BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

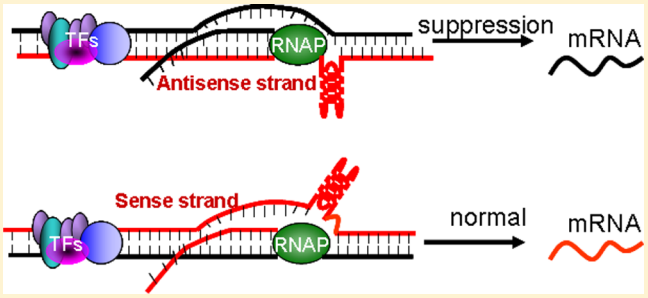
In the Sense of Transcription Regulation by G-Quadruplexes: Asymmetric Effects in Sense and Antisense Strands

Tani Agarwal,[†] Saumya Roy,[‡] Santosh Kumar,[†] Tushar Kanti Chakraborty,*[,‡,§] and Souvik Maiti*[,†,∥]

[†]Proteomics and Structural Biology Unit, CSIR-Institute of Genomics and Integrative Biology, Mall Road, Delhi 110 007, India
[‡]CSIR-Indian Institute of Chemical Technology, Uppal Road, Tarnaka, Hyderabad 500 007, India
[§]Department of Organic Chemistry, Indian Institute of Science, Bangalore 560 012, India
[∥]CSIR-National Chemical Laboratory, Dr. Homi Bhabha Road, Pune 411008, India

Supporting Information

ABSTRACT: G-Quadruplexes occupy important regulatory regions in the genome. DNA G-quadruplexes in the promoter regions and RNA quadruplexes in the UTRs (untranslated regions) have been individually studied and variously implicated at different regulatory levels of gene expression. However, the formation of G-quadruplexes in the sense and antisense strands and their corresponding roles in gene regulation have not been studied in much detail. In the present study, we have elucidated the effect of strand asymmetry in this context. Using biophysical methods, we have demonstrated the formation of stable G-quadruplex structure in vitro using CD and UV melting. Additionally, ITC was employed to demonstrate that a previously reported selective G-quadruplex ligand was able to bind and stabilize the G-quadruplex in the present sequence. Further, we have shown using reporter constructs that although the DNA G-quadruplex in either strand can reduce translation efficiency, transcriptional regulation differs when G-quadruplex is present in the sense or antisense strand. We demonstrate that the G-quadruplex motif in the antisense strand substantially inhibits transcription, while when in the sense strand, it does not affect transcription, although it does ultimately reduce translation. Further, it is also shown that the G-quadruplex stabilizing ligand can enhance this asymmetric transcription regulation as a result of the increased stabilization of the G-quadruplex.



In nature, guanine-rich stretches are found in several important genomic regions such as telomeres, centromeres, immunoglobulin switch regions, mutational hot spots, and promoter elements. These guanine-rich sequences may have the potential to form four-stranded secondary structures called G-quadruplexes stabilized by non-Watson−Crick interactions.[1] Hoogsteen bond stabilized G-quadruplex structures are often associated with important regulatory roles during transcription and translation. Several studies have demonstrated an extensive role of DNA G-quadruplexes in controlling transcription. Specific proteins like QUAD,[2−5] which selectively bind with G-quadruplex structures, have been discovered impinging on the biological roles of quadruplexes. Moreover, studies have shown that several proto-oncogenes genes like c-MYC, c-KIT, VEGF, PDGF, BCL-2, and HIF bear G-quadruplex motifs in their promoters.[6] These motifs may be efficiently targeted by G-quadruplex selective ligands to alter their molecular recognition and hence affect their downstream expression.[7−10] More recently discovered RNA G-quadruplexes are under extensive research owing to their higher feasibility of occurrence in the absence of competing complementary strands. Recent studies have demonstrated the influence of individual RNA G-quadruplexes found in 5′-UTR of natural genes or artificially introduced G-quadruplexes in prokaryotic or eukaryotic mRNAs,[11,12] thereby showing their significance in regulating translation processes. In all of these studies, RNA G-quadruplex in the 5′-UTR of the mRNA have been shown to suppress protein synthesis. In a previous study, we too have demonstrated that 5′-UTR G-quadruplex mediated suppression of protein synthesis in Zic-1 is due to the repression of translation rather than a consequence of reduced transcription,[13] which was subsequently confirmed by similar studies involving other genes.[14,15]

Being a well studied secondary structural motif implicated in gene regulation, several structural and functional aspects of the G-quadruplex have been investigated in detail including the effect of different sequences, loop lengths,[16,17] flank lengths,[18] the presence of various ions,[19,20] ligands,[21−26] and molecular crowding agents.[27−29] However, limited studies have been dedicated to understanding the importance of strand orientation in this context. Structural motifs in DNA can be present either on the sense/coding strand or on the antisense/template strand. Hence, strand orientation detailing as to which

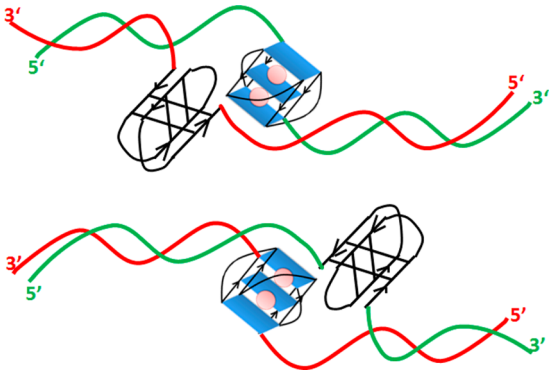
Received: October 25, 2013
Revised: May 20, 2014
Published: May 22, 2014

   dx.doi.org/10.1021/bi401451q | *Biochemistry* 2014, 53, 3711−3718

strand bears the G-quadruplex motif, sense or antisense, could greatly determine the impact of G-quadruplex upon downstream expression. In this context, bioinformatic studies by Du et al. have shown the striking predominance of DNA G-quadruplex motifs in the coding (sense) strand than in the template (antisense) strand in the vicinity of TSS (transcription start site) suggestive of asymmetric distribution of G-quadruplex motifs.[30] They suggest that as many as 5152 genes exhibit a strand asymmetric distribution of the PG4 (potential G-quadruplex) motif close to the TSS with a bigger subset comprising 3298 genes with more PG4 motifs in the coding strand than on the template strand near TSS. While the asymmetric distribution per se has been described and discussed, experimental exploration into the implications of this asymmetry on gene regulation remains to be reported. In order to understand this broad question, it would also be relevant to ask if the effect of the G-quadruplex motif could be asymmetric while sitting on different strands.

