Solution Structure of a MYC Promoter G-Quadruplex with 1:6:1 Loop Length

Jonathan Dickerhoff,* Buket Onel,† Luying Chen,‡ Yuwei Chen,§ and Danzhou Yang†,*,∥

†College of Pharmacy, Medicinal Chemistry and Molecular Pharmacology, Purdue University, 575 W Stadium Avenue, West Lafayette, Indiana 47907, United States
‡Department of Chemistry and Biochemistry, College of Science, University of Arizona, 1306 E. University Blvd, Tucson, Arizona 85721, United States
§Center for Cancer Research, Purdue University, 201 S University Street, West Lafayette, Indiana 47906, United States
||Purdue Institute for Drug Discovery, 720 Clinic Dr, West Lafayette, Indiana 47907, United States

ABSTRACT: The important MYC oncogene is deregulated in many cancer cells and comprises one of the most prominent G-quadruplex (G4) forming sequences in its promoter regions, the NHE III1 motif. Formation of G4s suppresses MYC transcription and can be modulated by drug binding, establishing these DNA structures as promising targets in cancer therapy. The NHE III1 motif can fold into more than one parallel G4s, including 1:2:1 and 1:6:1 loop length conformers, with the 1:2:1 conformer shown as the major species under physiological conditions in solution. However, additional factors such as protein interactions may affect the cellular folding equilibrium. Nucleolin, a protein shown to bind MYC G4 and repress MYC transcription, is reported herein to preferably bind to the 1:6:1 loop length conformer suggesting a physiological significance of this species. The high-resolution NMR solution structure of the 1:6:1 conformer is determined, which reveals a 5′-capping structure distinctive from the 1:2:1 form, with the 6 nt central loop playing an essential role for this specific capping structure. This suggests that each parallel G-quadruplex likely adopts unique capping and loop structures determined by the specific central loop and flanking sequences. The resulting structural information at the molecular level will help to understand protein recognition of different G4s, contribution of G4 polymorphism to gene regulation, and to rationally design small molecules selectively targeting the 1:6:1 MYC G4.

INTRODUCTION

G-Quadruplexes (G4) have emerged as one of the most exciting nucleic acid secondary structures. These highly diverse structures are formed by guanosine-rich sequences when four Gs associate in a cyclic array connected by Hoogsteen hydrogen bonds, called G-tetrads. Most reported structures consist of a G-core with three stacked tetrads that coordinate monovalent cations, that is, potassium or sodium through their carbonyl groups. The rich variety of conformations is achieved by numerous combinations of connecting loop structures and enclosing flanking sequences.1

In recent years, a significant number of potential G4 forming sequences were identified throughout the human genome and their existence in cell was shown.2−4 Strikingly, G4 motifs are often concentrated in promoter regions of highly regulated genes related to tumorigenesis, suggesting their role in transcriptional control and making G4s potential drug targets for cancer therapeutics.5

One of the most prominent examples is the G-rich nucleosome hypersensitive element (NHE III1) within the promoter region of the MYC oncogene.6,7 The encoded MYC protein is an influential transcription factor affecting cell growth and one of the most deregulated proteins in cancer cells.8 Repressing MYC transcription on the DNA level by ligand induced G4 stabilization is a promising strategy because the protein is short-lived and lacks clear sites for drug recognition.

