Published in final edited form as:

J Am Chem Soc. 2016 March 2; 138(8): 2563–2570. doi:10.1021/jacs.5b08596.

A new G quadruplex with hairpin loop immediately upstream of the human BCL2 P1 promoter modulates transcription

Buket Onel[†], **Megan Carver**[‡], **Guanhui Wu**[‡], **Daria Timonina**[†], **Salil Kalarn**[†], **Marti Larriva**[‡], and **Danzhou Yang**[‡,§,‖,†,*]

[‡]Department of Pharmacology and Toxicology, College of Pharmacy, University of Arizona, 1703 E. Mabel St, Tucson, AZ 85721, USA

[†]Department of Chemistry and Biochemistry

[§]BIO5 Institute

[‖]The Arizona Cancer Center

Abstract

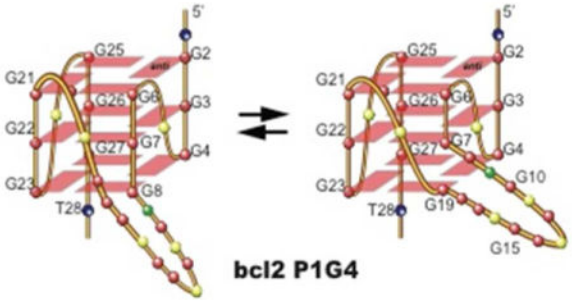
The abnormal over-expression of the BCL2 gene is associated with many human tumors. We found a new 28-mer G-quadruplex-forming sequence, P1G4, immediately upstream of the human BCL2 gene P1 promoter. The P1G4 is shown to be a transcription repressor using a promoter-driven luciferase assay; its inhibitory effect can be markedly enhanced by the G-quadruplex-interactive compound TMPyP4. G-quadruplex can readily form in the P1G4 sequence under physiological salt condition as shown by DMS footprinting. P1G4 and previously identified Pu39 G-quadruplexes appear to form independently in adjacent regions in the BCL2 P1 promoter. In the extended BCL2 P1 promoter region containing both Pu39 and P1G4, P1G4 appears to play a more dominant role in repressing the transcriptional activity. Using NMR spectroscopy, the P1G4 G-quadruplex appears to be a novel dynamic equilibrium of two parallel structures, one regular with two 1-nt loops and a 12-nt middle loop and another broken-strand with three 1-nt loops and a 11-nt middle loop; both structures adopt a novel hairpin (stem-loop duplex) conformation in the long loop. The dynamic equilibrium of two closely related structures and the unique hairpin loop conformation are specific to the P1G4 sequence and distinguish the P1G4 quadruplex from other parallel structures. The presence of P1G4 and Pu39 in adjacent regions of the BCL2 P1 promoter suggests a mechanism for precise regulation of BCL2 gene transcription. The unique P1G4 G-quadruplex may provide a specific target for small molecules to modulate BCL2 gene transcription.

TOC image

Corresponding Author: yang@pharmacy.arizona.edu.

SUPPORTING INFORMATION.

The effects of TMPyP4 and TMPyP2 on luciferase activities, Oligonucleotides for luciferase assay, CD spectra of P1G4 and P1G4KO, Non-denaturing EMSA gels of P1G4 and P1G4T, 1D ^{1}H NMR of P1G4T, selected imino proton assignments of P1G4, CD spectrum of P1G4T_G8T, and NMR binding of NSC 59276 to P1G4T and the effect of NSC 59276 on the P1G4-driven luciferase activities. The Supporting Information is available free of charge via the Internet at http://pubs.acs.org.



INTRODUCTION

The B-cell lymphoma-2 (BCL2) protein belongs to the BCL2 family of proteins and plays an essential role in the regulation of programmed cell death, or apoptosis.[1, 2] The abnormal overexpression of the BCL2 protein is linked to a large number of cancers.[3–7] Moreover, elevated levels of BCL2 are found to promote resistance to chemotherapy and gamma radiation.[8, 9] Therefore, BCL2 is considered to be an attractive target for cancer therapeutics. The most intensively pursued strategy is inhibiting BCL2 interactions with BH3-only proteins by small molecules[10, 11]; a BCL2–selective inhibitor, ABT-199, was recently developed to avoid thrombocytopenia caused by bcl-X_L inhibition and in Phase III clinical trials.[12] However, the dynamic nature of protein-protein interactions and acquired resistance present challenges for BCL2 inhibitors.[13–15] Modulation of BCL2 at the transcriptional level presents a compelling strategy for cancer therapeutics.

The major P1 promoter of the human BCL2 gene is located in the untranslated first exon, 1386 to 1432 base pairs upstream of the translation start site.[3, 16] It is a TATA-less, GC-rich promoter with multiple transcription start sites (-1394, -1399, -1406, -1410, and -1432) and is positioned in proximity to a nuclease hypersensitive site[16, 17] (Figure 1A). We have previously identified a 39-base-pair GC-rich region located 1489 to 1451 base pairs upstream of the translation start site (Figure 1A), whose G-rich strand (Pu39) can form two interchangeable G-quadruplex (G4) structures, a hybrid-type G-quadruplex[18, 19] and a parallel G-quadruplex with a 13-nt middle loop.[20] Stabilization of the BCL2 G-quadruplex by quindoline derivatives was reported to decrease mRNA and protein levels of BCL2 and lead to apoptosis in HL-60 cells.[21] Also reported is a G-rich sequence located 176 bp upstream of the P1 promoter which can form a stable G-quadruplex structure in the presence of peptide nucleic acid (PNA) molecules.[22]

