinity for poly C and binds to G-C-rich regions found in other sequences, including those present in promoters.[53] Upregulation of hnRNP K showed an increased level of transcription of the *Th* gene. The authors have also shown that cationic porphyrin TMPyP4 leads to suppression of transcription and have assumed that this must be due to stabilization of the GQ in the G-rich region. However, there are no data to directly show that TMPyP4 does stabilize the relevant GQ structures. As cationic porphyrin TMPyP4 is known to act either as a stabilizer or a destabilizer of GQ structures depending upon the sequence, it is not clear how the results of the two studies can be compared without a clarification of this critical point.[25,35] If TMPyP4 does indeed stabilize the GQ structures in the rat *Th*-proximal promoter that results in a reduced transcription product formation rate, then the findings of that study and ours would be in direct contrast to each other. If on the other hand TMPyP4 actually destabilizes the GQ, then it would be interesting to reevaluate the results of Banerjee et al. from this perspective. For example, it was shown that exposure to TMPyP4 reduces the hnRNP K occupancy in the *Th* promoter in SH-SY5Y cells. This might be explained by TMPyP4 destabilizing the GQ and promoting duplex formation between the C-rich and G-rich strands. As the C-rich strand would no longer be available, a decrease in the hnRNP K occupancy of the *Th* promoter would be expected. This would also imply that GQ formation in fact promotes *Th* transcription, as destabilizing the GQ via TMPyP4, resulting in a reduced level of transcription. In our studies, both CD and footprinting data show that TMPyP4 destabilizes the GQs present in wtTH49. In a reporter gene assay, we show that TMPyP4 reduces the level of transcription, which is in consonance with our findings that the G-rich wtTH49 sequence is required for optimal transcription. Assuming both studies to be accurate in their respective interpretations of their data would imply that TMPyP4 acts differently in rat and human *Th* promoter, stabilizing the GQs in one and destabilizing them in the other. This would not be surprising given the difference in the primary sequence encompassing the conserved region in the rat and human sequences.

In the promoter regions, G-rich sequences are always accompanied by C-rich sequences in the complementary strand. Nevertheless, formation of the i-motif *in vivo* is less favorable at physiological pH.[54] In addition, the higher affinity of hnRNP K toward single-stranded C-rich DNA further reduces the possibility of C-rich sequences folding into an i-motif *in vivo* and in turn affecting *TH* transcription.[55]

There is another G:C-rich sequence adjacent to TH49 that is located 17 nt upstream of the wtTH49 sequence.[32] Theoretically, it can form a two-tier GQ structure based upon the sequence of the bottom strand. Although two-tier GQ formation is less probable in the presence of its complementary sequence, further analysis will be required to determine its ability to form GQ and its possible functional role in *TH* transcription.[33]