The G-rich motif within the MYC promoter NHE III1 consists of five consecutive sequences with at least three Gs that can be assembled into different G4 conformers. Several high-resolution structures are reported that are all classified as parallel based on the sugar-phosphate backbone orientation of their G-cores.9−12 Analysis of dimethylsulfate (DMS) footprinting experiments under physiological conditions revealed the major conformation to have three propeller loops with respective loop lengths of 1:2:1.7 In addition to the 1:2:1 conformer, the formation of a 1:6:1 species was observed and shown to be thermodynamically favored. Later, it was
The evolutionary conservation of the long G-rich motif suggests an advantage in its inherent polymorphism and the resulting competition between different conformations. Transcription-generated dynamic superhelicity was shown to be the driving force of gene promoter G4 folding. In addition, DNA–protein interactions play an important role in G4 formation and regulation. Protein or small molecule interactions can shift the folding equilibrium toward a particular structure to fine-tune gene expression as an additional layer of transcriptional regulation. The nucleolin protein was found to specifically bind and stabilize the MYC promoter G4 and functions as transcriptional repressor. In this study, it is shown that nucleolin prefers the 1:6:1 G4 formed by the MYC promoter sequence NHE III, (Figure 1A) over the 1:2:1 conformer, the proposed physiologically major species. The high-resolution NMR structure of the 1:6:1 G4 was determined and revealed unique structural features with a 5′-capping triad ATG involving the last residue of the 6 nt central loop. This information can guide future studies to understand the fundamentals of G4–protein interactions based on structural motifs and will help to design small molecules specific for this 1:6:1 G4 by targeting this unique 5′-capping structure.

RESULTS

Competition EMSA Shows Nucleolin Preferably Binds the 1:6:1 over the 1:2:1 MYC G4. The 27 nt long oligonucleotide MycPu27 is often chosen as representation of the G4-forming motif NHE III, within the MYC promoter for in vitro studies (Figure 1B). On the basis of this fragment, sequences can be designed that adopt only a particular structure while maintaining the specific features of this conformation in the MycPu27 context. Generally, these are named based on the position of the G-stretches constituting their G-core and the last two 3′-terminal Gs are omitted because their involvement was not shown by DMS footprinting under physiological conditions. The proposed physiologically major species is Myc2345 (using G-stretches II, III, IV, and V) with propeller loops of 1:2:1 nt length, which is regularly used in drug studies to develop small molecules for MYC silencing. On the other hand, the formation of the longer loop conformer 1:6:1 can be induced by substituting the central loop residues G11–G14 with Ts, removal of the 3′-terminal G overhang, and addition of flanking Ts yielding the Myc1245_14T sequence. Binding of the nucleolin protein to both loop conformers was previously reported without direct comparison. Nucleolin is a 77 kDa multidomain protein, with an intrinsically disordered N-terminal domain, a central core of four RNA-binding domains, and a C-terminal RGG domain. It was shown that the N-terminal domain does not affect the MYC G4 binding. Therefore, an N-terminal truncated protein was used in the present study.

A competition EMSA was performed to analyze the binding of nucleolin to the 1:6:1 and 1:2:1 MYC G4s (Figure 1C). The band of radiolabeled Myc1245_14T is shifted upward upon nucleolin binding as seen in the second lane of the gel in comparison to the first control lane without protein. Addition of unlabeled competitor DNA in excess with similar or higher affinity displaces the labeled Myc1245_14T from nucleolin as shown by a reduced intensity of the complex band. In lane 3 and 4, the sequences Myc2345 and Myc1245_14T were added as competitor, respectively, with only the latter showing a significant attenuation of the upper complex band. The Myc1245_14T sequence serves as a control because addition of the identical unlabeled sequence with same affinity increases the overall receptor pool and decreases the population of nucleolin bound to radiolabeled DNA. However, Myc2345 binding by nucleolin is clearly disfavored compared to the longer loop conformer and it cannot displace Myc1245_14T even at high excess. This is consistent with the reported preference of nucleolin for G4 with longer loops.

Therefore, in direct competition, the 1:6:1 sequence with a longer central loop is preferably bound by the nucleolin protein and cannot be displaced by the 1:2:1 conformer, suggesting that protein binding might shift the in cell folding equilibrium toward the 1:6:1 conformer Myc1245_14T for the MYC promoter G4s.