Here, we intend to uncover the impact of strand asymmetry in the context of G-quadruplex mediated biological function. We have attempted to distinguish the downstream expression pattern depending on the orientation of the G-quadruplex bearing strand. Toward this objective, we have created a model system of plasmids in which a putative G-quadruplex sequence adopted from the forward strand of RASSF1 5′UTR has been strategically placed just a few nucleotides upstream of the start codon in the *Renilla* luciferase gene, separately in either sense or antisense strands (Schemes 1 and 2) in a dual reporter vector. Using quantitative PCR and a reporter expression assay, we have concluded that the G-quadruplex at the sense strand is ineffective in altering transcription, while transcription is significantly sensitive to an antisense strand G-quadruplex. At the same time, we also demonstrate that the G-quadruplex at either strand can significantly reduce translation efficiency in a dual reporter system. In this situation, when a G-rich sequence capable of forming a G-quadruplex is placed in one strand, the complementary C-rich strand may be expected to form an i-motif, which has been explained further.

Scheme 1. Typical Structural Scenario When a G-Rich Sequence on a Sense Strand (Green) of DNA May Lead to the Formation of the G-Quadruplex (Secondary Structure in Blue) while at the Same Time the C-Rich Sequence on the Complementary/Antisense Strand (Red) Could Form an i-Motif Structure (Black)[a]



[a]The case vice versa is also represented when strand orientations get inverted.

MATERIALS AND METHODS

Oligonucleotide Sequences (DNA and RNA). HPLC purified DNA and RNA oligonucleotides were purchased from SIGMA Aldrich and used as such. The sequences are as follows: G-rich, 5′-GGCCGGGCGGGGCCACAGGGCGGG; C-rich, 5′-CCCGCCCTGTGGCCCCGCCCGGCC. For the RNA counterparts, T was suitably replaced with U.

CD Spectroscopy. CD spectra were recorded in a Jasco spectropolarimeter 50 (model 715, Japan) equipped with a thermoelectrically controlled cell holder and a cuvette with a path length of 1 cm. The oligonucleotides were heated at 95 °C for 5 min followed by slow programmed cooling (0.2 °C/min) in 10 mM sodium cacodylate buffer with 100 mM KCl at the indicated pH. CD spectra for G-quadruplexes (5 μM) was recorded between 220 and 370 nm at 25 °C, and the spectra obtained were the average of three scans.

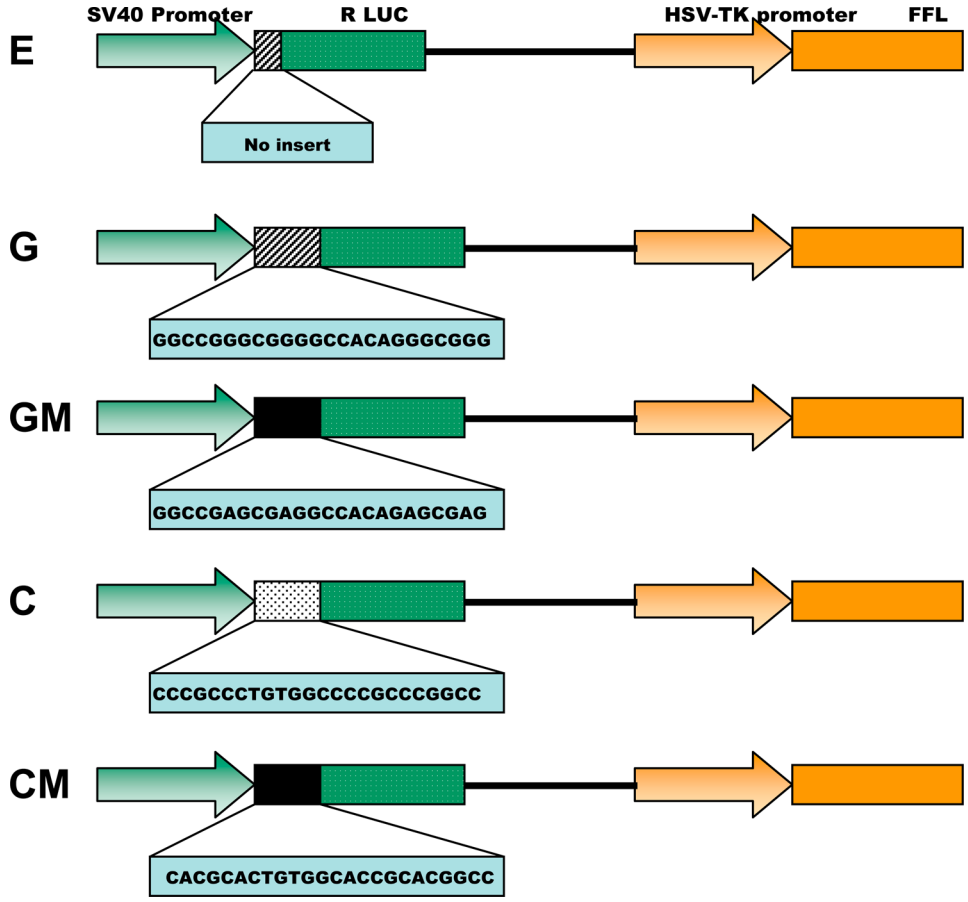
UV Melting Studies. Sample was prepared by heating the oligonucleotides in 10 mM sodium cacodylate buffer with 100 mM KCl, at the indicated pH, followed by slow cooling. The melting experiment was performed with 5 μM strand concentration in a Cary 400 (Varian) spectrophotometer equipped with a Hitachi SPR-10 thermo programmer at a heating rate of 0.2 °C/min. Data were collected at 295 or 260 nm. Typically, three replicate experiments were performed, and the average values are reported.