However, in our functional study of the BCL2 P1 promoter activity using a luciferase reporter system in tumor cells, we found that the construct with a complete G-quadruplex-knock-out Pu39 mutant was still affected by G4-interactive compounds. Careful examination of the BCL2 P1 promoter sequence revealed a 28-bp GC-rich region immediately upstream of the BCL2 P1 promoter (-1,439 to -1,412 bp) (P1G4, Figure 1A). Previous studies have shown that deletion of this promoter region increased the promoter activity more than 2-fold.[23] Several transcription factors have been reported or suggested to bind to this region, such as SP1 and AP2.[16] Herein we reported that a stable G-quadruplex can readily form in the P1G4 sequence under physiological salt condition and that the new P1G4 functions as a

transcription repressor. The P1G4 G-quadruplex appears to be a dynamic equilibrium of two parallel structures, one regular with two 1-nt loops and a 12-nt middle loop and another broken-strand with three 1-nt loops and a 11-nt middle loop; both structures adopt a novel hairpin (stem-loop duplex) structure in the long loop. The unique P1G4 G-quadruplex with a hairpin loop may provide a specific recognition site for small molecules. P1G4 and previously identified Pu39 G-quadruplexes appear to form independently in adjacent regions. In the extended BCL2 P1 promoter region containing both Pu39 and P1G4, P1G4 appears to play a more dominant role in repressing the transcriptional activity. The novel P1G4 quadruplex in the 28-bp GC-rich element immediately upstream of the BCL2 P1 promoter that acts as a transcriptional repressor adds to our understanding of BCL2 transcriptional control, and to the diversity of parallel G-quadruplexes.

RESULTS

BCL2 P1G4 functions as a transcriptional repressor

To determine the biological function of P1G4, a P1G4WT construct was created by inserting the (-1,390 to -1,444 bp) human BCL2 P1 promoter sequence in a promoter-less pGL4.13 vector upstream of the firefly luciferase reporter gene (Figure 1B). The SV40 promoter in pGL4.13 was deleted, as it was found to contain a GC-rich region that can be affected by G-quadruplex-interactive compounds. The Pu39 region was intentionally excluded to isolate the biological effect of P1G4 on gene transcription. We also generated a control construct with a mutated P1G4KO sequence that was not able to form G-quadruplex (Figure 1B). MCF-7 cells were used because they express high levels of BCL2.[24] Our results showed that the luciferase activity of the construct containing the P1G4KO sequence is more than 2.5-fold greater than that with the wild-type BCL2 P1G4 promoter sequence (Figure 1C left), indicating that the P1G4 functions as a transcriptional repressor.

We then carried out luciferase assays using P1G4WT and P1G4KO constructs in the presence of 5 μM TMPyP4, a G-quadruplex-interactive compound, and its positional isomer, TMPyP2, a poor G-quadruplex-interactive compound (Figure 1E). In the P1G4WT construct, TMPyP4 showed a pronounced inhibitory effect on the promoter activity whereas TMPyP2 showed a much weaker inhibition (Figure 1C middle, Supplemental Table 1). In contrast, in the P1G4KO construct, TMPyP4 showed a much-reduced inhibitory effect on the promoter activity (Figure 1C right, Supplemental Table 1). These results indicated that the inhibitory effect of P1G4, but not P1G4KO, can be markedly enhanced by the G-quadruplex-interactive compound TMPyP4.

To examine the function of both P1G4 and Pu39 in the BCL2 P1 promoter, we further generated a construct with an extended human BCL2 P1 promoter sequence that contains both Pu39 and P1G4 regions in the pGL4.13 vector (Dual-WT, Figure 1B). We then generated knock-out constructs with the Pu39KO sequence that was not able to form Pu39 G-quadruplex, or the P1G4KO sequence that was not able to form P1G4 G-quadruplex, or a double-knock-out sequence that was not able to form either P1G4 or Pu39 G-quadruplex, respectively (**Dual-Pu39KO, Dual-P1G4KO, Dual-KOKO**, Figure 1B). The luciferase assays in MCF-7 cells showed that the luciferase activity of the construct containing the P1G4KO sequence (Dual-P1G4KO) is again more than 2-fold greater than that with the

wild-type BCL2 promoter sequence Dual-WT (Figure 1D left), in agreement with the previous data that the P1G4 functions as a transcriptional repressor. Interestingly, the luciferase activity of the construct containing the Pu39KO sequence is about 50% of that with the wild-type BCL2 promoter sequence (Figure 1D left), suggesting that Pu39 may function as a transcriptional activator. When both Pu39 and P1G4 quadruplexes were knocked out (Dual-KOKO, Figure 1B), the luciferase activity was more than 1.5-fold greater than that with the wild-type BCL2 promoter sequence Dual-WT (Figure 1D left), indicating that P1G4 plays a more dominant role in repressing the transcriptional activity. The luciferase assays in the presence of TMPyP4 and TMPyP2 showed that TMPyP4 induced a clear inhibitory effect on the promoter activity whereas TMPyP2 did not appear to affect the promoter activity (Figure 1D right, Supplemental Table 1), indicating that the G-quadruplex-interactive compound TMPyP4 can markedly enhance the inhibitory effect of P1G4 on gene transcription. As TMPyP4 is known to be a nonselective G-quadruplex-interactive compound, it can stabilize both the Pu39 and P1G4 quadruplexes. The inhibitory effect of TMPyP4 was shown to be more pronounced in the Pu39KO construct where P1G4 is intact (Figure 1D right, Supplemental Table 1), again suggesting that P1G4 quadruplex plays a more dominant role in repressing the transcriptional activity in the BCL2 P1 promoter.