DOI: 10.1021/acs.biochem.5b00209
Biochemistry 2015, 54, 5533−5545

Our data on mutants that eliminate GQ formation at the 3′-end (5′GQ-MUT) and 5′-end (3′GQ-MUT) point to the significance of GQ formation at the 5′-end of the wild-type sequence. When GQ formation at the 5′-end of the wild-type sequence is prevented, as in 3′GQ-MUT, a significant decrease in transcription activity is observed. These data also imply that formation of GQ only at the 3′-end is not adequate to match wild-type activity. On the other hand, eliminating GQ formation at the 3′-end, as in 5′GQ-MUT, essentially reducing the competition faced during GQ formation at the 5′-end of the wild-type construct, results in enhanced activity as it facilitates formation of an essential GQ structure at the 5′-end. However, none of these mutants disrupt any of the functionally proven transcription factor-binding sites close to wtTH49. Of particular interest is the question of whether hnRNP K binding to the C-rich strand, which satisfies the binding requirement of three consecutive C's for poly-C-binding proteins,[56] is affected. This should not be an issue when DNA is in duplex form as it was shown that hnRNP K can be recruited to the promoter by CREB when the promoter is in its duplex form.[32] However, recruitment of hnRNP K to the promoter might be inhibited during transcription when the duplex is unwound to single-stranded DNA. As G's are mutated to A's to disrupt GQ formation in the G-rich strand, the C's in the C-rich strand are mutated to T's to preserve complementarity. This would in principle inhibit hnRNP K recruitment during transcription. In addition, Micholetti and co-workers also have shown that interaction between hnRNP K and TATA-binding protein (TBP) on the basal promoter facilitates gene transcription.[55] This effect would also in principle result in a reduction in the level of transcription if hnRNP K is not recruited to the promoter as efficiently in the mutant as it was in the wild-type sequence. Therefore, further analysis will be required to confirm the effect of mutants on hnRNP K binding with single-stranded C-rich sequence that has fever C's in the mutant DNA constructs. However, distinct activities of 5′GQ-MUT and 3′GQ-MUT imply that mutation on the C-rich strand does not have a significant effect on the *TH* promoter. Along these lines, because only three consecutive C's are adequate for recruiting hnRNP K, even the mutant DNA constructs, except TH49fm, will have a number of sites that can accommodate hnRNP K binding. Whether these observations can be understood in the context of the functionally significant segments located downstream of the wild-type structure is an important question that needs to be studied further. On the basis of our data, it would be tempting to propose a model in which the enhancement of transcription can be modulated by forming GQ in the 5′-end versus the 3′-end. Such a model would require a switching between the two GQs or a mechanism that selectively allows formation of one GQ over the other. This can be a novel mechanism for regulating the synthesis of dopamine and may have implications for diseases that are linked to the aberrant dopamine level. Additionally, the GQs in the *TH* promoter create potential sites that can be targeted by small molecules to regulate dopamine production via stabilization or destabilization of the structure.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.5b00209.

One table listing the sequences of the DNA constructs and six figures showing (1) electrophoretic mobility data of wtTH49 and TH49fm, (2) smFRET data on wtTH49 at different salt concentrations and sample time traces showing the dynamics of the system, (3) the luciferase activity of wtTH49 and the empty pGL3 basic vector, (4) CD data on wtTH49 and mutants with and without K^+, (5) CD data on wtTH49 showing TMPyP4 and TMPyP2 titration, and (6) DMS structural mapping of wtTH49 in the presence of various concentrations of TMPyP4 and TMPyP2 (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: sbasu@kent.edu.

Present Address

§M.T.: Department of Biochemistry and Biophysics, University of Pennsylvania School of Medicine, Philadelphia, PA 19104.

Funding

This work was supported by National Institutes of Health Grant 1R15GM096285-01A1 to S.B. and H.B.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Dr. L. Iacovitti (Thomas Jefferson University, Philadelphia, PA) for her kind gift of the human tyrosine hydroxylase promoter clone.

■ REFERENCES

(1) Onyshchenko, M. I., Gaynutdinov, T. I., Englund, E. A., Appella, D. H., Neumann, R. D., and Panyutin, I. G. (2009) Stabilization of G-quadruplex in the BCL2 promoter region in double-stranded DNA by invading short PNAs. *Nucleic Acids Res. 37*, 7570−7580.

(2) De Armond, R., Wood, S., Sun, D., Hurley, L. H., and Ebbinghaus, S. W. (2005) Evidence for the presence of a guanine quadruplex forming region within a polypurine tract of the hypoxia inducible factor 1α promoter. *Biochemistry 44*, 16341−16350.

(3) Rankin, S., Reszka, A. P., Huppert, J., Zloh, M., Parkinson, G. N., Todd, A. K., Ladame, S., Balasubramanian, S., and Neidle, S. (2005) Putative DNA quadruplex formation within the human c-kit oncogene. *J. Am. Chem. Soc. 127*, 10584−10589.

