ACS chemical biology

pubs.acs.org/acschemicalbiology

Articles

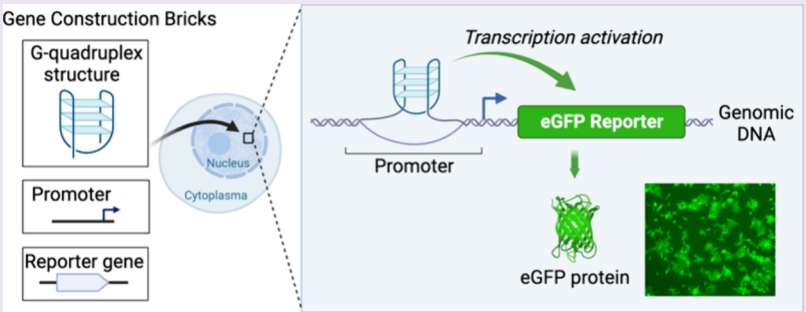
An Upstream G-Quadruplex DNA Structure Can Stimulate Gene Transcription

Yuqi Chen, Angela Simeone, Larry Melidis, Sergio Martinez Cuesta, David Tannahill, and Shankar Balasubramanian*

Cite This: *ACS Chem. Biol.* 2024, 19, 736–742

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



ABSTRACT: Four-stranded G-quadruplexes (G4s) are DNA secondary structures that can form in the human genome. G4 structures have been detected in gene promoters and are associated with transcriptionally active chromatin and the recruitment of transcription factors and chromatin remodelers. We adopted a controlled, synthetic biology approach to understand how G4s can influence transcription. We stably integrated G4-forming sequences into the promoter of a synthetic reporter gene and inserted these into the genome of human cells. The integrated G4 sequences were shown to fold into a G4 structure within a cellular genomic context. We demonstrate that G4 structure formation within a gene promoter stimulates transcription compared to the corresponding G4-negative control promoter in a way that is not dependent on primary sequence or inherent G-richness. Systematic variation in the stability of folded G4s showed that in this system, transcriptional levels increased with higher stability of the G4 structure. By creating and manipulating a chromosomally integrated synthetic promoter, we have shown that G4 structure formation in a defined gene promoter can cause gene transcription to increase, which aligns with earlier observational correlations reported in the literature linking G4s to active transcription.

INTRODUCTION

Certain guanine-rich DNA sequences can fold into intramolecular fold-back four-stranded secondary structures called G-quadruplexes (G4s).[1] G4s comprise stacked, Hoogsteen hydrogen-bonding G-tetrads that can be stabilized by a centrally coordinated cation (Figure 1a), with connecting loops of varying lengths and sequences that can be arranged in different orientations. Intramolecular G4s can be thermally stable under conditions of physiological salt, pH, and temperature. Computational prediction[2,3] and a G4-sensitive DNA sequencing approach[4] have suggested the potential for folded G4 structures to form at hundreds of thousands of sites in the human genome. Methods developed to detect G4 structures in cellular chromatin have observed only thousands of G4 structures in human cells,[5−9] suggesting that chromatin context suppresses the folding of most G4 structures. Computational prediction, sequencing experiments, and chromatin mapping approaches all show the enrichment of G4s immediately before the transcription start site (TSS) of many protein coding genes in gene regulatory regions called promoters.[3,10,11]

Multiple lines of evidence highlight promoter G4s as having biological importance for transcription. In vitro, G4s forming at a transcribed gene can stall RNA polymerase transit.[12,13] Cellular studies using transfected plasmid gene constructs suggest that natural G4 motifs from oncogene promoters, such as MYC, KRAS, and KIT, can modulate transcription of a downstream reporter gene, as compared to mutated G4 motifs.[14−18] Transcription of VEGF and KRAS oncogenes is also stimulated by promoter G4 folding arising from oxidative damage to DNA bases in promoters.[19,20] The addition of G4-stabilizing small molecules to cells can alter expression of

Received: December 15, 2023
Revised: February 12, 2024
Accepted: February 13, 2024
Published: February 28, 2024