Isothermal Titration Calorimetry (ITC). ITC measurements were performed in a VP-ITC titration calorimeter (MicroCal, Northampton, MA). Before loading, the solutions were thoroughly degassed. The preformed quadruplex (5 μM) was kept in the sample cell, and the ligand (300 μM) in the same buffer was filled in a syringe of volume 300 μL. Ligand solution was added sequentially in 8 μL aliquots (for a total of 20 injections, 15 s duration each) at 3 min intervals at 25 °C. Sequential titrations were performed to ensure full occupancy of the binding sites by loading and titrating with the same ligand without removing the samples from the cell until the titration signal was essentially constant. The respective heat of dilution was subtracted from the corresponding binding experiments prior to curve fitting. Typically three replicate titration experiments were performed.

Cloning Methodology. The psiCHECK2 vector (E) was digested with the Nhe I restriction enzyme (single restriction site) producing sticky ends which were ligated with the double stranded insert. Nhe1 recognition sites were strategically placed on the 5′ ends of the mentioned sequences and were hybridized with their complementary strands at 37 °C to prepare the double stranded inserts. 5′-CTAGCGGCCGGGCGGGGC-CACAGGGCGGG-3′ (for plasmid G) and cytosine-rich sequence 5′-CTAGCCCGCCCTGTGGCCCCGCCCGGCC-3′ (for plasmid C) were inserted into the sense strand of the vector to generate plasmids G and C, respectively. The ligated product was transformed into *E. coli* cells, and the plasmid was isolated using an RBC mini-prep kit. The plasmid was then sequenced and further used for transfection. Furthermore, each triad of contiguous guanines in the insert were interrupted by adenines to generate mutant plasmids, (GM and CM) in a site directed mutagenesis (SDM) experiment according to the manufacturer's protocol using a Stratagene SDM kit.

RNA Isolation and Real Time PCR. HEK cells were transfected with the above plasmids using Lipofectamine 2000 from Invitrogen. Total RNA was isolated from 1 × 10^6 transfected cells using Tri-Reagent (Sigma Chemical Co.) as

Scheme 2. Scheme of Construction of Plasmids[a]



[a]R refers to *Renilla* luciferase, and FFL refers to the firefly luciferase.

recommended by the manufacturer. DNase treatment was performed according to the manufacturer's instructions using DNase from Invitrogen. One microgram of total RNA was reverse transcribed in a 20 μL reaction volume using the Revert Aid enzyme from Fermentas. The cDNA was hence amplified using SYBR Green Real Time PCR Master mix from SAFLABS in the Real Time PCR reactions. The amplification was done in LC 480 from ROCHE with the following cycling conditions for PCR: 94 °C for 2 min, followed by 35 cycles of 94 °C for 30 s, 58 °C for 30 s, and 72 °C for 30 s. Amplification of the firefly gene was performed as a control using the same PCR conditions.

Luciferase Reporter Assay. HEK cells grown to 70% confluency in a 24 well plate were either without or with increasing concentrations of the ligand (up to 100 μM). The cells were transfected with 250 ng of any one of the 5 above plasmids (E, G, GM, C, or CM) using Lipofectamine 2000 from Invitrogen. After 4 h, media were replaced with complete DMEM. The expression of firefly luciferase and *Renilla* luciferase was measured in a dual luciferase assay (Promega), 24 h after transfection (as described by the manufacturer). The cell lysate (20 μL) was mixed with 100 μL of reconstituted luciferase assay reagent, and light output was measured for 12 s with a FB12 Luminometer (Berthold).

RESULTS

We first performed in vitro studies using standard biophysical methods to assess the structures assumed by the above oligonucleotides and their stabilities. Using CD spectroscopy, we observed that the G-rich sequences were able to fold into a stable G-quadruplex at physiological pH (Figure 1a). G-rich DNA as well its RNA counterpart sequences showed a strong positive peak at 263 nm forming stable parallel G-quadruplexes. We have also performed an UV melting experiment and found that G-rich oligonucleotides, DNA as well as RNA, form stable G-quadruplexes at pH 7 with T_m of 61 and 71 °C, respectively (Figure 1b).

However, the C-rich sequence showed a limited potential to fold into a tetraplex structure at neutral pH. Nevertheless, as shown in Figure 1a, at acidic pH conditions this C-rich DNA sequence shows a peak at 295 nm, a signature characteristic of i-motif formation. Moreover, for the C-rich RNA counterpart, the peaks at 210, 240, and 270 nm indicated alternative RNA structure other than the i-motif, probably resembling the A-form of RNA. This result is well in accordance with previous reports suggesting the incapability of RNA oligonucleotides to form an i-motif. This failure of RNA to fold into the i-motif has been attributed to the impediment offered by the 2′-OH in its ribose sugar.[31,32] Further, we show in a UV melting experiment in Figure 1b that the cytosine-rich sequences at pH 5 showed a hyperchromic signal at 260 nm (Figure 1b). The T_m was recorded to be only 46 and 47 °C for DNA and RNA sequences, respectively.