(4) Dexheimer, T. S., Sun, D., and Hurley, L. H. (2006) Deconvoluting the structural and drug-recognition complexity of the G-quadruplex-forming region upstream of the bcl-2 P1 promoter. *J. Am. Chem. Soc. 128*, 5404−5415.

(5) Perrone, R., Nadai, M., Frasson, I., Poe, J. A., Butovskaya, E., Smithgall, T. E., Palumbo, M., Palù, G., and Richter, S. N. (2013) A dynamic G-quadruplex region regulates the HIV-1 long terminal repeat promoter. *J. Med. Chem. 56*, 6521−6530.

(6) Nagatsu, T., Levitt, M., and Udenfriend, S. (1964) Tyrosine hydroxylase: The initial step in norepinephrine biosynthesis. *J. Biol. Chem. 239*, 2910−2917.

(7) Zigmond, R. E., Schwarzschild, M. A., and Rittenhouse, A. R. (1989) Acute regulation of tyrosine hydroxylase by nerve activity and by neurotransmitters via phosphorylation. *Annu. Rev. Neurosci. 12*, 415−461.

(8) Arenzana, F. J., Arévalo, R., Sánchez-González, R., Clemente, D., Aijón, J., and Porteros, A. (2006) Tyrosine hydroxylase immunoreactivity in the developing visual pathway of the zebrafish. *Anat. Embryol. 211*, 323−334.

(9) Pickel, V. M., Joh, T. H., Field, P. M., Becker, C. G., and Reis, D. J. (1975) Cellular localization of tyrosine hydroxylase by immunohistochemistry. *J. Histochem. Cytochem. 23*, 1−12.