Optimizing the Myc1245 Sequence for NMR Analysis. With this renewed physiological interest in Myc1245, it is desirable to learn its structural features at atomic resolution. A
Single mutations of the non-native T14 central loop residue closest to the S′-tetrad were evaluated to overcome this interaction. Two sequences, Myc1245 and Myc1245_14G with the two purine bases, A and the native G, respectively, were tested. Indeed, no broadening was observed for both sequences confirming T14’s involvement in dynamic interactions in Myc1245. Surprisingly, the imino proton spectra of Myc1245 and Myc1245_14G are nearly superimposable and resemble the high temperature spectrum of Myc1245_14T (Figure 1D). This indicates a strong structural similarity of all sequences and the non-involvement of the position 14 purine base in a S′-capping. Ultimately, Myc1245 was chosen as sequence for NMR structure determination based on the additional base proton H2 compared to G that can provide more structural information without altering the structure.

**Determining the High-Resolution NMR Structure of Myc1245.** A set of 2D NMR spectra were recorded for Myc1245 including NOE, COSY, and a DQF-COSY, and a 1H−13C HSQC of the aromatic region (Figures 2, S1, and S2). On the basis of the reported parallel topology of the G-core, standard assignment strategies following the sequential contacts were employed (Table S1). Generally, this cross peak pattern is assignment strategies following the sequential contacts were generally expected for a three tetrad G4 corresponding to the 12 hydrogen bonded Gs within the core. Clearly, Myc1245_14T adopts only one major conformation; however, broadening is observed for the G7 and G16 imino protons of the S′-tetrad that are in close proximity to the central 6 nt loop. This highly localized effect, not observed for a sequence missing the terminal T0, suggests the presence of dynamic interactions involving the mutated central loop and the S′-terminus. In addition, this possibly artificial dynamic is temperature-dependent and disappears at higher temperatures (Figure 1D).

1D NMR spectrum was recorded for Myc1245_14T to evaluate if the spectral quality was sufficient for NMR structure determination (Figure 1D). Twelve imino proton resonances are generally expected for a three tetrad G4. Measured at 25 °C with 100 mM K+, pH 7.

**Figure 2.** (A) 2D NOESY spectral region of Myc1245 with 300 ms mixing time showing the sequential H8−H1′ contacts traced by solid lines. Missing NOE cross peaks are marked with an asterisk. 2D 1H−13C HSQC of Myc1245 showing the (B) H6−C6/H8−C8 peaks for all bases and the (C) adenine H2−C2 contacts. All spectra measured at 25 °C with 100 mM K+, pH 7.

**Figure 3.** Superposition of the 10 lowest energy structures for Myc1245 (PDB: 6NEB).

**Structure of the Myc1245 1:6:1 G4.** The centerpiece of the S′-capping is an ATG triad, consisting of A15, T1, and G2, that connects the central loop and the S′-flanking. Within the 6 nt central loop, only the syn conformer A15 stacks upon the S′-tetrad and is bound by T1 to form a Hoogsteen hydrogen-bonded T1−A15 base pair (Figure 4A). Thereby, about half of the tetrad including G16 and G20 is covered. The orientation of this AT base pair is clearly defined by several NOE cross peaks (Tables S2 and S3). For example, T1’s methyl group shows dipolar coupling with G16−H1 and G20−H8. Also, A15’s position is restrained by several interactions of its aromatic protons. While A15−H8 has NOE cross peaks with...
G7−H1, A15−H2 interacts with both G7−H1 and G7−H1′. Furthermore, the T1−A15 base pair layer upon the 5′-tetrad is complemented by G2 which is stacking on G3, as defined by a G2−H8/G20−H1 cross peak and the standard sequential NOE pattern between the two adjacent G. Slight broadening of the G2 NOESY resonances may suggest a more dynamic position as formation of only one hydrogen bond with the AT base pair is indicated by the proximity of G2’s carbonyl function and the A15 exocyclic amino group. Finally, T0 is stacked on top of the ATG triad and located close to the central loop as shown by NOE interactions between the T0 methyl group and A15’s H2, H5′, and H5″ protons (Figure 4B and Table S2).