https://doi.org/10.1021/acschembio.3c00775
ACS Chem. Biol. 2024, 19, 736–742

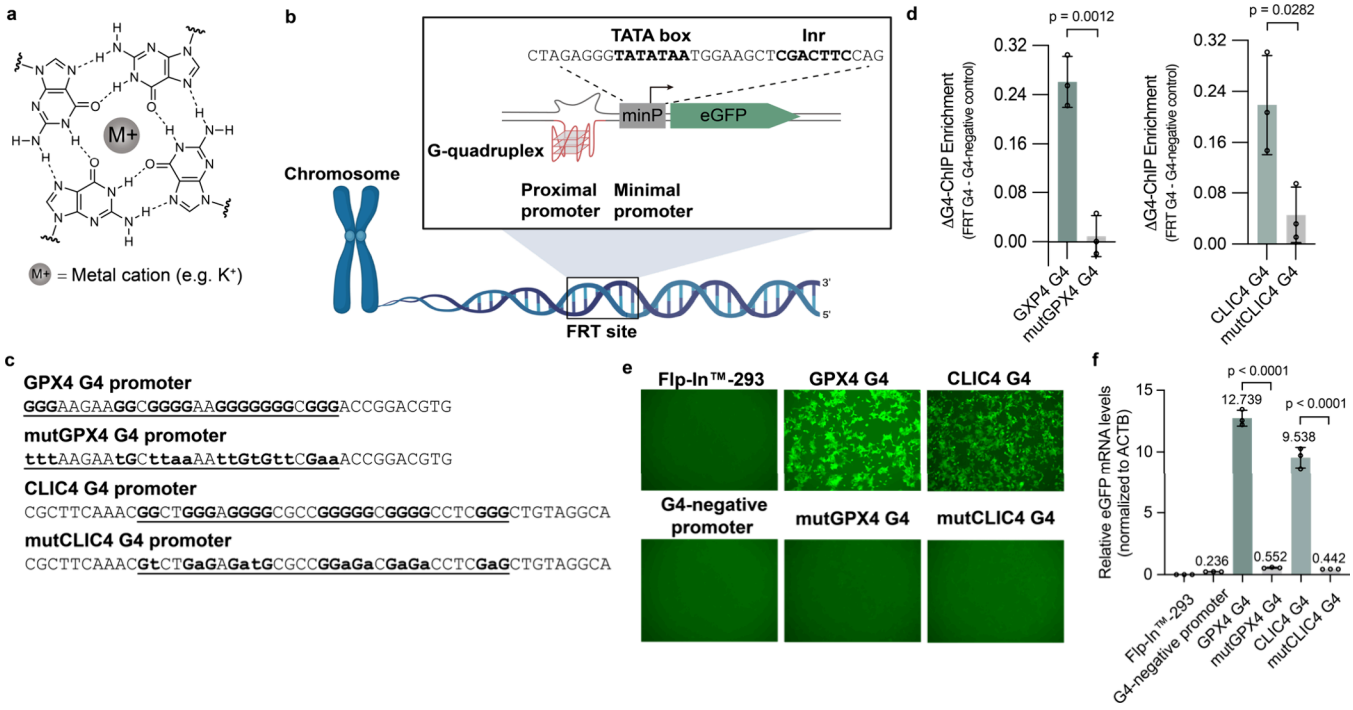


Figure 1. GPX4 and CLIC4 promoter sequences fold into G4 structures in cells and promote eGFP transcription. (a) G-tetrad stabilized by Hoogsteen base pairing and a monovalent cation. (b) Architecture of the G4 promoter eGFP reporter system. A G4 is placed at the proximal promoter region on the template strand. A minP sequence containing a TATA box and Inr acts as a core promoter to mediate transcription initiation and is placed between the G4 and eGFP protein coding sequences. The sequences of TATA box and Inr are in bold. Panel (b) was created with BioRender.com. (c) Sequences of GPX4 G4/mutated GPX4 G4 and CLIC4 G4/mutated CLIC4 G4. Point mutations for mutated G4s are indicated by lowercase letters. (d) G4-ChIP-qPCR analysis of G4 formation at the CLIC4 and GPX4 G4 promoters normalized to a positive control G4 in the host genome (RBBP4) and after background signal subtraction. (e) Representative fluorescence microscopy images of Flp-In expression cell lines showing that promoter GPX4 and CLIC4 G4s drive eGFP protein expression. (f) Quantification of eGFP RNA expression by RT-qPCR. G4-negative promoter: the Flp-In-293 expression cell line in which G4 is absent from the eGFP promoter. Flp-In-293: the parental Flp-In-293 cell line without the integration of the eGFP reporter (mean ± s.d.; two-tailed unpaired *t* test).

oncogenes, such as SRC, BCL2, and KIT, and also induce DNA damage to promote cell killing.[21−23] Detection of folded G4 structures with structure-specific probes shows that folded G4s are consistently detected within accessible chromatin of human cells and their incidence correlates with active gene transcription.[5,24−26] Furthermore, the differentiation of human stem cells into defined lineages revealed that the dynamic alteration of where folded G4 structures were retained, gained, or lost had a positive correlation with transcriptional activity, active histone marks, and open chromatin.[24] Proteins involved in the regulation of gene expression, including several transcription factors, can directly bind G4 structures with high affinity.[26−30] Collectively, such studies lend support for the formation of promoter G4s as positive regulators of transcription.

There are limitations in the existing data that link G4s to transcription. Interpretations rely largely on correlations. Many studies did not include the actual detection of folded G4s and rely on inference from the G4 sequence motifs. Plasmid-based studies do not position the G4 structure within the native locus in a natural chromatin context. Studies from our own laboratory and other laboratories have shown that G4-targeting small molecules can lower transcription at genes with G4 motifs in promoters. A possible interpretation of such studies is that stabilization of G4s inhibits transcription, and therefore, G4s are negative regulators of transcription. There are several reasons to be cautious about interpreting such experiments. Small molecules can bind to thousands of folded G4 targets in the cellular genome with many potential downstream effects. Small molecules can compete off G4-bound proteins (e.g., transcription factors),[28] and so a small molecule-G4 complex is not just a more stable G4. Last, many small molecules that target G4s, for example, pyridostatin, cause proximal DNA strand cleavage and a downstream cellular DNA damage response,[22] which confounds a simple interpretation of transcriptional changes in a single gene. Thus, small molecule experiments cause complex effects that can preclude a clear and reliable interpretation. There is a need for studies that can more clearly and directly resolve the relationship between a promoter G4 and transcription.

Herein, we describe the design and assembly of a synthetic promoter-reporter gene regulatory system that was exploited to study the consequential effects of a promoter-G4 on transcription. The observations show that the formation of G4 structures in this context positively regulates gene transcription, with transcription levels being related to the intrinsic thermal stability of the G4 structure.