DOI: 10.1021/acs.biochem.5b00209
Biochemistry 2015, 54, 5533−5545

(10) Reddy, S. D. N., Rayala, S. K., Ohshiro, K., Pakala, S. B., Kobori, N., Dash, P., Yun, S., Qin, J., O'Malley, B. W., and Kumar, R. (2011) Multiple coregulatory control of tyrosine hydroxylase gene transcription. *Proc. Natl. Acad. Sci. U. S. A. 108*, 4200−4205.
(11) Furlong, R. A., Rubinsztein, J. S., Ho, L., Walsh, C., Coleman, T. A., Muir, W. J., Paykel, E. S., Blackwood, D. H. R., and Rubinsztein, D. C. (1999) Analysis and metaanalysis of two polymorphisms within the tyrosine hydroxylase gene in bipolar and unipolar affective disorders. *Am. J. Med. Genet. 88*, 88−94.
(12) Ishiguro, H., Arinami, T., Saito, T., Akazawa, S., Enomoto, M., Mitushio, H., Fujishiro, H., Tada, K., Akimoto, Y., Mifune, H., Shiozuka, S., Hamaguchi, H., Toru, M., and Shibuya, H. (1998) Systematic search for variations in the tyrosine hydroxylase gene and their associations with schizophrenia, affective disorders, and alcoholism. *Am. J. Med. Genet. 81*, 388−396.
(13) Kunugi, H., Kawada, Y., Hattori, M., Ueki, A., Otsuka, M., and Nanko, S. (1998) Association study of structural mutations of the tyrosine hydroxylase gene with schizophrenia and parkinson's disease. *Am. J. Med. Genet. 81*, 131−133.
(14) Moore, D. J., West, A. B., Dawson, V. L., and Dawson, T. M. (2005) Molecular pathophysiology of Parkinson's disease. *Annu. Rev. Neurosci. 28*, 57−87.
(15) Kessler, M. A., Yang, M., Gollomp, K. L., Jin, H., and Iacovitti, L. (2003) The human tyrosine hydroxylase gene promoter. *Mol. Brain Res. 112*, 8−23.
(16) Romano, G., Suon, S., Jin, H., Donaldson, A. E., and Iacovitti, L. (2005) Characterization of five evolutionary conserved regions of the human tyrosine hydroxylase (TH) promoter: Implications for the engineering of a human TH minimal promoter assembled in a self-inactivating lentiviral vector system. *J. Cell. Physiol. 204*, 666−677.
(17) Gandelman, K. Y., Coker, G. T., Moffat, M., and O'Malley, K. L. (1990) Species and regional differences in the expression of cell-type specific elements at the human and rat tyrosine hydroxylase gene loci. *J. Neurochem. 55*, 2149−2152.
(18) Kim, K. S., Kim, C. H., Hwang, D. Y., Seo, H., Chung, S., Hong, S. J., Lim, J. K., Anderson, T., and Isacson, O. (2003) Orphan nuclear receptor Nurr1 directly transactivates the promoter activity of the tyrosine hydroxylase gene in a cell-specific manner. *J. Neurochem. 85*, 622−634.
(19) Kim, T. E., Park, M. J., Choi, E. J., Lee, H. S., Lee, S. H., Yoon, S. H., Oh, C. K., Lee, B. J., Kim, S. U., Lee, Y. S., and Lee, M. A. (2003) Cloning and cell type-specific regulation of the human tyrosine hydroxylase gene promoter. *Biochem. Biophys. Res. Commun. 312*, 1123−1131.
(20) Matsushita, N., Okada, H., Yasoshima, Y., Takahashi, K., Kiuchi, K., and Kobayashi, K. (2002) Dynamics of tyrosine hydroxylase promoter activity during midbrain dopaminergic neuron development. *J. Neurochem. 82*, 295−304.
(21) Kumer, S. C., and Vrana, K. E. (1996) Intricate regulation of tyrosine hydroxylase activity and gene expression. *J. Neurochem. 67*, 443−462.
(22) Henderson, E., Hardin, C. C., Walk, S. K., Tinoco, I., Jr, and Blackburn, E. H. (1987) Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine·guanine base pairs. *Cell 51*, 899−908.
(23) Williamson, J. R., Raghuraman, M. K., and Cech, T. R. (1989) Monovalent cation-induced structure of telomeric DNA: The G-quartet model. *Cell 59*, 871−880.
(24) Blackburn, E. H. (1991) Structure and function of telomeres. *Nature 350*, 569−573.
(25) Siddiqui-Jain, A., Grand, C. L., Bearss, D. J., and Hurley, L. H. (2002) Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci. U. S. A. 99*, 11593−11598.
(26) Sun, D., Guo, K., Rusche, J. J., and Hurley, L. H. (2005) Facilitation of a structural transition in the polypurine/polypyrimidine tract within the proximal promoter region of the human VEGF gene by the presence of potassium and G-quadruplex-interactive agents. *Nucleic Acids Res. 33*, 6070−6080.