The central loop of 6 nt connects the outer tetrads by spanning the G-core. A single nucleotide is the minimal length of this structural motif, so the five additional residues can significantly increase the loop’s conformational flexibility. However, the overall structure is well converged with an RMSD of 1.23 Å, probably due to the interaction of A15 with the 5′-flanking. This is the sole interaction of a loop residue with flanking motifs, and because A15 is stacked upon the tetrad, the loop is stretched toward the 5′-terminus. In this way, analogous interactions with the 3′-flanking are excluded and the adoptable conformational space is limited. Overall, the solvent exposure of the loop DNA bases is reduced by their positioning within the groove as hydrophobic dehydration is described as an important energetic contribution to G4 folding (Figure 4C, D). Consequently, the sugar phosphate backbone is turned outside and electrostatic repulsion between the negative phosphates of loop and G-core are minimized by the loop’s central position. With A15 stacked upon the 5′-tetrad, A14 and T10 are the outer residues of the flexible loop segment and closest to the G-core. As already indicated by the resemblance of Myc1245 and Myc1245_14G NMR spectra, A14 is not located above the 5′-tetrad and instead points into the groove. There, it is in close proximity to T10 based on a T10-Me/A14-H2 NOE cross peak. Numerous additional dipolar contacts between T10’s H6 or methyl protons and the core residues G8, G9, G16, and G17 confirm its position (Table S2). The residues T10 and A14 are covered by T11 and T12/T13, respectively, with their bases pointing to opposite directions. This separation of the central loop into two parts comprising T10−T11 and T12−A14 is also reflected by the further attenuated NOE cross peak pattern between T11−T12 (Table S3).

In contrast to the 5′-terminus, the 3′-flanking sequence does not interact with the central loop and instead forms an internal

---

**Table 1. NMR Restraints and Structural Statistics for the Myc1245 Quadruplex**

| NOE-Based Distance Restraints |  |
|-------------------------------|---|
| intraresidue                  | 373 |
| inter-residue                 | 175 |
| sequential                    | 55  |

| Other Restraints              |   |
|-------------------------------|---|
| hydrogen bonds                | 56  |
| torsion angles                | 46  |
| G-tetrad planarity            | 36  |

| Structural Statistics         |   |
|-------------------------------|---|
| pairwise heavy atom RMSD (Å)  | 0.45 ± 0.11 |
| G-tetrad core                 | 1.23 ± 0.31 |
| all residues                  | 0.74 ± 0.21 |
| without central loop          | 0.001 ± 0.008 |
| mean NOE restraint violation (Å) | 0.19 |
| max. NOE restraint violation (Å) | 0.01 ± 0.00 |
| deviations from idealized geometry | 2.19 ± 0.02 |

---

G7−H1, A15−H2 interacts with both G7−H1 and G7−H1′. Furthermore, the T1−A15 base pair layer upon the 5′-tetrad is complemented by G2 which is stacking on G3, as defined by a G2−H8/G20−H1 cross peak and the standard sequential NOE pattern between the two adjacent G. Slight broadening of the G2 NOESY resonances may suggest a more dynamic position as formation of only one hydrogen bond with the AT base pair is indicated by the proximity of G2’s carbonyl function and the A15 exocyclic amino group. Finally, T0 is stacked on top of the ATG triad and located close to the central loop as shown by NOE interactions between the T0 methyl group and A15’s H2, H5′, and H5″ protons (Figure 4B and Table S2).

The central loop of 6 nt connects the outer tetrads by spanning the G-core. A single nucleotide is the minimal length of this structural motif, so the five additional residues can significantly increase the loop’s conformational flexibility. However, the overall structure is well converged with an RMSD of 1.23 Å, probably due to the interaction of A15 with the 5′-flanking. This is the sole interaction of a loop residue with flanking motifs, and because A15 is stacked upon the tetrad, the loop is stretched toward the 5′-terminus. In this way, analogous interactions with the 3′-flanking are excluded and the adoptable conformational space is limited. Overall, the solvent exposure of the loop DNA bases is reduced by their positioning within the groove as hydrophobic dehydration is described as an important energetic contribution to G4 folding (Figure 4C, D). Consequently, the sugar phosphate backbone is turned outside and electrostatic repulsion between the negative phosphates of loop and G-core are minimized by the loop’s central position. With A15 stacked upon the 5′-tetrad, A14 and T10 are the outer residues of the flexible loop segment and closest to the G-core. As already indicated by the resemblance of Myc1245 and Myc1245_14G NMR spectra, A14 is not located above the 5′-tetrad and instead points into the groove. There, it is in close proximity to T10 based on a T10-Me/A14-H2 NOE cross peak. Numerous additional dipolar contacts between T10’s H6 or methyl protons and the core residues G8, G9, G16, and G17 confirm its position (Table S2). The residues T10 and A14 are covered by T11 and T12/T13, respectively, with their bases pointing to opposite directions. This separation of the central loop into two parts comprising T10−T11 and T12−A14 is also reflected by the further attenuated NOE cross peak pattern between T11−T12 (Table S3).