RESULTS AND DISCUSSION

Experimental Design. We developed a synthetic gene regulatory construct to directly evaluate the relationship between G4s in the promoter and transcription (Figure 1b,c). Using the Flp-In-293 human cell line, we placed G4 sequences, which are in front of a minimal core promoter (minP) construct that drives the expression of a green fluorescent protein (eGFP) reporter (see Supplementary

https://doi.org/10.1021/acschembio.3c00775
ACS Chem. Biol. 2024, 19, 736−742

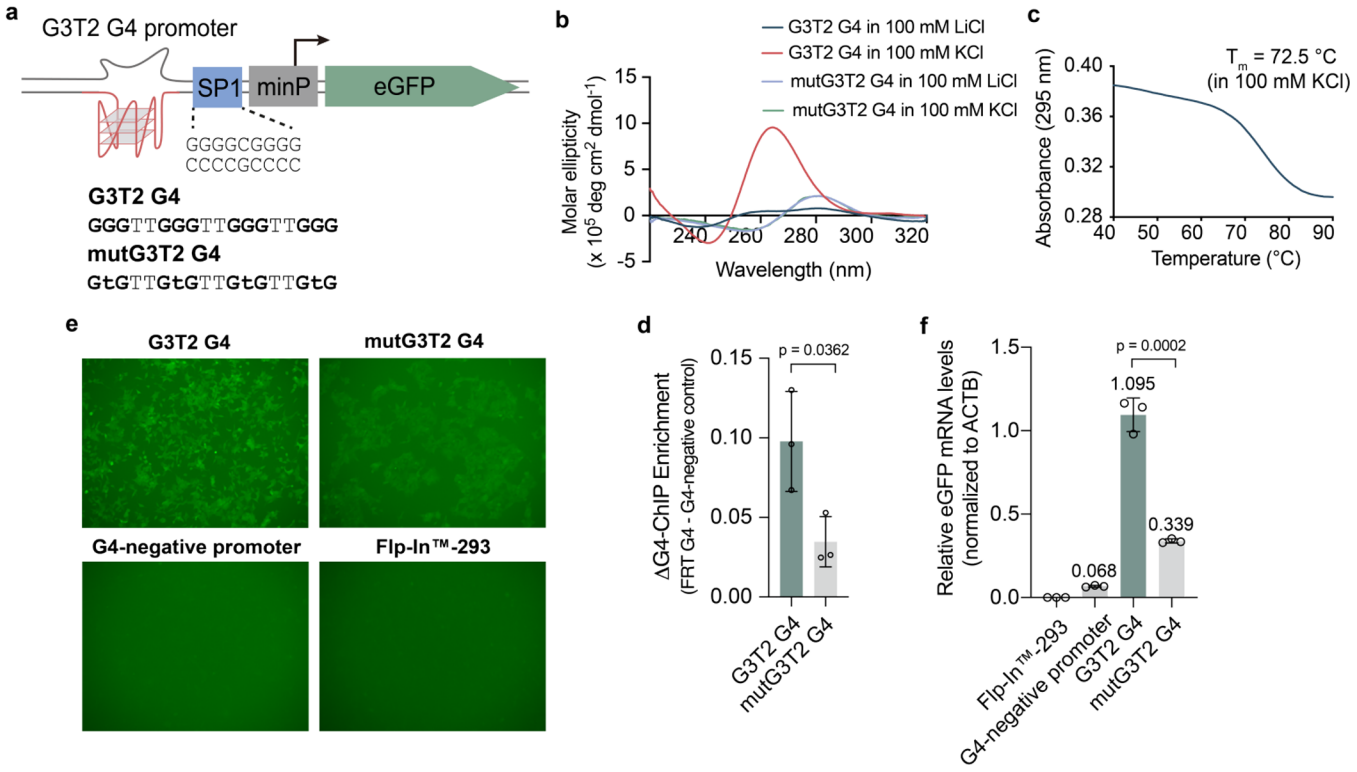


Figure 2. A G4 structure lacking SP1 consensus sequences still promotes enhanced eGFP expression. (a) Architecture of the eGFP reporter system. A synthetic $(G3T2)_3GGG$ G4 or a mutated version (termed G3T2 or mutG3T2, respectively) is placed in the proximal promoter. A SP1 binding site is also placed between the proximal and core promoters. (b, c) Biophysical characterization of the G3T2 G4 structure. Left, circular dichroism spectroscopy for G3T2 or mutG3T2 G4 oligonucleotides showing spectra characteristic of a G4 structure for G3T2 but not mutG3T2 (buffer conditions: 10 mM Tris−HCl (pH 7.4) with 100 mM KCl or LiCl; oligonucleotide concentration: 10.0 μM). Right, UV-melting curve for the G3T2 G4 oligonucleotide. T_m is indicated on the graph (buffer conditions: 10 mM Tris−HCl (pH 7.4) with 100 mM KCl; oligonucleotide concentration: 5.0 μM). (d) G4-ChIP-qPCR quantification of G4 formation for G3T2, as described in Figure 1, compared to mutG3T2 cells. (e) Representative fluorescence microscopy images showing that the G3T2 G4 has elevated eGFP protein expression compared to mutG3T2, G4-negative, and host cell line expression. (f) Quantification of the increase in eGFP transcription for G3T2 cells compared to mutG3T2 cells by RT-qPCR as described in Figure 1 (mean ± s.d.; two-tailed unpaired t test).