(27) Guo, K., Pourpak, A., Beetz-Rogers, K., Gokhale, V., Sun, D., and Hurley, L. H. (2007) Formation of pseudosymmetrical G-quadruplex and i-motif structures in the proximal promoter region of the RET oncogene. *J. Am. Chem. Soc. 129*, 10220−10228.
(28) Cogoi, S., and Xodo, L. E. (2006) G-quadruplex formation within the promoter of the KRAS proto-oncogene and its effect on transcription. *Nucleic Acids Res. 34*, 2536−2549.
(29) Qin, Y., Rezler, E. M., Gokhale, V., Sun, D., and Hurley, L. H. (2007) Characterization of the G-quadruplexes in the duplex nuclease hypersensitive element of the PDGF-A promoter and modulation of PDGF-A promoter activity by TMPyP4. *Nucleic Acids Res. 35*, 7698−7713.
(30) Palumbo, S. L., Memmott, R. M., Uribe, D. J., Krotova-Khan, Y., Hurley, L. H., and Ebbinghaus, S. W. (2008) A novel G-quadruplex-forming GGA repeat region in the c-myb promoter is a critical regulator of promoter activity. *Nucleic Acids Res. 36*, 1755−1769.
(31) Lewis, E. J., Harrington, C. A., and Chikaraishi, D. M. (1987) Transcriptional regulation of the tyrosine hydroxylase gene by glucocorticoid and cyclic AMP. *Proc. Natl. Acad. Sci. U. S. A. 84*, 3550−3554.
(32) Banerjee, K., Wang, M., Cai, E., Fujiwara, N., Baker, H., and Cave, J. W. (2014) Regulation of tyrosine hydroxylase transcription by hnRNP K and DNA secondary structure. *Nat. Commun. 5*, 5769−5781.
(33) Pandey, S., Agarwala, P., and Maiti, S. (2013) Effect of loops and G-quartets on the stability of RNA G-quadruplexes. *J. Phys. Chem. B 117*, 6896−6905.
(34) Joachimi, A., Benz, A., and Hartig, J. S. (2009) A comparison of DNA and RNA quadruplex structures and stabilities. *Bioorg. Med. Chem. 17*, 6811−6815.
(35) Morris, M. J., Wingate, K. L., Silwal, J., Leeper, T. C., and Basu, S. (2012) The porphyrin TMPyP4 unfolds the extremely stable G-quadruplex in MT3-MMP mRNA and alleviates its repressive effect to enhance translation in eukaryotic cells. *Nucleic Acids Res. 40*, 4137−4145.
(36) Weitzmann, M. N., Woodford, K. J., and Usdin, K. (1996) The Development and Use of a DNA Polymerase Arrest Assay for the Evaluation of Parameters Affecting Intrastrand Tetraplex Formation. *J. Biol. Chem. 271*, 20958−20964.
(37) Ray, S., Qureshi, Mohammad, H., Malcolm, Dominic, W., Budhathoki, Jagat, B., Çelik, U., and Balci, H. (2013) RPA-mediated unfolding of systematically varying G-quadruplex structures. *Biophys. J. 104*, 2235−2245.
(38) Sen, D., and Gilbert, W. (1990) A sodium-potassium switch in the formation of four-stranded G4-DNA. *Nature 344*, 410−414.
(39) Rujan, I. N., Meleney, J. C., and Bolton, P. H. (2005) Vertebrate telomere repeat DNAs favor external loop propeller quadruplex structures in the presence of high concentrations of potassium. *Nucleic Acids Res. 33*, 2022−2031.
(40) Aboul-ela, F., Murchie, A. I. H., and Lilley, D. M. J. (1992) NMR study of parallel-stranded tetraplex formation by the hexadeoxynucleotide d(TG4T). *Nature 360*, 280−282.
(41) Matsugami, A., Okuizumi, T., Uesugi, S., and Katahira, M. (2003) Intramolecular higher order packing of parallel quadruplexes comprising a G:G:G:G tetrad and a G(:A):G(:A):G(:A):G heptad of GGA triplet repeat DNA. *J. Biol. Chem. 278*, 28147−28153.
(42) Nambiar, M., Srivastava, M., Gopalakrishnan, V., Sankaran, S. K., and Raghavan, S. C. (2013) G-quadruplex structures formed at the HOX11 breakpoint region contribute to its fragility during t(10;14) translocation in T-cell leukemia. *Mol. Cell. Biol. 33*, 4266−4281.
(43) Ray, S., Bandaria, J. N., Qureshi, M. H., Yildiz, A., and Balci, H. (2014) G-quadruplex formation in telomeres enhances POT1/TPP1 protection against RPA binding. *Proc. Natl. Acad. Sci. U. S. A. 111*, 2990−2995.
(44) Han, H., Hurley, L. H., and Salazar, M. (1999) A DNA polymerase stop assay for G-quadruplex-interactive compounds. *Nucleic Acids Res. 27*, 537−542.
(45) Kim, S. M., Yang, J. W., Park, M. J., Lee, J. K., Kim, S. U., Lee, Y. S., and Lee, M. A. (2006) Regulation of human tyrosine hydroxylase