P1G4 forms a stable G-quadruplex under physiologically relevant ionic conditions

To confirm that the P1G4 forms a stable G-quadruplex under physiologically relevant ionic conditions, we carried out dimethyl sulfate (DMS) footprinting experiments. The guanine N7 in a G-tetrad is involved in Hoogsten hydrogen bonding (Figure 1E) and is protected from DMS methylation and piperidine cleavage, in contrast to the guanines in single- or double-stranded DNA.[19] Hence, DMS footprinting provides a useful tool to determine the tetrad guanines of a G-quadruplex. The P1G4 sequence consists of five runs of three consecutive guanines (Figure 2A). The results of DMS footprinting analysis on the wild-type P1G4 in the absence and presence of 140 mM K^+ at pH 7 are shown in Figure 2B left. G-runs I, II, IV and V, each consisting of three consecutive guanines, appeared to be protected from DMS methylation. G8 appeared to be less protected from cleavage compared to other guanines in G-run II. G10, G12, G14, and G15 showed complete cleavage, indicating that these guanines are not involved in G-tetrad formation. G-run III appeared to be cleaved; the mutant P1G4 GrunIII-G/T sequence (Figure 2A), in which all guanines in G-run III were mutated to thymine, showed a similar DMS footprinting cleavage pattern to that of the wild-type BCL2 P1G4 in the presence of 140 mM K^+, but with G8 more protected in the mutant sequence (Figure 2B). As a control, we performed DMS footprinting on a complete knock-out mutant sequence P1G4-KO (Figure 2A), in which the middle guanine of each G-run was mutated to a thymine residue. The result showed complete cleavage at each guanine position, both in the absence and presence of K^+ (Figure 2B), indicating G-quadruplex was not able to form in the P1G4-KO sequence. The disruption of the P1G4 quadruplex in P1G4-KO was also supported by CD data (Figure S1).

We also performed DMS footprinting on a 94-nt extended BCL2 promoter sequence that contains both Pu39 and P1G4, in the absence and presence of 140 mM K^+ at pH 7 (Figure 3). Even in the presence of Pu39, G-quadruplexes were clearly formed in the P1G4 segment (Figure 3 left). The DMS protection pattern is similar to that observed for the P1G4 itself,

but with more protection on G-run III in the extended 94-nt sequence. On the other hand, the DMS cleavage pattern of the Pu39 segment (Figure 3 right) is also similar to what was previously observed for Pu39 itself.[20] Therefore, Pu39 and P1G4 G-quadruplexes appeared to form independently in the extended BCL2 promoter sequence.

The P1G4 G-quadruplex is a dynamic equilibrium of two parallel structures, both containing a stem-loop hairpin in the middle loop

To determine the folding structure of the major G-quadruplex formed in the BCL2 P1G4 sequence, we performed NMR studies in combination with CD and EMSA. The P1G4 with five runs of three consecutive guanines (Figure 4A) can form multiple G-quadruplexes. 1D ^{1}H NMR spectrum of P1G4 in 50 mM K^+ solution showed sharp peaks at 10.5–12 ppm, characteristic of imino protons of tetrad guanines (Figure 4B top). However, higher order structures appeared to form in P1G4 as indicated by elevated baselines.

The tendency for the formation of higher-order G-quadruplexes has been observed in other G-rich quadruplex-forming sequences.[20, 25] The P1G4T sequence (Figure 4A), with thymine at both ends, gave rise to improved NMR spectral quality with decreased higher-order structures (Figure 4B middle). This was confirmed by EMSA experiments: P1G4 showed clear formation of higher-order structures, which is increased at higher K^+ concentration, while P1G4T showed more predominant formation of the monomeric structure (Figure S2). The CD spectra of P1G4 and P1G4T are almost identical, showing a positive peak at 264 nm and a negative peak at 240 nm in 50 mM K^+ solution at pH 7 (Figure 4C), suggesting a parallel G-quadruplex.[26] DMS footprinting of the monomeric structure of P1G4T showed a very similar cleavage pattern as compared to that of P1G4 (Figure 2B), suggestive of the same major monomeric G-quadruplex formation. The T_m of P1G4T in 50 mM K^+ solution at pH 7 was determined by CD melting experiment to be ~ 68.5 °C (Table 1), consistent with the variable temperature experiment by 1D ^{1}H NMR (Figure 4D). The T_m of P1G4T appears to be concentration-independent (Table 1), in agreement with a monomeric structure. To achieve better NMR spectral quality at a higher DNA concentration, 25 mM K^+ was used for further NMR analyses (Figure S3).