Methods) into human chromatin. The G4 sequences were always positioned on the template strand. Cell lines were generated by integrating each construct into the same genomic position by site-directed recombination (Supplementary Methods) and confirmed to have a single copy of the correct sequence by Sanger DNA sequencing (Supplementary Methods, Supporting Information).

G4 folding of test sequences was detected and quantified using a folded-G4 structure-specific antibody BG4[25] to affinity-capture folded G4 structures from isolated chromatin,[5] followed by quantitative polymerase chain reaction (G4 ChIP-qPCR) using primers adjacent to the G4 site. Fluorescence microscopy or flow cytometry (see Supplementary Methods) of live cells by detecting eGFP fluorescence was used to estimate expression levels. Accurate quantification of transcription was achieved from isolated RNA by quantitative reverse transcription qPCR (RT-qPCR) with gene-specific primers for eGFP using actin expression as a reference.

Selection of Natural Promoter G4s for Insertion. Experimental mapping of folded G4s in human cell lines[31] shows that promoter G4s are enriched in front of the transcription start sites of actively transcribing genes (see Supplementary Methods and Figure S1). Glutathione peroxidase 4 (GPX4) and chloride intracellular channel 4 (CLIC4) G4s were selected as representative single G4s present in the proximal promoters of genes undergoing high transcription (see Supplementary Methods and Figure S2). Circular dichroism (CD) spectroscopy[32,33] confirmed that GPX4 and CLIC4 G4 oligonucleotides fold into structures consistent with parallel (~265 nm maxima and ~245 nm minima) or hybrid (~295 nm, ~265 nm maxima and ~245 nm minima) G4 conformations, respectively (Figure S3a,c).[34−36] Folded G4 formation was disrupted either when tetrad-forming G bases were mutated or when folded in 100 mM LiCl. UV thermal melting spectroscopy[37,38] indicating that GPX4 and CLIC4 G4s have high thermal stability with melting temperatures (T_m) of 73.5 and 81.0 °C, respectively (Figure S3b,d).

Promoter G4s Stimulate Transcription. We assessed whether the insertion of GPX4 or CLIC4 G4 sequences, or their mutated variants, causes folded G4 structures in our cell lines as determined by G4 ChIP-qPCR. The RBBP4 G4, situated elsewhere in the genome, was used as a positive reference for G4 folding, and background correction was performed by subtracting the signal detected from cells carrying only the minimal promoter without the G4 insertion. Cells with the GPX4 or CLIC4 G4 sequence exhibited a detectable, folded G4 structure, which was lost upon mutation of critical G bases (Figure 1d). We have therefore created detectable folded G4 structures at specific sites by inserting G4 sequences into the cellular human genome.

We then investigated whether G4 formation was linked to active transcription. For this, we quantified eGFP expression