DOI: 10.1021/acs.biochem.5b00209
Biochemistry 2015, 54, 5533−5545

gene by neuron-restrictive silencer factor. *Biochem. Biophys. Res. Commun. 346*, 426−435.
(46) Guerrero-Cazares, H., Alatorre-Carranza, M. d. P., Delgado-Rizo, V., Duenas-Jimenez, J. M., Mendoza-Magana, M. L., Morales-Villagran, A., Ramirez-Herrera, M. A., Guerrero-Hernandez, A., Segovia, J., and Duenas-Jimenez, S. H. (2007) Dopamine release modifies intracellular calcium levels in tyrosine hydroxylase-transfected C6 cells. *Brain Res. Bull. 74*, 113−118.
(47) Grand, C. L., Han, H., Munoz, R. M., Weitman, S., Von Hoff, D. D., Hurley, L. H., and Bearss, D. J. (2002) The cationic porphyrin TMPyP4 down-regulates c-MYC and human telomerase reverse transcriptase expression and inhibits tumor growth in vivo. *Mol. Cancer. Ther. 1*, 565−573.
(48) Martino, L., Pagano, B., Fotticchia, I., Neidle, S., and Giancola, C. (2009) Shedding light on the interaction between TMPyP4 and human telomeric quadruplexes. *J. Phys. Chem. B 113*, 14779−14786.
(49) Mitchell, T., Ramos-Montoya, A., Di Antonio, M., Murat, P., Ohnmacht, S., Micco, M., Jurmeister, S., Fryer, L., Balasubramanian, S., Neidle, S., and Neal, D. E. (2013) Downregulation of androgen receptor transcription by promoter G-quadruplex stabilization as a potential alternative treatment for castrate-resistant prostate cancer. *Biochemistry 52*, 1429−1436.
(50) Sun, D., Liu, W. J., Guo, K., Rusche, J. J., Ebbinghaus, S., Gokhale, V., and Hurley, L. H. (2008) 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. *Mol. Cancer Ther. 7*, 880−889.
(51) Zhang, A. Y., Bugaut, A., and Balasubramanian, S. (2011) A sequence-independent analysis of the loop length dependence of intramolecular RNA G-quadruplex stability and topology. *Biochemistry 50*, 7251−7258.
(52) Risitano, A., and Fox, K. R. (2004) Influence of loop size on the stability of intramolecular DNA quadruplexes. *Nucleic Acids Res. 32*, 2598−2606.
(53) Uribe, D. J., Guo, K., Shin, Y.-J., and Sun, D. (2011) Heterogeneous nuclear ribonucleoprotein K and nucleolin as transcriptional activators of the vascular endothelial growth factor promoter through interaction with secondary DNA structures. *Biochemistry 50*, 3796−3806.
(54) Fernández, S., Eritja, R., Aviñó, A., Jaumot, J., and Gargallo, R. (2011) Influence of pH, temperature and the cationic porphyrin TMPyP4 on the stability of the i-motif formed by the 5′-(C3TA2)4−3′ sequence of the human telomere. *Int. J. Biol. Macromol. 49*, 729−736.
(55) Michelotti, E. F., Michelotti, G. A., Aronsohn, A. I., and Levens, D. (1996) Heterogeneous nuclear ribonucleoprotein K is a transcription factor. *Mol. Cell. Biol. 16*, 2350−2360.
(56) Du, Z., Lee, J. K., Tjhen, R., Li, S., Pan, H., Stroud, R. M., and James, T. L. (2005) Crystal structure of the first KH domain of human poly(C)-binding protein-2 in complex with a C-rich strand of human telomeric DNA at 1.7 A. *J. Biol. Chem. 280*, 38823−38830.

DOI: 10.1021/acs.biochem.5b00209
Biochemistry 2015, 54, 5533−5545