We prepared site-specifically labeled P1G4T oligonucleotides with 8% ^{15}N-guanine for the detection of imino protons (H1) of each guanine (Figure 5). The H1 protons of guanines involved in the G-tetrad connected with Hoogsteen hydrogen bonding have a characteristic chemical shift in the region of 10.5–12 ppm, as compared to 13–14 ppm for guanines involved in Watson-Crick hydrogen bonding.[27] The imino protons of guanines that are not involved in hydrogen bonding interactions cannot be seen in ^{1}H NMR due to fast exchange with bulk H_2O. The H1 proton of guanine is one-bond connected to N1 (Figure 1E) and can be readily detected for the labeled guanine by 1D ^{15}N-edited NMR experiments.[28] Of the 19 guanines on P1G4T, imino protons of G2–G4, G6–G8, G21–G23, and G25–G27 were observed in the 10.5–12 ppm region (Figure 5), indicating the involvement of these guanines in the tetrad formation, consistent with the DMS footprinting data (Figure 2B). Combined with the CD data, the P1G4T G-quadruplex appears to be a three-tetrad parallel-stranded structure involving G-runs I, II, IV, and V, with 1-nt first and third loops and a 12-nt middle loop (P1G4A, Figure 6 left). Interestingly, the imino protons of G10 and G15 were shown to

be around 13 ppm (Figure 5), characteristic of Watson-Crick base pairing, indicating a stem-loop duplex conformation adopted by the 12-nt middle loop (Figure 6). Although the ^{1}H NMR spectrum of P1G4 is not as well resolved as P1G4T, we were able to detect several guanine imino protons in ^{15}N-G-labeled P1G4 (Figure S4), which were shown to be at the same positions as those in P1G4T, suggesting the same major G-quadruplex conformation.

Surprisingly, G19 of G-run III showed a clear imino proton peak at ~ 11 ppm (Figure 5), indicating the potential involvement of G19 in a tetrad formation. On the other hand, of the three guanines on G-run II, the imino proton peak of G8 appeared to be weaker than those of G6 and G7 (Figure 5), suggestive of a more dynamic conformation of G8. The dynamic nature of G8 was also shown in the DMS footprinting data, in that G8 is less protected from DMS cleavage than the other two guanines on G-run II (Figure 2B). Upon the G-to-T mutation of run-III including G19, the partial cleavage of G8 disappeared, indicating complete involvement of G8 in tetrad formation (Figure 2B). It thus appeared that a second G-quadruplex structure forms in P1G4T, where G8 of the bottom tetrad is substituted by G19 resulting in a parallel structure containing a broken strand, with three 1-nt loops and one 11-nt loop (P1G4B, Figure 6 right). Indeed, P1G4T with G8-to-T mutation (P1G4T_G8T) shows a ^{1}H NMR spectrum similar to that of P1G4T (Figure 4B bottom), and a very similar CD spectrum (Figure S5), indicative of a similar major G-quadruplex formation. The two parallel G-quadruplexes differed only in the bottom G-tetrad and appeared to be in dynamic equilibrium (Figure 6), which may provide stability to the major P1G4T G-quadruplex. The same Watson-Crick imino peaks around 13 ppm were observed in P1G4T_G8T, indicating the stem loop conformation also formed in the P1G4B structure with a broken strand (Figure 6 right). While G18 did not show an imino proton detectable in the 10.5–12 ppm region, G17 of G-run III showed a weak peak around 10.5 ppm, suggesting that G17 may be involved in a hydrogen-bonded capping structure below the 3′ G-tetrad. Depending on the hydrogen-bonding conformation, the imino proton of a guanine residue involved in a capping structure can have a chemical shift around 10.5 ppm, as shown in the major G-quadruplex formed in the VEGF promoter.[29] The NMR variable temperature study showed that, while the P1G4 core quadruplex melting temperature was about 65 °C, the melting temperature was around 45 °C for the stem-loop structure of P1G4, and below 35 °C for G17 and thus the 3′ capping structure (Figure 4D). With only two detectable Watson-Crick base pairs in the stem-loop hairpin structure of P1G4, a melting temperature of 45 °C is quite remarkable.

It has been recently reported that a pregnanol derivative, NSC 59276 (IMC-76), can bind to the flexible hairpin to shift the dynamic equilibrium between the alternative hairpin and the i-motif formed in the BCL2 P1 promoter.[30] We are interested in finding out whether NSC 59276 has an effect on the P1G4 quadruplex-induced transcriptional repression. We first tested the binding of NSC 59276 with the P1G4 quadruplex by NMR. NSC 59276 did not appear to specifically stabilize the hairpin loop within the P1G4 structure (Figure S6A). It did not appear to stabilize the P1G4 quadruplex either, as shown by CD melting of P1G4T in the presence of 4 equivalence of NSC 59276, whose T_m is similar to that of free P1G4T DNA. We then examined the effect of NSC 59276 on the luciferase activities of P1G4-WT and P1G4KO constructs. Consistently, NSC 59276 did not show a clear inhibitory effect on the promoter activity in either construct (Figure S6B).

DISCUSSION

It is significant that a stable G-quadruplex is found to form in the P1G4 sequence immediately upstream of the P1 promoter and act as a transcriptional repressor. P1G4 is only 13 nt away from the previously identified Pu39 region, with both regions capable of forming independent G-quadruplexes in the BCL2 P1 promoter sequence. The presence of stable G-quadruplexes in adjacent regions could be important for the precise regulation of BCL2 gene transcription. The major P1 promoter contains multiple transcription start sites (-1394, -1399, -1406, -1410, and -1432); depending on which transcription start site is used, different G4 may be formed. For example, transcription from all start sites except for -1432 allows formation of P1G4 quadruplex, but transcription from the start site -1432 can only permit Pu39 G-quadruplex formation (Figure 1A). In addition, as the transcription-associated negative superhelicity is the driving force for the G-quadruplex formation in proximal promoter regions, the formation of P1G4 and Pu39 G-quadruplexes likely depends on the level of transcriptional activity.