However, we confirmed the in vitro formation of the i-motif structure by the C-rich DNA sequence from pH dependent enhancement of the folding in a CD titration experiment

dx.doi.org/10.1021/bi401451q | *Biochemistry* 2014, 53, 3711−3718

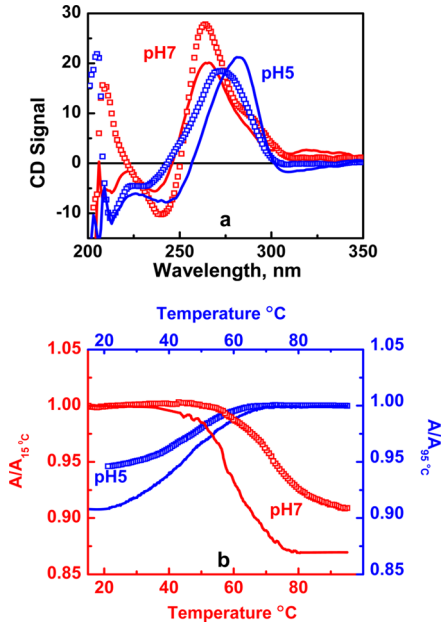


Figure 1. Structure and stability of the oligonucleotides. (a) Different CD spectra depicted by G-rich (red) and C-rich (blue) oligonucleotides show the signature for the G-quadruplex and i-motif, respectively. (b) UV melting profiles showed varying thermal stabilities of secondary structures formed by G-rich (red) and C-rich (blue) oligonulceotides. G-DNA is indicated as red lines and G-RNA as red squares. C-DNA is shown in blue lines and C-RNA as blue squares. All experiments were performed in 10 mM sodium cacodylate buffer at indicated pH.

(Figure 2a), whereas the C-rich RNA counterpart in a similar experiment showed collapsing CD signals with decreasing pH

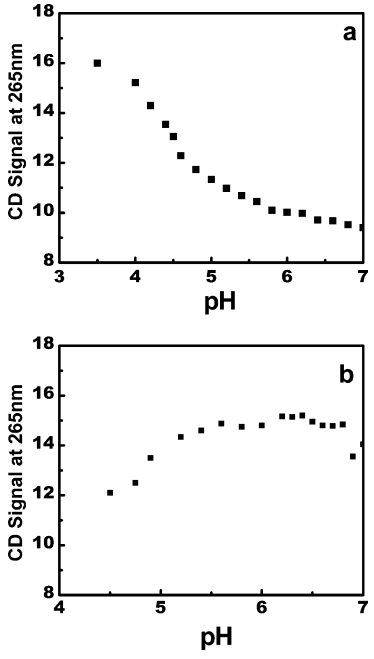


Figure 2. CD titration as a function of pH. (a) Formation of a stable i-motif in C-rich DNA and (b) disruption of structure (A-form) in C-rich RNA as a function of decreasing (acidic) pH. CD spectra were recorded in 10 mM sodium cacodylate buffer and 100 mM KCl at 25 °C.

(Figure 2b) showing that the structure is sufficiently destroyed. This clearly suggests that the secondary conformation adopted by the C-rich RNA could not be an i-motif. Additionally, CD melting curves (acidic pH) confirmed the pH dependence of the secondary structure formed by the DNA counterpart. (Supporting Information, Figure 1).

Next, in order to confirm that the G-rich sequences could indeed form a G-quadruplex structure and could be targeted and further stabilized by the G-quadruplex selective ligand, we performed indicative and confirmatory binding experiments. In a previous study by our group, this furan based cyclic homo-oligopeptide ligand (see Supporting Information, Scheme 1 for chemical structure) has been demonstrated to be highly selective for the G-quadruplex motif, thereby reducing the transcription of c-myc by targeting the G-quadruplex in the c-myc promoter in HeLa cells (ligand 2 from our previous studies).[33] Also, several molecules from the same family have been previously reported by our group to be selective G-quadruplex binders.[34,35] To demonstrate the direct binding of this furan based cyclic homo-oligopeptide ligand with the G-quadruplex motif in the present system, CD, UV melting, and ITC experiments were performed, and binding parameters are also reported.

In the presence of equimolar amounts of ligand, increase in peak amplitude in CD spectra (Figure 3a) and thermal

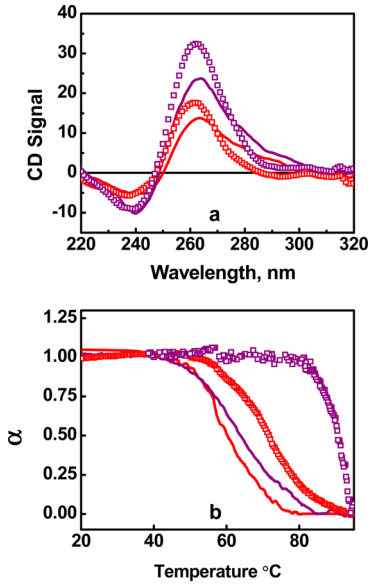


Figure 3. Structure and stability of the oligonucleotides upon interaction with the G-quadruplex ligand. (a) CD spectra and (b) UV melting traces of G-rich DNA (lines) and RNA (squares) oligonucleotides in the absence (red) and presence of an equimolar concentration of ligand 2 (purple) show increased stabilization of the G-quadruplex structure. All experiments were performed in 10 mM sodium cacodylate containing 100 mM KCl buffer at pH 7.

stabilization (ΔT_m = 4 °C) were recorded for G-rich DNA, indicating quadruplex stabilization as a result of ligand binding (Figure 3b). Notably, in the case of G-rich RNA, a complete melting domain could not be observed in the UV melting profile across the 25−95 °C range, indicating quadruplex stabilization as a result of ligand binding. Further, we corroborated the direct binding of the G-quadruplex with ligand in a thermodynamic binding study using ITC. A strong binding in the order of 3.3 (±0.3) × 10^5 M^{-1} and 3.9 (±0.4) ×

dx.doi.org/10.1021/bi401451q | Biochemistry 2014, 53, 3711−3718

10^6 M^{-1} was recorded with DNA and RNA G-quadruplex, respectively, at a 1:1 G-quadruplex to ligand stoichiometry. The binding enthalpies in these two cases were −3.2 (±0.2) and −3.0 (±0.3) kcal/mol. The binding isotherms and the thermodynamic profile of the above binding are shown in Figure 4. Also, in Supporting Information, Figure 2, we show