In contrast to the 5′-terminus, the 3′-flanking sequence does not interact with the central loop and instead forms an internal

---

**Figure 4.** (A) Top view of the ATG triad covering the 5′-tetrad and (B) complete 5′-capping structure. (C + D) Side views of the central loop emphasizing its two parts. (E) Top view of the base pair covering the 3′-tetrad and (F) complete 3′-capping structure.
capping structure (Figure 4E). The terminal T26 folds back and stacks upon the 3′-tetrad, forming a reverse wobble base pair with G23 that diagonally spans the tetrad from G9 to G22. Standard sequential NOE cross peaks observed between G22 and G23 demonstrate that they are stacked as a continuation of the G-core (Table S3). On the other hand, T26’s orientation is unambiguously restrained by dipolar interactions of its methyl protons to G9−H8 as well as to both G5 and G9 imino protons. Thus, the T26 sugar points toward the groove formed by G9 and G18. The two adenosines A24 and A25 link the GT base pair and show no interaction with the 3′-tetrad (Figure 4F). However, their bases are oriented toward the G5/G9 site based on weak interactions with T26-Me.

DISCUSSION

The high-resolution structure of the 1:6:1 species Myc1245 provides new molecular insights into the structural diversity of MYC G4s. Interestingly, both Myc1245 and the 1:2:1 Myc2345 conformations have the same 3′-flanking sequence in the genomic context and will probably share a common capping structure. However, the 3′-capping motifs differ strongly dependent on the central loop length and thereby provide distinctive recognition sites for proteins and small molecules. Most of the 3′-flanking residues within the Myc2345 structure simply stack upon the tetrad without specific interactions; thus, the first flanking residue A6 5′ to the G-core is easy to access and often recruited by end-stacking drugs. In the case of 1:6:1 Myc1245, a far more complex capping structure is formed by the interplay of the longer central loop and the 3′-flanking. Interestingly, this kind of interaction is observed among parallel G4s with a central loop of at least 3 nt, which are often found in promoter regions, such as VEGF, hTERT, or c-kit.

The 3′ ATG triad of the 1:6:1 conformer Myc1245 provides a unique structural profile for ligand recognition compared to the 1:2:1 species. In addition, the higher localization of residues around the 3′-tetrad allows the constitution of more complex binding pockets that are often observed for nonparallel structures with lateral or diagonal loops.

On the basis of the now gained information about the 1:6:1 MYC G4, future studies can give insight which structural motif provides new molecular insights into the structural diversity of MYC G4s. Indeed, both Myc1245 and the 1:2:1 Myc2345 conformations have the same 3′-flanking sequence in the genomic context and will probably share a common capping structure. However, the 3′-capping motifs differ strongly dependent on the central loop length and thereby provide distinctive recognition sites for proteins and small molecules. Most of the 3′-flanking residues within the Myc2345 structure simply stack upon the tetrad without specific interactions; thus, the first flanking residue A6 5′ to the G-core is easy to access and often recruited by end-stacking drugs. In the case of 1:6:1 Myc1245, a far more complex capping structure is formed by the interplay of the longer central loop and the 3′-flanking. Interestingly, this kind of interaction is observed among parallel G4s with a central loop of at least 3 nt, which are often found in promoter regions, such as VEGF, hTERT, or c-kit.