Surprisingly, using luciferase assays in MCF-7 cells, we found that in the extended BCL2 P1 promoter region containing both P1G4 and Pu39, P1G4 quadruplex appears to play a major role in repressing transcriptional activity. The G-quadruplex-interactive compound, TMPyP4, also induced more pronounced inhibitory effect on the BCL2 P1 promoter activity through the P1G4 quadruplex. It has recently been shown that the complementary C-rich strand of Pu39 forms a i-motif, which is highly dynamic and exists in equilibrium with a flexible hairpin species.[30] Stabilization of the i-motif structure results in upregulation, whereas stabilization of the flexible hairpin results in downregulation of the BCL2 gene. Therefore, the effect of Pu39/Py39 region on the BCL2 gene transcription may be more complicated and related to both quadruplex and i-motif on the complementary strands.

Parallel-stranded structures with three tetrads are found to be common in the promoter G-quadruplexes, while most parallel-stranded promoter G-quadruplexes contain three chain-reversal loops with 1-nt first and last loops and a variable-length middle loop.[31] It is intriguing that the P1G4 G-quadruplex appears to be a dynamic equilibrium of two parallel structures, one broken-strand, both with a novel hairpin conformation adopted by the middle loop. DNA quadruplexes with duplex motifs have been shown to be quite tolerable.[32] Unlike two inter-changeable G-quadruplexes with distinct structures formed from Pu39, the two P1G4 conformations are very similar whose dynamic equilibrium may provide entropy and stability to the formation of the P1G4 G-quadruplex. This dynamic equilibrium of two closely related conformations and the unique hairpin loop are specific to the P1G4 sequence, distinguishing the P1G4 quadruplex from other parallel structures, and may be recognized by proteins or small molecules. The unique hairpin loop within the P1G4 quadruplex may provide a specific target for small molecule recognition.

MATERIALS AND METHODS

Oligonucleotides

Site-specific ^{15}N-labeled and unlabeled DNA oligonucleotides were synthesized using β-cyanoethylphosphoramidite solid phase chemistry (Applied Biosystem Expedite™ 8909) as

described previously.[18, 29, 33] ^{15}N-labeled guanine phosphoramidite was purchased from Cambridge Isotope Laboratories, Inc. (Andover, MA). 8% of ^{15}N-labeled guanine phosphoramidite was mixed with 92% of unlabeled guanine phosphoramidite in site-specifically ^{15}N-labeled DNA synthesis. The synthesized oligonucleotides were eluted from the column with a 50%:50% mixture of 40% methylamine:ammonia, heated for 10 min at 65°C, purified on reverse-phase Micropure II columns (BioSearch Technologies) and subjected to sequential dialysis through 10 mM NaOH, water, 150 mM NaCl, and water before lyophilization.

NMR Spectroscopy Study

NMR experiments were performed on a Bruker DRX-600 spectrometer. DNA oligonucleotides were prepared in either 50 mM or 25 mM K^+-containing K-phosphate buffer (pH 7) solution, with a total final concentration of 0.1–2.5 mM DNA. The samples were prepared in 10/90% D_2O/H_2O solution. Each sample was annealed by heating at 95°C for 5 min and cooling slowly to room temperature. 1D ^{1}H spectra were recorded at 25°C with 512 scans. One-dimensional ^{15}N-edited gradient HMQC experiments with site-specific ^{15}N-labeling were performed to identify guanine imino H1 and aromatic H8 protons. The water signal was suppressed using a Watergate pulse sequence.

DMS Footprinting

Gel-purified Bcl-2 oligonucleotides (with 7T flanking ends) 5′-end labeled using [γ-^{32}P] ATP in the presence of T4 polynucleotide kinase were denatured at 90°C for 5 min, then slowly cooled to room temperature in 10 mM Tris-HCl buffer (pH=7.6) in the presence or absence of 140 mM KCl. The samples were then methylated by treatment with 0.5% (final concentration) DMS, 1 μg calf thymus DNA for 7 minutes at room temperature. The reaction was stopped by the addition of β-mercaptoethanol and glycerol to final concentrations of 7.7% and 2.6%, respectively. In order to separate the single-stranded DNA, intramolecular G-quadruplex, and intermolecular G-quadruplexes, an 8% native PAGE gel was used. The DNA was recovered from the gel, ethanol-precipitated and subjected to cleavage with 10% piperidine for 16 minutes at 90 °C. A Speedvac was used to remove the piperidine and two successive water washes. The cleaved products (5500cpm/5ul) were analyzed on a 16% sequencing (denaturing) PAGE gel. Due to the length of the 94-nt sequence, the same DMS footprinting sample series was loaded on the same gel at two different times to allow the best separation of 5′ (Figure 3 left) and 3′ (Figure 3 right) bands.