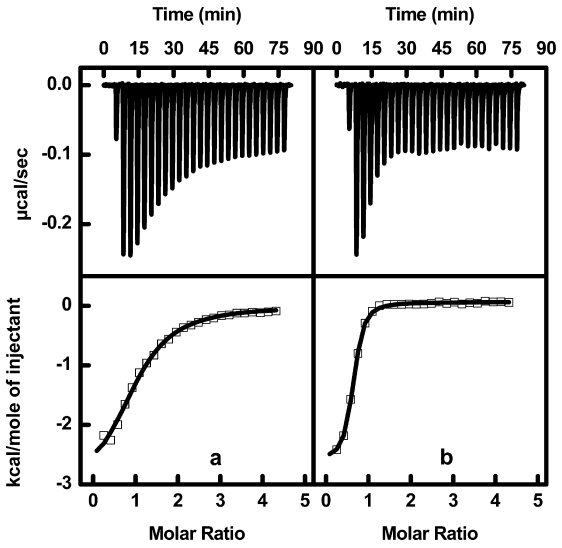


Figure 4. Direct binding of ligand with the G-quadruplex: ITC thermograms showing the binding of the furan based cyclic homo-oligopeptide ligand with a G-quadruplex motif in DNA (a) and RNA (b) sequences. The experiment was performed in 10 mM sodium cacodylate buffer containing 100 mM KCl at 25 °C. Heat of dilution of the ligand was correspondingly subtracted.

the ITC binding of the ligand with a mutated DNA counterpart of the G-rich DNA with single base substitutions to disrupt G-quartets. The thermodynamic profile indicated much weaker interactions of the order of 8.0 (±0.5) × 10^4 M^{-1}, hence suggesting selectivity for G-quadruplex conformation during the binding event.

We next performed molecular level studies based on an in-cellulo system to gain a closer look at the intracellular events. In order to assess if the expression of a gene could change owing to the presence of a G-quadruplex structure right upstream of the start site, dual reporter constructs were used in which the G-quadruplex motif was placed just upstream of the *Renilla* gene either at the sense (plasmid G) or the antisense strand (plasmid C), whereas the firefly gene was unchanged. Firefly expression was hence used as normalizing control (Scheme 2). First, we employed quantitative real time PCR to better understand the gene expression changes from the transcriptional standpoint. This technique of assessing transcript levels under the influence of the G-quadruplex is well established and has been used previously in several studies.[7,13−15,33]As shown in Figure 5, with the sense strand G-quadruplex (construct G), no significant change in the levels of *Renilla* transcripts could be observed, but significant reduction of ∼50% in transcript levels in the case of the antisense G-quadruplex (construct C) was seen. Additionally, constructs bearing quartet destabilizing base mutations in the inserted G-quadruplex motif placed either at the sense strand (GM) or at the antisense strand (CM) restored expression similar to that of the empty vector. This provided a clear indication that the change in transcript levels, if any, can be attributed to G-quadruplex mediated interference

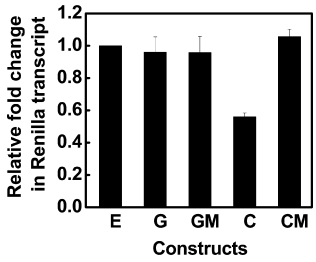


Figure 5. Relative fold change in *Renilla* transcript to measure the effect of the G-quadruplex placed at both strands individually on transcriptional efficiency. G-Quadruplex at the sense strand (G) and antisense strand (C) shows different effects on transcript levels. The mutations in the quadruplex sequence (GM and CM) restore transcription similar to that in the empty psi-check-2 vector (E). Error bars show standard deviation from three individual experiments.

with the transcriptional machinery and more so in a strand orientation specific manner.

To further validate this idea, we also performed similar experiments in the presence of a G-quadruplex selective ligand. In the present model system, it was observed that despite treatment with ligand at 100 μM concentration, there was no sensitivity to transcription in the case of any plasmid except for the antisense strand G-quadruplex (plasmid C). Incubation of plasmid C transfected cells in media containing increasing concentrations of ligand in the order of 0, 50 μM, and 100 μM slashed down relative transcript levels to 55%, 45%, and 25%, respectively, as compared to that in the empty vector (Figure 6). It was inferred that the ligand could effectively diminish the

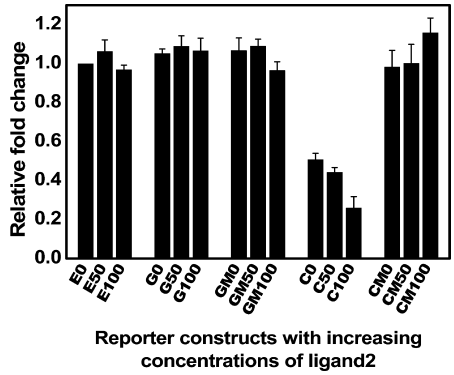


Figure 6. Relative fold change in the *Renilla* transcript to measure the effect of different concentrations of the G-quadruplex selective ligand (0, 50, and 100 μM) on transcriptional efficiency mediated by the G-quadruplex. The ligand shows no effect when the G-quadruplex was placed at the sense strand (G) and a concentration-dependent effect with the G-quadruplex at the antisense strand (C). The mutations in the quadruplex sequence (GM and CM) restored transcription similar to that in the empty psi-check-2 vector (E). Error bars show standard deviation from three individual experiments.

transcription of *Renilla* luciferase gene in a concentration dependent manner only in the case of the antisense G-quadruplex, whereas transcription in the sense strand G-quadruplex (G) along with the mutated motif plasmids (GM and CM) remained unaffected at even the highest concentration of ligand. This evasion from ligand mediated effect suggested two things: one that the effect was most probably governed by the structure because in absence of the potential motif, even the highest concentration of the selective G-

dx.doi.org/10.1021/bi401451q | *Biochemistry* 2014, 53, 3711−3718

quadruplex stabilizing ligand was ineffective in inducing the effect. Second and most important, the effect was quite sensitive to the orientation of the strand which harbors the G-quadruplex motif, bringing to light the concept of strand asymmetry in the context of G-quadruplex mediated gene regulation.