In conclusion, the high-resolution structure of the 1:6:1 species Myc1245 reported herein is an important addition to the ensemble of MYC G4 structures as well as parallel G4s. Nucleolin’s preference for this conformation over the 1:2:1 loop conformer Myc2345 strongly suggests its physiological significance and might imply how protein binding can fine-tune gene expression by stabilization of a particular G4.
Structures were analyzed and visualized with PyMOL and energy structures were selected for the energy-minimized for 500 steps in vacuum, and the 10 lowest energy structures were deposited in the Protein Data Bank (6NEB).

The OL15 version of the Amber force field for DNA was used that further modifies the parmbsc0 version. Structures were initially equilibrated for 5 ps at 100 K and then heated to 1000 K during 1 ps. The system was cooled down to 100 K during 45 ps after 30 ps of high temperature equilibration. Finally, a cool down to 0 K was performed in the last 10 ps.

Dihedral restraints were applied for glycosidic torsion angles with 170°–310° and 200°–280° for anti-conformers in loop regions and within the G-core, respectively. The single syn conformer A15 was restricted to 0°–120°. The DQF-COSY spectrum was used to confirm the south sugar pucker based on the difference in signal intensity for H1’–H2’ and H1’–H2” cross peaks. This could not be achieved with certainty for residues 6, 8, 9, 13, 16, 23, 24, 26 due to spectral crowding and isochronicity of H2’/H2” resonances. The pseudorotation phase angle of all other residues was restricted to 144°–180° for a south-type sugar during simulated annealing. A distance geometry simulated annealing protocol was used in Xplor-NIH 2.48 to generate 100 starting structures. The sander module of the Amber 16 package was employed for simulated annealing of the 10 starting structures in implicit water. Finally, the last 500 ps of the trajectories were averaged and assigned in case of overlapped and thus ambiguous resonances. NOESY and DQF-COSY spectra; table with chemical shifts; stereo view of determined G4 structure; and tables with sequential and long-range NOE interactions were deposited in the Protein Data Bank (6NEB).

Data Deposition. The coordinates of the Myc1245 molecule interaction with a five-guanine-tract G-quadruplex structure were deposited in the Protein Data Bank (6NEB).

ASSOCIATED CONTENT

Supporting Information
The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acsomega.8b03580.

NOESY and DQF-COSY spectra; table with chemical shifts; stereo view of determined G4 structure; and tables with sequential and long-range NOE interactions are provided (PDF).

AUTHOR INFORMATION

Corresponding Author
E-mail: yangdz@purdue.edu. Phone: (765) 494-8148. Fax: (765) 494-8352.

ORCID

Jonathan Dickerhoff: 0000-0002-3197-9472
Danzhou Yang: 0000-0001-7489-7111

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This research was supported by the National Institutes of Health (R01CA177585 (D.Y.) and P30CA023168 (Purdue Center for Cancer Research)). We thank Dr. Clement Lin for proofreading the manuscript.