Circular Dichroism (CD) Spectroscopy

A Jasco-810 spectropolarimeter (Jasco Inc, Easton, MD) equipped with a thermodynamically controlled cell holder was used to obtain circular dichroism spectra. 10 μM samples were prepared in 50 mM K^+-containing K-phosphate buffer. Measurements were taken in a 1 mm optical path length quartz cell. The CD spectra were three averaged scans between 230–330 nm at 25°C after application of a baseline correction to remove signal contributions from the buffer. CD melting experiments to obtain T_m values were performed at 264 nm with a heating rate of 2°C/min and 1 s response time between 25–93°C.

Preparation of Promoter Reporter Constructs

The P1G4 firefly luciferase construct was generated by inserting the human bcl-2 P1 promoter DNA sequence (-1,390 to -1,444 bp), containing the P1G4 sequence upstream of the firefly luciferase reporter gene in a promoter-less pGL4.13 vector in the same orientation as in the human bcl-2 gene (Figure 1B), using an oligonucleotide cassette cloning approach. The SV40 promoter normally present in pGL4.13 was deleted, as we found that the SV40 promoter contains a GC-rich region that can be affected by G-quadruplex-interactive small molecules. The control constructs (KO) were generated with a mutated G-quadruplex sequence that was not able to form G-quadruplex. To analyze the respective function of Pu39G4 and P1G4, we generated several other pGL4.13 –SV40 luciferase constructs containing Pu39G4-P1G4 (Dual-WT), Pu39KO-P1G4 (Dual-Pu39KO), Pu39G4-P1G4KO (Dual-P1G4KO), and Pu39KO-P1G4KO (Dual-KOKO) sequences (Figure 1B). Most dual constructs have mutations to generate a BssHII site between Pu39 and P1G4 sequences to allow a cassette approach to construction. No significant difference in promoter activities was observed between wild-type and BssHII constructs.

For all promoter constructs, complementary custom DNA oligonucleotide pairs (Sigma, see Supplemental Table 2) at equimolar ratios were 5′-phosphorylated using T4 polynucleotide kinase (New England Biolabs, NEB) with T4 DNA ligase buffer (NEB) for 2 hr at 37°C, then heated at 95°C for 5 min and allowed to cool slowly, inactivating enzymes and annealing the strands to provide appropriate sticky ends. For double inserts, resulting equimolar dsDNA oligonucleotide pairs were ligated with T4 DNA ligase for 16 hr at room temperature (RT). pGL4.13 –SV40 plasmid was digested with SacI and XhoI, and the resulting ends were dephosphorylated with Antarctic phosphatase (NEB) for 2 hr at 37°C. The cut plasmid was then isolated via QiaQuick PCR purification kit. T4 DNA ligase (NEB) was used to ligate dsDNA oligonucleotide insert with cut plasmid at RT for 16 hr. Successful clones were verified by sequencing.

Transfection and Luciferase Assay

MCF-7 cells (8×10^4) were cultured in 500μl RPMI including 10% FBS and 1% penicillin/streptomycin antibiotics overnight using a 24 well plate at 37°C, 5% CO_2. Cell viability was assessed (>90%) by trypan blue prior to experiments. 2 μM of each plasmid construct was transfected into MCF-7 cells using Fugene HD (Promega). A Renilla luciferase plasmid, pRL-TK, was co-transfected with each construct for normalization. Transfection was performed in OptiMem serum-free medium, which was replaced by 10% FBS, 1% penicillin/streptomycin-supplemented RPMI after a five-hour incubation at 37°C, 5% CO_2. After a further 24-hour incubation, cells were lysed with Passive Lysis Buffer, and Firefly and Renilla luciferase activities were measured using the Dual-Luciferase Reporter assay kit (Promega).

For drug studies, MCF-7 cells were transfected with the constructs and treated with TMPyP4 and TMPyP2 5 hours post-transfection at 5 μM, lower than their IC_{50} values.[34] At 24 hours after small molecule compound treatment, cells were lysed and tested for Firefly and Renilla luciferase activity using the Dual-Luciferase Reporter assay kit (Promega). Luciferase activities expressed as relative luciferase units (RLUs) were normalized to total

protein concentrations. Each transfection was run in triplicate or duplicate in each of three independent experiments.

Native Gel Electrophoresis

Electrophoresis experiments were performed with a 1.5mm thick 10×7 cm native gel containing 16% acrylamide (Acrylamide: Bis-acrylamide=29:1) in TBE buffer, pH=8 supplemented with12.5mM KCl. Each sample contains 0.4 nmol DNA. DNA bands were visualized using ethidium bromide staining.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research is supported by the National Institutes of Health (GM083117 and CA177585 to DY). We thank Clement Lin for proofreading the manuscript.