Further, as a proof of concept we extended the study to investigate the influence of the G-quadruplex motif ultimately at the level of translation regulation. Hence, we utilized a dual luciferase assay to determine the translation efficiency of the constructs used in the study. Relative luminescence levels were measured, and the results are shown in Figure 7. It was

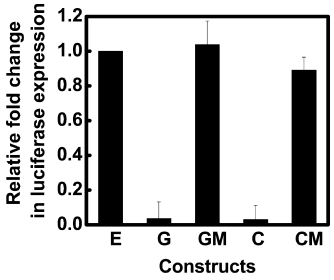


Figure 7. Relative fold change in luciferase expression under the G-quadruplex placed at both strands individually to measure the effect on translation efficiency. G-Quadruplex at the sense strand (G) and antisense strand (C) both show similar reduction in luciferase expression levels. The mutations in the quadruplex sequence (GM and CM) restore transcription similar to that in the empty psi-check-2 vector (E). Error bars show standard deviation from three individual experiments.

observed that the relative expression of the *Renilla* reporter was significantly reduced to only about 10% as compared to the empty vector, in both the cases when the G-quadruplex motif was located at the sense (construct G) and the antisense strands (construct C). Although a similar observation in the case of the sense strand G-quadruplex has been documented previously,[10,12] diminished luminescence with the antisense G-quadruplex, however, was rather surprising. Additionally, the plasmids containing the mutated G-quadruplex motif at the sense strand (GM) and antisense strand (CM) showed a dramatic return to luminescence levels similar to those of the empty vector. It is worth mentioning here that both the GM and CM sequences do not show any melting domain in the UV melting curves (Supporting Information, Figure 3a). Moreover, they do not show the strong CD signature too (Supporting Information Figure 3b). This rescue of luminescence due to the disruption of the G-quadruplex is a significant contribution to the pool of proofs that the effect was essentially due to the G-quadruplex motif.

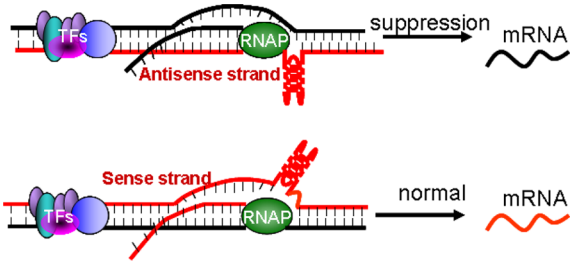
From this, it is inferred that the presence of the G-quadruplex at the sense strand is effective in decreasing the translation efficiency, while at the antisense strand, it is liable to exert a greater control over the ultimate expression of protein by affecting both transcription and translation processes.

DISCUSSION

Strand bias in the functional context of G-quadruplex mediated altered transcription is a novel finding. Although numerous studies until date have reported G-quadruplex controlled altered gene expression, they have not yet focused on the element of strand asymmetry. Nonetheless, it is interesting to discuss here the mechanism of this asymmetric behavior. We suggest that the origin of this strand asymmetry is most probably due to the physical blockage to the transcription machinery posed by the G-quadruplex fold. The RNAPII machinery which reads the template/antisense strand in the 3′ to 5′ orientation is faced with a physical barrier when it encounters the G-quadruplex at the template/antisense strand.[36] This hampers the reading and proper templating for transcription leading to reduced transcript levels in turn. Conversely, if the template strand is devoid of any such barrier, transcription remains unaltered, whether or not the G-quadruplex was present at the complementary (sense or nontemplate) strand, thereby making the presence of the G-quadruplex at the sense or nontemplate strand irrelevant to efficiency at the transcription level. Tornaletti and colleagues have demonstrated G-quadruplex mediated arrest and inhibition of RNAPII in two different studies.[37,38] However, they have used single stranded substrates for RNAPII, suggesting similar but somewhat incomplete observations. Whereas by systematically constructing double stranded plasmids with the G-quadruplex sequence at the same locus in each of the two strands and comparing the results from reporter assays, we provide more direct evidence. A similar system has been used in the context of TCR (transcription coupled repair) to show the blocking of RNAPII upon encounter with 8-oxoguanine.[39] Authors have shown using a run-off assay that RNAPII is efficiently blocked when encountering a lesion on the antisense strand, whereas the sense strand lesion is easily bypassed.

Additionally, as for translation efficiency, it is presumable that with fewer transcripts on account of diminished transcription, translation (in the present model system) may also be compromised. However, in the case of the sense strand G-quadruplex, translation may have been decreased as a result of the formation of the G-quadruplex structure in the 5′ UTR of the resulting transcript. It represents the case of the inhibitory role of the RNA G-quadruplex on translation surfaces once again. No other study in the past has reported the cumulative effect of strand orientation based influence of the G-quadruplex on translation, which seems to be similarly diminished in either case, but the mechanism seems absolutely different. The events may be summarized as depicted in Scheme 3.

Scheme 3. Result of Strand Asymmetry in G-Quadruplex Mediated Transcription



In this context, it is also noteworthy that the G-quadruplex motif on one (G-rich) strand may be accompanied by the i-motif (although only at acidic pH) on the complementary C-rich strand. In the case represented above, with a sense strand G-quadruplex, the antisense strand could have an i-motif, which being a secondary structure may also pose an impediment to RNAPII reading. However, in the absence of any effect on transcription with a sense strand G-quadruplex, it is presumable that the i-motif structure (on the complementary strand) was

probably not encountered intracellularly. Conversely, if at all the i-motif was encountered inside the cell, the outcome on transcription due to physical block would have been similar in both cases, not withstanding the presence of the G-quadruplex at the sense or antisense strand. However, in the absence of the same it may be inferred that most probably the converse strand was either devoid of any secondary structure (in this case the i-motif), or the structure was unable to manifest a measurable effect. Hence, we provide circumstantial evidence that most probably the i-motif does not form in vivo or at least does not alter transcription like G-quadruplex does.