REFERENCES

1. Zhang, S.; Wu, Y.; Zhang, W. G-Quadruplex Structures and Their Interaction Diversity with Ligands. ChemMedChem 2014, 9, 899–911.
2. Biffi, G.; Tannahill, D.; McCallery, J.; Balasubramanian, S. Quantitative visualization of DNA G-quadruplex structures in human cells. Nat. Chem. 2013, 5, 182–186.
3. Hänsel-Herths; R.; Beraldi, D.; Lensing, S. V.; Marsico, G.; Zynier, K.; Parry, A.; Di Antonio, M.; Pike, J.; Kimura, H.; Narita, M.; Tannahill, D.; Balasubramanian, S. G-Quadruplex structures mark human regulatory chromatin. Nat. Genet. 2016, 48, 1267–1272.
4. Hänsel, R.; Löh, F.; Trantirek, L.; Dötsch, V. High-resolution insight into G-overhang architecture. J. Am. Chem. Soc. 2013, 135, 2816–2824.
5. Balasubramanian, S.; Hurley, L. H.; Neidle, S. Targeting G-quadruplexes in gene promoters: A novel anticancer strategy? Nat. Rev. Drug Discovery 2011, 10, 261–275.
6. González, V.; Hurley, L. H. The c-MYC NHE III1: Function and Regulation. Annu. Rev. Pharmacol. Toxicol. 2010, 50, 111–129.
7. 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. 2002, 99, 11590–11598.
8. Dang, C. V. MYC on the path to cancer. Cell 2012, 149, 22–35.
9. Mathad, R. I.; Hatzakis, E.; Dai, J.; Yang, D. c-MYC promoter G-quadruplex formed at the 5′-end NHE III 1 element: insights into biological relevance and parallel-stranded G-quadruplex stability. Nucleic Acids Res. 2011, 39, 9023–9033.
10. Stump, S.; Mou, T.-C.; Sprang, S. R.; Natale, N. R.; Beall, H. D. Crystal structure of the major quadruplex formed in the promoter region of the human c-MYC oncogene. PLoS One 2018, 13, No. e0205584.
11. Ambras, A.; Chen, D.; Dai, J.; Jones, R. A.; Yang, D. Solution Structure of the Biologically Relevant G-Quadruplex Element in the Human c-MYC Promoter. Implications for G-Quadruplex Stabilization. Biochemistry 2005, 44, 2048–2058.
12. Phan, A. T.; Kuruvyi, V.; Gav, H. Y.; Patel, D. J. Small-molecule interaction with a five-guanine-tract G-quadruplex structure from the human MYC promoter. Nat. Chem. Biol. 2005, 1, 167–173.
13. Phan, A. T.; Modi, Y. S.; Patel, D. J. Propeller-Type Parallel-Stranded G-Quadruplexes in the Human-mycPromoter. J. Am. Chem. Soc. 2004, 126, 8710–8716.
14. Kouzine, F.; Sanford, S.; Elisha-Feil, Z.; Levens, D. The functional response of upstream DNA to dynamic supercoiling in vivo. Nat. Struct. Mol. Biol. 2008, 15, 146–154.
15. Kouzine, F.; Levens, D. Supercoid-driven DNA structures regulate genetic transactions. Front. Biosci. 2007, 12, 4409–4423.
16. Zheng, K.-w.; He, Y.-d.; Liu, H.-h.; Li, X.-m.; Hao, Y.-h.; Tan, Z. Superhelicity constrains a localized and R-loop-dependent DNA structure with synthetic applications. ACS Omega 2019, 4, 2533–2539.
formation of G-quadruplexes at the upstream region of transcription. ACS Chem. Biol. 2017, 12, 2609–2618.

(17) Xia, Y.; Zheng, K.-W.; He, Y.-d.; Liu, H.-h.; Wen, C.-j.; Hao, Y.-h.; Tan, Z. Transmission of dynamic supercoiling in linear and multi-way branched DNAs and its regulation revealed by a fluorescent G-quadruplex torsion sensor. Nucleic Acids Res. 2018, 46, 7418–7424.

(18) González, V.; Guo, K.; Hurley, L.; Sun, D. Identification and Characterization of Nucleolin as a c-myc G-quadruplex-binding Protein. J. Biol. Chem. 2009, 284, 23622–23635.

(19) Sutherland, C.; Cui, Y.; Mao, H.; Hurley, L. A Mechanosensor Mechanism Controls the G-Quadruplex/α-Motif Molecular Switch in the MYC Promoter NHE III. J. Am. Chem. Soc. 2016, 138, 14138–14151.

(20) Dai, J.; Carver, M.; Hurley, L. H.; Yang, D. Solution Structure of a 2:1 Quindoline-c-MYC G-Quadruplex: Insights into G-Quadruplex-Interactive Small Molecule Drug Design. J. Am. Chem. Soc. 2011, 133, 17673–17680.

(21) Lago, S.; Tosoni, E.; Nada, M.; Palumbo, M.; Richter, S. N. The cellular protein nucleolin preferentially binds long-looped G-quadruple nucleic acids. Biochim. Biophys. Acta BBA - Gen. Subj. 2017, 1861, 1371–1381.