References

1. Chao DT, Korsmeyer SJ. Annual Review of Immunology. 1998; 16:395–419.
2. Letai A, Bassik MC, Walensky LD, Sorcinelli MD, Weiler S, Korsmeye SJ. Cancer Cell. 2002; 2:183–192. [PubMed: 12242151]
3. Tsujimoto Y, Croce CM. Proceedings of the National Academy of Sciences of the United States of America. 1986; 83:5214–5218. [PubMed: 3523487]
4. Marschitz I, Tinhofer I, Hittmair A, Egle A, Kos M, Greil R. American Journal of Clinical Pathology. 2000; 113:219–229. [PubMed: 10664624]
5. Pezzella F, Turley H, Kuzu I, Tungekar MF, Dunnill MS, Pierce CB, Harris A, Gatter KC, Mason DY. N Engl J Med. 1993; 329:690–694. [PubMed: 8393963]
6. McDonnell TJ, Troncoso P, Brisbay SM, Logothetis C, Chung LW, Hsieh JT, Tu SM, Campbell ML. Cancer Research. 1992; 52:6940–6944. [PubMed: 1458483]
7. Joensuu H, Pylkkanen L, Toikkanen S. American Journal of Pathology. 1994; 145:1191–1198. [PubMed: 7977649]
8. Reed JC, Kitada S, Takayama S, Miyashita T. Annals of Oncology. 1994; 5(Suppl 1):61–65. [PubMed: 8172820]
9. Renault TT, Chipuk JE. Ann N Y Acad Sci. 2013; 1285:59–79. [PubMed: 23527542]
10. Oltersdorf T, Elmore SW, Shoemaker AR, Armstrong RC, Augeri DJ, Belli BA, Bruncko M, Deckwerth TL, Dinges J, Hajduk PJ, Joseph MK, Kitada S, Korsmeyer SJ, Kunzer AR, Letai A, Li C, Mitten MJ, Nettesheim DG, Ng S, Nimmer PM, O'Connor JM, Oleksijew A, Petros AM, Reed JC, Shen W, Tahir SK, Thompson CB, Tomaselli KJ, Wang B, Wendt MD, Zhang H, Fesik SW, Rosenberg SH. Nature. 2005; 435:677–681. [PubMed: 15902208]
11. Tse C, Shoemaker AR, Adickes J, Anderson MG, Chen J, Jin S, Johnson EF, Marsh KC, Mitten MJ, Nimmer P, Roberts L, Tahir SK, Xiao Y, Yang X, Zhang H, Fesik S, Rosenberg SH, Elmore SW. Cancer Res. 2008; 68:3421–3428. [PubMed: 18451170]
12. Souers AJ, Leverson JD, Boghaert ER, Ackler SL, Catron ND, Chen J, Dayton BD, Ding H, Enschede SH, Fairbrother WJ, Huang DCS, Hymowitz SG, Jin S, Khaw SL, Kovar PJ, Lam LT, Lee J, Maecker HL, Marsh KC, Mason KD, Mitten MJ, Nimmer PM, Oleksijew A, Park CH, Park CM, Phillips DC, Roberts AW, Sampath D, Seymour JF, Smith ML, Sullivan GM, Tahir SK, Tse C, Wendt MD, Xiao Y, Xue JC, Zhang H, Humerickhouse RA, Rosenberg SH, Elmore SW. Nat Med. 2013; 19:202–208. [PubMed: 23291630]
13. Zinzalla G, Thurston DE. Future Medicinal Chemistry. 2009; 1:65–93. [PubMed: 21426071]

14. Choudhary GS, Al-harbi S, Mazumder S, Hill BT, Smith MR, Bodo J, Hsi ED, Almasan A. Cell Death and Disease. 2015
15. Juin P, Geneste O, Gautier F, Depil S, Campone M. Nat Rev Cancer. 2013; 13:455–465. [PubMed: 23783119]
16. Seto M, Jaeger U, Hockett RD, Graninger W, Bennett S, Goldman P, Korsmeyer SJ. EMBO Journal. 1988; 7:123–131. [PubMed: 2834197]
17. Young RL, Korsmeyer SJ. Mol Cell Biol. 1993; 13:3686–3697. [PubMed: 8388542]
18. Dai J, Dexheimer TS, Chen D, Carver M, Ambrus A, Jones RA, Yang DZ. J Am Chem Soc. 2006; 128:1096–1098. [PubMed: 16433524]
19. Dexheimer TS, Sun D, Hurley LH. J Am Chem Soc. 2006; 128:5404–5415. [PubMed: 16620112]
20. Agrawal P, Lin C, Mathad RI, Carver M, Yang D. J Am Chem Soc. 2014; 136:1750–1753. [PubMed: 24450880]
21. Wang XD, Ou TM, Lu YJ, Li Z, Xu Z, Xi C, Tan JH, Huang SL, An LK, Li D, Gu LQ, Huang ZS. J Med Chem. 2010; 53:4390–4398. [PubMed: 20481493]
22. Onyshchenko MI, Gaynutdinov TI, Englund EA, Appella DH, Neumann RD, Panyutin IG. Nucleic Acids Res. 2009; 37:7570–7580. [PubMed: 19820116]
23. Heckman C, Mochon E, Arcinas M, Boxer LM. Journal of Biological Chemistry. 1997; 272:19609–19614. [PubMed: 9235968]
24. Beck MT, Peirce SK, Chen WY. Oncogene. 2002; 21:5047–5055. [PubMed: 12140755]
25. Chen Y, Agrawal P, Brown RV, Hatzakis E, Hurley L, Yang D. J Am Chem Soc. 2012; 134:13220–13223. [PubMed: 22866911]
26. Hatzakis E, Okamoto K, Yang D. Biochemistry. 2010; 49:9152–9160. [PubMed: 20849082]
27. Zhang Z, Dai J, Veliath E, Jones RA, Yang DZ. Nucl Acids Res. 2010; 38:1009–1021. [PubMed: 19946019]
28. Szewczak AA, Kellogg GW, Moore PB. FEBS Letters. 1993; 327:261–264. [PubMed: 7688695]
29. Agrawal P, Hatzakis E, Guo K, Carver M, Yang D. Nucleic Acids Res. 2013; 41:10584–10592. [PubMed: 24005038]
30. Kendrick S, Kang HJ, Alam MP, Madathil MM, Agrawal P, Gokhale V, Yang D, Hecht SM, Hurley LH. J Am Chem Soc. 2014; 136:4161–4171. [PubMed: 24559410]
31. Onel B, Lin C, Yang D. Sci China Chem. 2014; 57:1605–1614. [PubMed: 27182219]
32. Lim KW, Phan AT. Angewandte Chemie (International ed in English). 2013; 52:8566–8569. [PubMed: 23794476]
33. Dai J, Chen D, Jones RA, Hurley LH, Yang D. Nucleic Acids Res. 2006; 34:5133–5144. [PubMed: 16998187]
34. Izbicka E, Wheelhouse RT, Raymond E, Davidson KK, Lawrence RA, Sun DY, Windle BE, Hurley LH, Von Hoff DD. Cancer Research. 1999; 59:639–644. [PubMed: 9973212]