https://doi.org/10.1021/acschembio.3c00775
ACS Chem. Biol. 2024, 19, 736−742

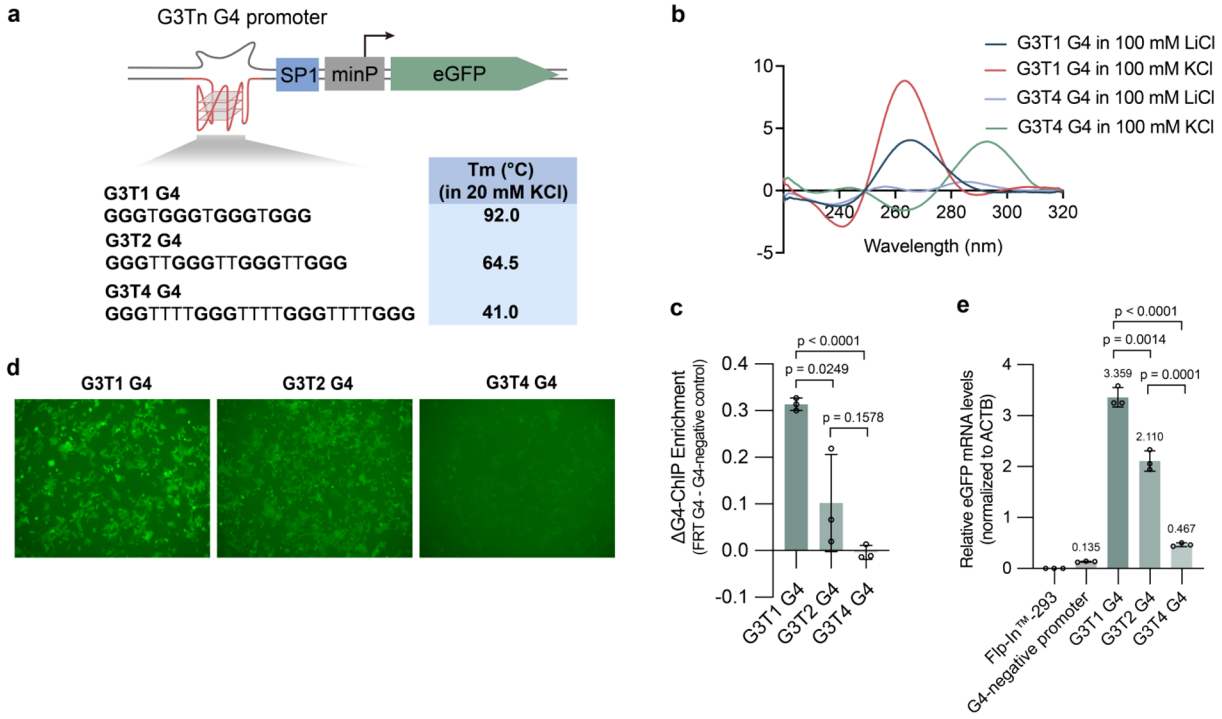


Figure 3. Increased transcriptional activity correlates with increased G4 structural stability in promoters. (a) Sequences of the synthetic G4s G3T1, G3T2, and G3T4 and the corresponding decrease in thermal stability as assessed by oligonucleotide UV melting. (b) Circular dichroism spectra confirming G4 folding for G3T1 and G3T4 in 100 mM KCl or LiCl as per Figure 2 (data for G3T2 is in Figure 2b). (c) Confirmation that increasing G4 stability leads to increased G4 formation in cells, as assessed by G4-ChIP-qPCR. (d) Representative fluorescence microscopy images of the synthetic G4 cell lines showing that eGFP protein expression correlates with G4 stability (data for G3T2 is from Figure 2e). (e) Quantification of the eGFP expression for synthetic G4 cell lines by RT-qPCR showing that expression levels correlate with G4 stability (mean ± s.d.; two-tailed unpaired *t* test.).

driven from GPX4 or CLIC4 G4-containing promoters compared to mutated promoters unable to form a G4. Fluorescence microscopy of live cells revealed that cells with either a GPX4 or CLIC4 G4 promoter had easily detectable eGFP fluorescence compared to those lacking a G4 or carrying a mutated G4 (Figure 1e). With expression analysis by reverse transcription followed by quantitative polymerase chain reaction of cDNA (RT-qPCR), cells with a GPX4 or CLIC4 G4 showed more than a 40-fold increase in eGFP mRNA compared to cells with only the minimal promoter reporter lacking a G4 (Figure 1f). A similar G4-driven increase (70-fold) in expression was observed when we measured the mean fluorescence intensity (MFI) for eGFP by flow cytometry (Figure S4). Therefore, in our synthetic regulatory system, the addition of a folded G4 structure in the proximal promoter stimulates transcription. When we generated cells carrying a CLIC4 G4 but lacking the minimal core promoter, eGFP reporter expression was ~90% lower than when both components were present (Figure S5).

The combination of a minimal core promoter and G4 is thus required for maximal transcriptional output.

To rule out whether the activation of gene expression in the G4-positive cell lines is due to intrinsic duplex G-richness, we shuffled the CLIC4 G4 sequence to break up the G4 motif but preserve the overall nucleotide composition. We confirmed that the oligonucleotide for the shuffled G4 did not fold into a G4 structure, as judged by CD spectroscopy (Figure S6a,b). Insertion of the shuffled sequence into the proximal promoter in cells led to a loss of eGFP reporter expression by over 97% to basal levels compared to CLIC4 G4-containing cells (Figure S6c–e). These findings demonstrate that it is the folded G4 structure rather than intrinsic G-richness that stimulates transcription.

G4 Stimulate Transcription Independent of the SP1 Consensus Motif. Many G4 sequence motifs overlap with the consensus duplex DNA binding site for the transcription factor SP1, which makes it challenging to distinguish the potential effects of each motif.[39] We noted that the CLIC4 G4 and GPX4 G4 sequences, described earlier, also contain a SP1 consensus target sequence, which may confound our explanation. We therefore designed a second minimal core promoter system in which a fixed, separate, SP1 consensus sequence was incorporated with an unnatural G4 sequence in front to rule out effects due to simultaneously changing both the G4 and the SP1 binding site (Figure 2a).

We designed a simple G4, $(G3T2)_3GGG$ (**G3T2**), that excludes motifs for the SP1 canonical duplex binding. Affinity pull-down of the SP1 transcription factor from Flp-In-293 cell lysates using G3T2 G4 and duplex G3T2 (dsG3T2) oligonucleotides and analyzed by western blotting shows that SP1 binds strongly to a folded, but not duplex, G4. As a control, SP1 was shown to bind to G4 structures (G4Myc) and the SP1 duplex consensus (Figure S7). The **G3T2** oligonucleotide folds into a stable G4 structure as assessed by CD spectroscopy and UV thermal melting spectroscopy (T_m = 72.5 °C) (Figure 2a–c). G4 folding was also abolished in 100 mM LiCl or on mutation of the central tetrad Gs to T ($(GTGT2)_3GTG$).

We next measured G4 folding and eGFP expression in cell lines in which either the **G3T2** G4 or the mutated G4 was

https://doi.org/10.1021/acschembio.3c00775
ACS Chem. Biol. 2024, 19, 736–742

inserted into the promoter construct (Figure 2a). G4 ChIP-qPCR showed that cells carrying a **G3T2** G4 exhibit a stronger signal for folded G4 compared to cells carrying a mutated G4 motif (Figure 2d). This increase in detectable folded G4 was also accompanied by a ∼3-fold increase in eGFP RNA/protein expression (Figure 2e,f and Figure S4).

These findings confirm that a folded G4 can stimulate transcription independent of the canonical SP1 binding site.

G4 Stability Can Influence G4 Formation in Cells. We designed a series of unnatural G4s with progressively longer loop sizes of 1, 2, and 4 nucleotides, respectively, $(GGGT)_3GGG$ (**G3T1**), $(GGGTT)_3GGG$ (**G3T2**), and $(GGGTTTT)_3GGG$ (**G3T4**) (Figure 3a), designed to have progressively lower thermal stability based on biophysical studies.[40,41] G4 folding was verified for each sequence by CD spectroscopy of oligonucleotides[34−36] (Figures 2b and 3b), and UV thermal melting spectroscopy confirmed that thermal stability decreases with increasing loop length (T_m = 92.0, 64.5, and 41.0 °C for **G3T1**, **G3T2**, and **G3T4**, respectively (Figure 3a and Figure S8) at 20 mM $[K^+]$).