(22) Greene, K. L.; Wang, Y.; Live, D. Influence of the glycosidic torsion angle on 13C and 15N shifts in guanosine nucleotides: Investigations of G-tetrad models with alternating syn and anti bases. J. Biomol. NMR 1995, 5, 333–338.

(23) Boncina, M.; Vesnaveer, G.; Chaires, J. B.; Lah, J. Unraveling the thermodynamics of the folding and interconversion of human telomere G-quadruplexes. Angew. Chem., Int. Ed. 2016, 55, 10340–10344.

(24) Calabrese, D. R.; Chen, X.; Leon, E. C.; Gaikwad, S. M.; Phyo, Z.; Hewitt, W. M.; Alden, S.; Hilimire, T. A.; He, F.; Michalowski, A. M.; Simmons, K. L.; Saunders, L. B.; Zhang, S.; Connors, D.; Walters, K. J.; Mock, B. A.; Schneekloth, J. S., Jr. Chemical and structural studies provide a mechanistic basis for recognition of the MYC G-Quadruplex. J. Am. Chem. Soc. 2009, 131, 1783–1784.

(25) Agrawal, P.; Hatzakis, E.; Guo, K.; Carver, M.; Yang, D. Solution Structure of the major G-Quadruplex formed in the human c-kit2 promoter. Nucleic Acids Res. 2013, 41, 1–9.

(26) Lim, K. W.; Lacroix, L.; Yue, D. J. E.; Lim, J. K. C.; Lim, J. M. W.; Phan, A. T. Coexistence of two distinct G-quadruplex conformations in the hTERT promoter. J. Am. Chem. Soc. 2010, 132, 12331–12342.

(27) Kuryavyi, V.; Phan, A. T.; Patel, D. J. Solution structures of all parallel-stranded monomeric and dimeric G-quadruplex scaffolds of the human c-kit2 promoter. Nucleic Acids Res. 2010, 38, 6757–6773.

(28) Lin, C.; Wu, G.; Wang, K.; Onel, B.; Sakai, S.; Shao, Y.; Yang, D. Molecular recognition of the hybrid-2 human telomeric G-quadruplex by epiberberine: Insights into conversion of telomeric G-quadruplex structures. Angew. Chem., Int. Ed. 2018, 57, 10888–10893.

(29) Liu, W.; Zhong, Y.-F.; Liu, L.-Y.; Shen, C.-T.; Zeng, W.; Wang, F.; Yang, D.; Mao, Z.-W. Solution structures of multiple G-quadruplex complexes induced by a platinum(II)-based tripod reveal dynamic binding. Nat. Commun. 2018, 9, 3496.

(30) Islam, B.; Stadlbauer, P.; Gil-Ley, A.; Pérez-Hernández, G.; Haider, S.; Neidle, S.; Bussi, G.; Banas, P.; Otyepka, M.; Sponer, J. Exploring the dynamics of propeller loops in human telomeric DNA quadruplexes using atomistic simulations. J. Chem. Theory Comput. 2017, 13, 2458–2480.

(31) Johansson, C.; Finger, L. D.; Trantirek, L.; Mueller, T. D.; Kim, S.; Laird-Offringa, I. A.; Feigon, J. Solution structure of the complex formed by the two N-terminal RNA-binding domains of nucleolin and a pre-rRNA Target. J. Mol. Biol. 2004, 337, 799–816.

(32) Onel, B.; Carver, M.; Wu, G.; Timonina, D.; Kalam, S.; Larriva, M.; Yang, D. A new G-quadruplex with hairpin loop immediately upstream of the human BCL2 PI promoter modulates transcription. J. Am. Chem. Soc. 2016, 138, 2563–2570.

(33) Vranken, W. F.; Boucher, W.; Stevens, T. J.; Fogh, R. H.; Pajon, A.; Linas, M.; Ulrich, E. L.; Markley, J. L.; Ionides, J.; Laue, E. D. The CCPN data model for NMR spectroscopy: Development of a software pipeline. Proteins Struct. Funct. Bioinforma. 2005, 59, 687–696.