CONCLUSIONS

In the present study, we have endeavored to understand and explain the role of G-quadruplex structures in altering transcription. On the foreground, we have shown that G-quadruplexes exhibit precision in their biological role of regulation of gene expression by identifying strand orientation. The present study unravels an important factor of strand asymmetric behavior in their biological function. G-quadruplexes in the sense strand do not interfere with transcription. While in the antisense strand, G-quadruplexes create a physical blockage for RNAPII and thus reduce transcription (Scheme 1). In summary, the G-quadruplex in the sense strand affects only translation, whereas in the antisense stand, they can affect transcription as well as translation. Lastly, we also suggest that the i-motif most probably does not form in vivo, though it needs further investigation.

An interesting and surprisingly similar context may be drawn from the case of the 27 mer G-quadruplex from the human c-myc promoter (P1) which happens to be located on the template/noncoding/antisense strand in the genome (Ensembl genome browser). It is well demonstrated that this G-quadruplex is also sufficiently capable of down-regulating transcription,[7,33] whereas RNA G-quadruplexes like those reported in Zic1 or MT3-MMP[13,14] originate from the sense strand in DNA but do not exhibit transcriptional down-regulation but only suppress translation. However, we caution that apart from the above examples, it may not be appropriate to generalize here that all functional DNA G-quadruplexes may be curated to the template strand of the corresponding gene without further experimentation and validation. Overall, G-quadruplexes may function as sophisticated tools to fine tune gene regulation at the level of transcription or translation based on its strand position.

ASSOCIATED CONTENT

Supporting Information

Structure of the G-quadruplex binding ligand; changes in CD signal as a function of temperature; ITC thermogram; and UV melting and CD curves. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Authors

*(T.K.C.) Phone: +91-80- 2293 2215. Fax: +91-80-2360-0529. E-mail: tushar@orgchem.iisc.ernet.in.

*(S.M.) Phone: +91-11-2987-9108. Fax: +9127667471. E-mail: souvik@igib.res.in.

Funding

This work was financially supported by Department of Science and Technology (DST) (Swarnajayanti project to S.M.), Government of India. The fellowships from Council of Scientific and Industrial Research (CSIR), Government of India (to T.A and S.R), are gratefully acknowledged.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Frank-Kamenetskii, M. (1989) DNA structure. The turn of the quadruplex? *Nature 342*, 737.

(2) Weisman-Shomer, P., and Fry, M. (1993) QUAD, a protein from hepatocyte chromatin that binds selectively to guanine-rich quadruplex DNA. *J. Biol. Chem. 268*, 3306−3312.

(3) Sanders, C. M. (2010) Human Pif1 helicase is a G-quadruplex DNA-binding protein with G-quadruplex DNA-unwinding activity. *Biochem. J. 430*, 119−128.

(4) Zaug, A. J., Podell, E. R., and Cech, T. R. (2005) Human POT1 disrupts telomeric G-quadruplexes allowing telomerase extension in vitro. *Proc. Natl. Acad. Sci. U.S.A. 102*, 10864−10869.

(5) Takahama, K., Kino, K., Arai, S., Kurokawa, R., and Oyoshi, T. (2011) Identification of Ewing's sarcoma protein as a G-quadruplex DNA- and RNA-binding protein. *FEBS J. 278*, 988−998.

(6) Simonsson, T. (2001) G-quadruplex DNA structures–variations on a theme. *Biol. Chem. 382*, 621−628.

(7) 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.

(8) Gunaratnam, M., Swank, S., Haider, S. M., Galesa, K., Reszka, A. P., Beltran, M., Cuenca, F., Fletcher, J. A., and Neidle, S. (2009) Targeting human gastrointestinal stromal tumor cells with a quadruplex-binding small molecule. *J. Med. Chem. 52*, 3774−3783.

(9) 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.

(10) Wang, X. D., Ou, T. M., Lu, Y. J., Li, Z., Xu, Z., Xi, C., Tan, J. H., Huang, S. L., An, L. K., Li, D., Gu, L. Q., and Huang, Z. S. (2010) Turning off transcription of the bcl-2 gene by stabilizing the bcl-2 promoter quadruplex with quindoline derivatives. *J. Med. Chem. 53*, 4390−4398.

(11) Wieland, M., and Hartig, J. S. (2007) RNA quadruplex-based modulation of gene expression. *Chem. Biol. 14*, 757−763.

(12) Kumari, S., Bugaut, A., Huppert, J. L., and Balasubramanian, S. (2007) An RNA G-quadruplex in the 5′ UTR of the NRAS proto-oncogene modulates translation. *Nat. Chem. Biol. 3*, 218−221.

(13) Arora, A., Dutkiewicz, M., Scaria, V., Hariharan, M., Maiti, S., and Kurreck, J. (2008) Inhibition of translation in living eukaryotic cells by an RNA G-quadruplex motif. *RNA 14*, 1290−1296.

(14) Morris, M. J., and Basu, S. (2009) An unusually stable G-quadruplex within the 5′-UTR of the MT3 matrix metalloproteinase mRNA represses translation in eukaryotic cells. *Biochemistry 48*, 5313−5319.

(15) Halder, K., Wieland, M., and Hartig, J. S. (2009) Predictable suppression of gene expression by 5′-UTR-based RNA quadruplexes. *Nucleic Acids Res. 37*, 6811−6817.