We constructed cell lines with each unnatural G4 sequence inserted in front of the minimal promoter of our assay system with a constant SP1 consensus binding sequence. We then quantified folded G4 folding by G4 ChIP-qPCR. The level of folded G4 folding detected at this site in cells varied systematically with loop length and thermal stability (Figure 3c). The greatest G4 ChIP-qPCR signal was seen in cells with **G3T1**, the signal was reduced in cells with **G3T2**, and no G4 folding was discernible for cells with **G3T4**. In the context of our synthetic cellular system, the extent of folded G4 in cells is therefore directly related to the thermal stability of the G4.

Increased G4 Stability Leads to Increased Transcription. We then assessed transcriptional output for the **G3T1**, **G3T2**, and **G3T4** cellular insertions by RT-qPCR of eGFP mRNA. The level of eGFP reporter transcription varied systematically with loop length and thermal stability with **G3T1** cells exhibiting the highest transcription, followed by **G3T2** and **G3T4** cells giving ∼1.6- and 9.7-fold lower transcription, respectively (Figure 3d,e). Similar results were obtained using flow cytometry measurement of eGFP fluorescence (Figure S4; ∼1.5- and ∼12.7-fold reduction in MFI for 2- and 4-loop G4s compared to the 1-loop G4). Increased transcription therefore correlates with increased thermal stability of the G4 and the level of folded G4 formation in cells.

CONCLUSIONS

We have presented a systematic study of G4 folding in the context of a synthetic gene promoter that drives a reporter gene in chromatin of human cells. The insertion of natural and unnatural G4 sequence motifs led to a measurable folded G4 structure at the insertion site, with a concomitant increase in transcription of the proximal gene. Systematic variation in the thermal stability of folded G4s, by controlling loop lengths, showed that higher stability gives rise to an increased level of folded G4 signal in cells and higher level of transcription of the reporter gene. The main outcomes support the finding that the presence of folded G4 structures in front of transcribed genes can have a positive effect on transcription, probably through the recruitment of proteins such as transcription factors and chromatin remodelers.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acschembio.3c00775.

Detailed description of experimental procedures, materials, and additional figures as mentioned in the text (PDF)

AUTHOR INFORMATION

Corresponding Author

Shankar Balasubramanian − *Yusuf Hamied Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, U.K.; Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge CB2 0RE, U.K.; School of Clinical Medicine, University of Cambridge, Cambridge CB2 0SP, U.K.;* orcid.org/0000-0002-0281-5815; Email: SB10031@cam.ac.uk

Authors

Yuqi Chen − *Yusuf Hamied Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, U.K.;* orcid.org/0000-0002-5635-3409

Angela Simeone − *Yusuf Hamied Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, U.K.; Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge CB2 0RE, U.K.;* orcid.org/0000-0002-0663-0121

Larry Melidis − *Yusuf Hamied Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, U.K.; Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge CB2 0RE, U.K.;* orcid.org/0000-0001-6853-2722

Sergio Martinez Cuesta − *Yusuf Hamied Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, U.K.; Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge CB2 0RE, U.K.;* Present Address: Data Sciences and Quantitative Biology, Discovery Sciences, R&D, AstraZeneca, Cambridge CB2 0AA, U.K; orcid.org/0000-0001-9806-2805

David Tannahill − *Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge CB2 0RE, U.K.;* orcid.org/0000-0002-3811-6864

Complete contact information is available at: https://pubs.acs.org/10.1021/acschembio.3c00775

Author Contributions

Y.C., D.T., and S.B. conceived and designed the experiments. Y.C. performed the biophysical, cell, and molecular biology experiments. Computational analysis was performed by A.S., L.M., and S.M.C. All authors interpreted the results. Y.C., D.T., and S.B. wrote the manuscript with contributions from all authors.

Funding

This work was supported by Cancer Research UK Programme funding (C9681/A29214) and Core funding (C9545/A29580) and Herchel Smith funds (University of Cambridge). S.B. is a Wellcome Trust Senior Investigator (209441/Z/17/Z).

Notes

The authors declare the following competing financial interest(s): S.B. is a founder and shareholder of Biomodal Ltd, Inflex Ltd. and RNAvate Ltd. L.M. is a consultant for Inflex Ltd.

■ ACKNOWLEDGMENTS

We thank the members of the Balasubramanian laboratory for helpful discussions. We thank the Research Instrumentation core facilities at the Cancer Research UK Cambridge Institute.