(16) Arora, A., and Maiti, S. (2009) Stability and molecular recognition of quadruplexes with different loop length in the absence and presence of molecular crowding agents. *J. Phys. Chem. B 113*, 8784−8792.

(17) Guédin, A., Gros, J., Alberti, P., and Mergny, J. L. (2010) How long is too long? Effects of loop size on G-quadruplex stability. *Nucleic Acids Res. 38*, 7858−7868.

(18) Arora, A., Nair, D. R., and Maiti, S. (2009) Effect of flanking bases on quadruplex stability and Watson-Crick duplex competition. *FEBS J. 276*, 3628−3640.

(19) Marathias, V. M., and Bolton, P. H. (1999) Determinants of DNA quadruplex structural type: sequence and potassium binding. *Biochemistry 38*, 4355−4364.
(20) Schultze, P., Hud, N. V., Smith, F. W., and Feigon, J. (1999) The effect of sodium, potassium and ammonium ions on the conformation of the dimeric quadruplex formed by the Oxytricha nova telomere repeat oligonucleotide d(G(4)T(4)G(4). *Nucleic Acids Res. 27*, 3018−3028.
(21) Monchaud, D., and Teulade-Fichou, M. P. (2008) A hitchhiker's guide to G-quadruplex ligands. *Org. Biomol. Chem. 6*, 627−636.
(22) Phan, A. T., Kuryavyi, V., Gaw, H. Y., and Patel, D. J. (2005) Small-molecule interaction with a five-guanine-tract G-quadruplex structure from the human MYC promoter. *Nat. Chem. Biol. 1*, 167−173 Erratum in *Nat. Chem. Biol. 1*, 234.
(23) Ou, T. M., Lu, Y. J., Zhang, C., Huang, Z. S., Wang, X. D., Tan, J. H., Chen, Y., Ma, D. L., Wong, K. Y., Tang, J. C., Chan, A. S., and Gu, L. Q. (2007) Stabilization of G-quadruplex DNA and down-regulation of oncogene c-myc by quindoline derivatives. *J. Med. Chem. 50*, 1465−1474.
(24) Tera, M., Ishizuka, H., Takagi, M., Suganuma, M., Shin-ya, K., and Nagasawa, K. (2008) Macrocyclic hexaoxazoles as sequence- and mode-selective G-quadruplex binders. *Angew. Chem., Int. Ed. 47*, 5557−5560.
(25) Han, H., Langley, D. R., Rangan, A., and Hurley, L. H. (2001) Selective interactions of cationic porphyrins with G-quadruplex structures. *J. Am. Chem. Soc. 123*, 8902−8913.
(26) Kim, M. Y., Gleason-Guzman, M., Izbicka, E., Nishioka, D., and Hurley, L. H. (2010) The different biological effects of telomestatin and TMPyP4 can be attributed to their selectivity for interaction with intramolecular or intermolecular G-quadruplex structures. *Cancer. Res. 63*, 3247−3256.
(27) Miyoshi, D., Karimata, H., and Sugimoto, N. (2006) Factors regulating thermodynamic stability of DNA structures under molecular crowding conditions. *Nucleic Acids Symp. Ser. 50*, 203−204.
(28) Zhang, D. H., Fujimoto, T., Saxena, S., Yu, H. Q., Miyoshi, D., and Sugimoto, N. (2010) Monomorphic RNA G-quadruplex and polymorphic DNA G-quadruplex structures responding to cellular environmental factors. *Biochemistry 49*, 4554−4563.
(29) Heddi, B., and Phan, A. T. (2011) Structure of human telomeric DNA in crowded solution. *J. Am. Chem. Soc. 133*, 9824−9833.
(30) Du, Z., Zhao, Y., and Li, N. (2008) Genome-wide analysis reveals regulatory role of G4 DNA in gene transcription. *Genome Res. 18*, 233−241.
(31) Lacroix, L., Mergny, J. L., Leroy, J. L., and Hélène, C. (1996) Inability of RNA to form the i-motif: implications for triplex formation. *Biochemistry 35*, 8715−8722.
(32) Guéron, M., and Leroy, J. L. (2000) The i-motif in nucleic acids. *Curr. Opin. Struct. Biol. 10*, 326−331.
(33) Agarwal, T., Roy, S., Chakraborty, T. K., and Maiti, S. (2010) Selective targeting of G-quadruplex using furan-based cyclic homooligopeptides: effect on c-MYC expression. *Biochemistry 49*, 8388−8397.
(34) Chakraborty, T. K., Arora, A., Roy, S., Kumar, N., and Maiti, S. (2007) Furan based cyclic oligopeptides selectively target G-quadruplex. *J. Med. Chem. 50*, 5539−5542.
(35) Agarwal, T., Roy, S., Chakraborty, T. K., and Maiti, S. (2010) Furan based cyclic homo-oligopeptides bind G-quadruplex selectively and repress c-MYC transcription. *Bioorg. Med. Chem. Lett. 20*, 4346−4349.
(36) Lodish, H., Berk, A., Matsudaira, P., Kaiser, C. A., Kreiger, M., Scott, M., Zipursky, S. L., and Darnell, J. (1987) *Molecular Cell Biology*, 5th ed., p 110, W. H. Freeman and Company, New York.
(37) Broxson, C., Beckett, J., and Tornaletti, S. (2011) Transcription arrest by a G quadruplex forming-trinucleotide repeat sequence from the human c-myb gene. *Biochemistry 50*, 4162−4172.
(38) Tornaletti, S., Park-Snyder, S., and Hanawalt, P. C. (2008) G4-forming sequences in the non-transcribed DNA strand pose blocks to T7 RNA polymerase and mammalian RNA polymerase II. *J. Biol. Chem. 283*, 12756−12762.
(39) Tornaletti, S., Maeda, L. S., Kolodner, R. D., and Hanawalt, P. C. (2004) Effect of 8-oxoguanine on transcription elongation by T7 RNA polymerase and mammalian RNA polymerase II. *DNA Repair 3*, 483−494.