■ REFERENCES

(1) Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S. Quadruplex DNA: sequence, topology and structure. *Nucleic acids res.* **2006**, *34* (19), 5402−5415.
(2) Huppert, J. L.; Balasubramanian, S. Prevalence of quadruplexes in the human genome. *Nucleic Acids Res.* **2005**, *33*, 2908−2916.
(3) Todd, A. K.; Johnston, M.; Neidle, S. Highly prevalent putative quadruplex sequence motifs in human DNA. *Nucleic Acids Res.* **2005**, *33*, 2901−2907.
(4) Chambers, V. S.; Marsico, G.; Boutell, J. M.; Di Antonio, M.; Smith, G. P.; Balasubramanian, S. High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nat. Biotechnol.* **2015**, *33* (8), 877−881.
(5) Hänsel-Hertsch, R.; Beraldi, D.; Lensing, S. V.; Marsico, G.; Zyner, K.; Parry, A.; Di Antonio, M.; Pike, J.; Kimura, H.; Narita, M.; et al. G-quadruplex structures mark human regulatory chromatin. *Nat. Genet.* **2016**, *48* (10), 1267−1272.
(6) Yu, Z.; Spiegel, J.; Melidis, L.; Hui, W. W. I.; Zhang, X.; Radzevičius, A.; Balasubramanian, S. Chem-map profiles drug binding to chromatin in cells. *Nat. Biotechnol.* **2023**, *41*, 1265−1271.
(7) Galli, S.; Melidis, L.; Flynn, S. M.; Varshney, D.; Simeone, A.; Spiegel, J.; Madden, S. K.; Tannahill, D.; Balasubramanian, S. DNA G-Quadruplex Recognition in Vitro and in Live Cells by a Structure-Specific Nanobody. *J. Am. Chem. Soc.* **2022**, *144*, 23096−23103.
(8) Liu, H. Y.; Zhao, Q.; Zhang, T. P.; Wu, Y.; Xiong, Y. X.; Wang, S. K.; Ge, Y. L.; He, J. H.; Lv, P.; Ou, T. M.; et al. Conformation Selective Antibody Enables Genome Profiling and Leads to Discovery of Parallel G-Quadruplex in Human Telomeres. *Cell Chem. Biol.* **2016**, *23*, 1261−1270.
(9) Feng, Y.; He, Z.; Luo, Z.; Sperti, F. R.; Valverde, I. E.; Zhang, W.; Monchaud, D. Side-by-side comparison of G-quadruplex (G4) capture efficiency of the antibody BG4 versus the small-molecule ligands TASQs. *iScience* **2023**, *26*, No. 106846.
(10) Huppert, J. L.; Balasubramanian, S. G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res.* **2007**, *35*, 406−413.
(11) Eddy, J.; Maizels, N. Gene function correlates with potential for G4 DNA formation in the human genome. *Nucleic Acids Res.* **2006**, *34*, 3887−3896.
(12) Broxson, C.; Beckett, J.; Tornaletti, S. Transcription Arrest by a G Quadruplex Forming-Trinucleotide Repeat Sequence from the Human c-myb Gene. *Biochemistry* **2011**, *50* (19), 4162−4172.
(13) Tornaletti, S.; Park-Snyder, S.; Hanawalt, P. C. G4-forming Sequences in the Non-transcribed DNA Strand Pose Blocks to T7 RNA Polymerase and Mammalian RNA Polymerase II. *J. Biol. Chem.* **2008**, *283* (19), 12756−12762.
(14) Siddiqui-Jain, A.; Grand, C. L.; Bearss, D. J.; Hurley, L. H. Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11593−11598.
(15) Ducani, C.; Bernardinelli, G.; Högberg, B.; Keppler, B. K.; Terenzi, A. Interplay of three G-quadruplex units in the KIT promoter. *J. Am. Chem. Soc.* **2019**, *141*, 10205−10213.
(16) Morgan, R. K.; Batra, H.; Gaerig, V. C.; Hockings, J.; Brooks, T. A. Identification and characterization of a new G-quadruplex forming region within the kRAS promoter as a transcriptional regulator. *Biochim. Biophys. Acta, Gene Regul. Mech.* **2016**, *1859* (2), 235−245.
(17) Paul, R.; Das, T.; Debnath, M.; Chauhan, A.; Dash, J. G-Quadruplex-Binding Small Molecule Induces Synthetic Lethality in Breast Cancer Cells by Inhibiting c-MYC and BCL2 Expression. *ChemBioChem.* **2020**, *21*, 963−970.
(18) Miyazaki, T.; Pan, Y.; Joshi, K.; Purohit, D.; Hu, B.; Demir, H.; Mazumder, S.; Okabe, S.; Yamori, T.; Viapiano, M.; et al. Telomestatin Impairs Glioma Stem Cell Survival and Growth through the Disruption of Telomeric G-Quadruplex and Inhibition of the Proto-oncogene, c-Myb. *Clin. Cancer Res.* **2012**, *18*, 1268.
(19) Fleming, A. M.; Ding, Y.; Burrows, C. J. Oxidative DNA damage is epigenetic by regulating gene transcription via base excision repair. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, 2604−2609.
(20) Cogoi, S.; Ferino, A.; Miglietta, G.; Pedersen, E. B.; Xodo, L. E. The regulatory G4 motif of the Kirsten ras (KRAS) gene is sensitive to guanine oxidation: implications on transcription. *Nucleic Acids Res.* **2018**, *46*, 661−676.
(21) Bejugam, M.; Sewitz, S.; Shirude, P. S.; Rodriguez, R.; Shahid, R.; Balasubramanian, S. Trisubstituted isoalloxazines as a new class of G-quadruplex binding ligands: Small molecule regulation of c-kit oncogene expression. *J. Am. Chem. Soc.* **2007**, *129*, 12926−12927.
(22) Rodriguez, R.; Miller, K. M.; Forment, J. V.; Bradshaw, C. R.; Nikan, M.; Britton, S.; Oelschlaegel, T.; Xhemalce, B.; Balasubramanian, S.; Jackson, S. P. Small-molecule−induced DNA damage identifies alternative DNA structures in human genes. *Nat. Chem. Biol.* **2012**, *8* (3), 301−310.
(23) Ohnmacht, S. A.; Marchetti, C.; Gunaratnam, M.; Besser, R. J.; Haider, S. M.; Di Vita, G.; Lowe, H. L.; Mellinas-Gomez, M.; Diocou, S.; Robson, M.; Šponer, J.; Islam, B.; Barbara Pedley, R.; Hartley, J. A.; Neidle, S.; et al. A G-quadruplex-binding compound showing anti-tumour activity in an in vivo model for pancreatic cancer. *Sci. Rep.* **2015**, *5* (1), 1−11.
(24) Zyner, K. G.; Simeone, A.; Flynn, S. M.; Doyle, C.; Marsico, G.; Adhikari, S.; Portella, G.; Tannahill, D.; Balasubramanian, S. G-quadruplex DNA structures in human stem cells and differentiation. *Nat. Commun.* **2022**, *13* (1), 1−17.
(25) Hänsel-Hertsch, R.; Simeone, A.; Shea, A.; Hui, W. W. I.; Zyner, K. G.; Marsico, G.; Rueda, O. M.; Bruna, A.; Martin, A.; Zhang, X.; et al. Landscape of G-quadruplex DNA structural regions in breast cancer. *Nat. Genet.* **2020**, *52* (9), 878−883.
(26) Lago, S.; Nadai, M.; Cernilogar, F. M.; Kazerani, M.; Domíniguez Moreno, H.; Schotta, G.; Richter, S. N. Promoter G-quadruplexes and transcription factors cooperate to shape the cell type-specific transcriptome. *Nat. Commun.* **2021**, *12* (1), 1−13.
(27) Raiber, E. A.; Kranaster, R.; Lam, E.; Nikan, M.; Balasubramanian, S. A non-canonical DNA structure is a binding motif for the transcription factor SP1 in vitro. *Nucleic Acids Res.* **2012**, *40*, 1499.
(28) Spiegel, J.; Cuesta, S. M.; Adhikari, S.; Hánsel-Hertsch, R.; Tannahill, D.; Balasubramanian, S. G-quadruplexes are transcription factor binding hubs in human chromatin. *Genome Biol.* **2021**, *22*, 1−15.
(29) Li, L.; Williams, P.; Ren, W.; Wang, M. Y.; Gao, Z.; Miao, W.; Huang, M.; Song, J.; Wang, Y. YY1 interacts with guanine quadruplexes to regulate DNA looping and gene expression. *Nat. Chem. Biol.* **2021**, *17* (2), 161−168.
(30) Zhang, X.; Spiegel, J.; Martínez Cuesta, S.; Adhikari, S.; Balasubramanian, S. Chemical profiling of DNA G-quadruplex-interacting proteins in live cells. *Nat. Chem.* **2021**, *13* (7), 626−633.
(31) The FANTOM Consortium and the RIKEN PMI and CLST (DGT). A promoter-level mammalian expression atlas. *Nature* **2014**, *507* (7493), 462−470.
(32) del Villar-Guerra, R.; Trent, J. O.; Chaires, J. B. G-Quadruplex Secondary Structure Obtained from Circular Dichroism Spectroscopy. *Angew. Chem., Int. Ed.* **2018**, *57*, 7171−7175.
(33) del Villar-Guerra, R.; Gray, R. D.; Chaires, J. B. Quadruplex DNA Structure Characterization by Circular Dichroism. *Curr. Protoc. Nucleic Acid Chem.* **2017**, *68*, 17.8.1−17.8.16.
(34) Balagurumoorthy, P.; Brahmachari, S. K.; Mohanty, D.; Bansal, M.; Sasisekharan, V. Hairpin and parallel quartet structures for telomeric sequences. *Nucleic Acids Res.* **1992**, *20*, 4061−4067.
(35) Berova, N.; Nakanishi, K.; Woody, R. *Circular dichroism: principles and applications*. John Wiley & Sons: New York, 2000 ISBN: 978−0-471−33003−5.
(36) Giraldo, R.; Suzuki, M.; Chapman, L.; Rhodes, D. Promotion of parallel DNA quadruplexes by a yeast telomere binding protein: a

https://doi.org/10.1021/acschembio.3c00775
ACS Chem. Biol. 2024, 19, 736−742

circular dichroism study. *Procs. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 7658−7662.
(37) Mergny, J.; Lacroix, L. UV Melting of G-Quadruplexes. *Curr. Protoc. Nucleic Acid Chem.* **2009**, *37*, 17.1.1−17.1.15.
(38) Mergny, J. L.; Lacroix, L. Analysis of Thermal Melting Curves. *Oligonucleotides* **2003**, *13* (6), 515−537.
(39) Todd, A. K.; Neidle, S. The relationship of potential G-quadruplex sequences in cis -upstream regions of the human genome to SP1-binding elements. *Nucleic Acids Res.* **2008**, *36*, 2700−2704.
(40) Bugaut, A.; Balasubramanian, S. A sequence-independent study of the influence of short loop lengths on the stability and topology of intramolecular DNA G-quadraplexes. *Biochemistry* **2008**, *47*, 689−697.
(41) Guédin, A.; Gros, J.; Alberti, P.; Mergny, J. L. How long is too long? Effects of loop size on G-quadruplex stability. *Nucleic Acids Res.* **2010**, *38*, 7858−7868.

https://doi.org/10.1021/acschembio.3c00775
ACS Chem. Biol. 2024, 19, 736−742