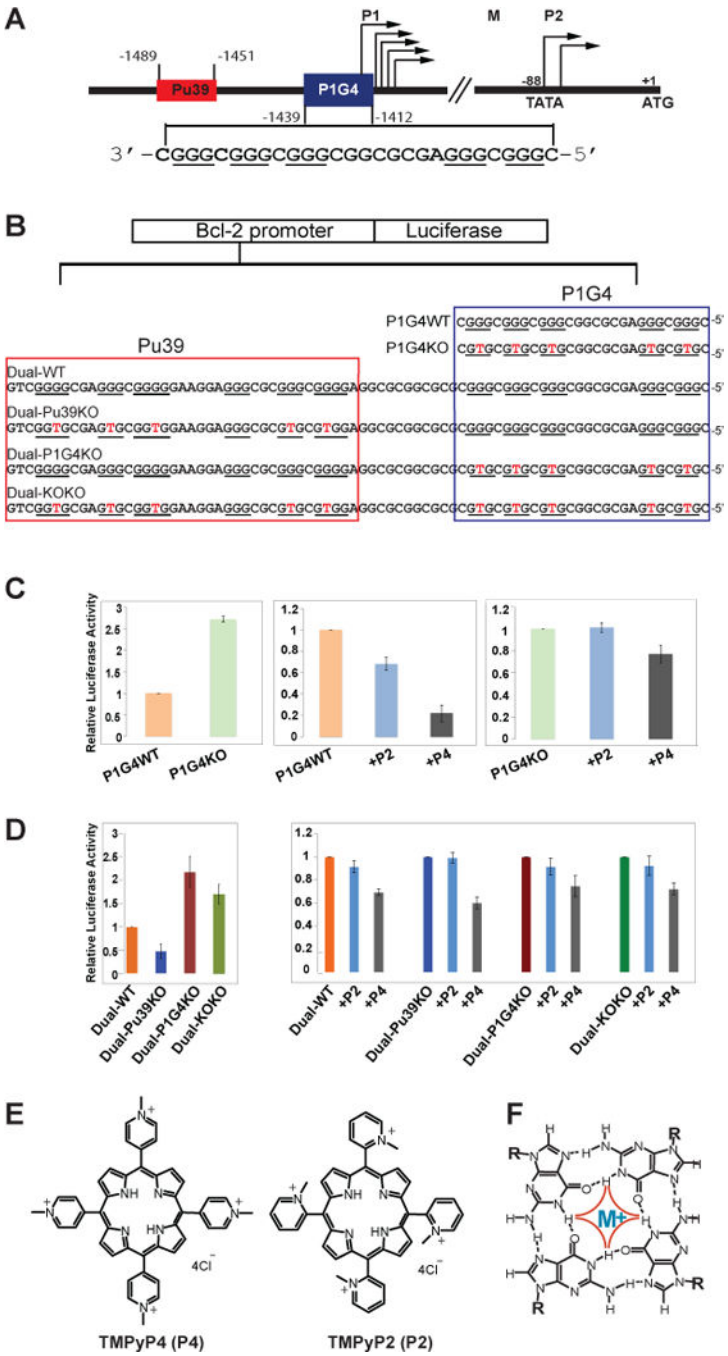


Figure 1.

(A) Promoter structure of the human *BCL2* gene. Shown in the inset is the P1G4 sequence with guanine runs of the purine-rich strand underlined. Transcriptional start sites of the P1 promoter are indicated using arrows. (B) P1G4-WT and P1G4KO constructs, as well as Dual-WT (containing both Pu39 and P1G4), Dual-Pu39KO (Pu39 G-quadruplex knock-out), Dual-P1G4KO (P1G4 G-quadruplex knock-out), and Dual-KOKO (double knock-out) used for promoter-driven luciferase assay. P1G4 and Pu39 sequences are enclosed by blue and red boxes, respectively. G-runs are underlined; the mutations in the G-quadruplex knock-out

constructs are shown in red. (C) Luciferase activity of P1G4-WT and P1G4KO constructs in the absence and presence of TMPyP4 and TMPyP2 at 5uM in MCF-7 cells. Each data point is the average of three independent experiments. (D) Luciferase activity of Dual-WT, Dual-Pu39KO, Dual-P1G4KO, and Dual-KOKO constructs in the absence and presence of 5uM TMPyP4 and TMPyP2 in MCF-7 cells. Each data point is the average of three independent experiments. (E) Structures of TMPyP4 and TMPyP2. (F) Structure of G-tetrad.

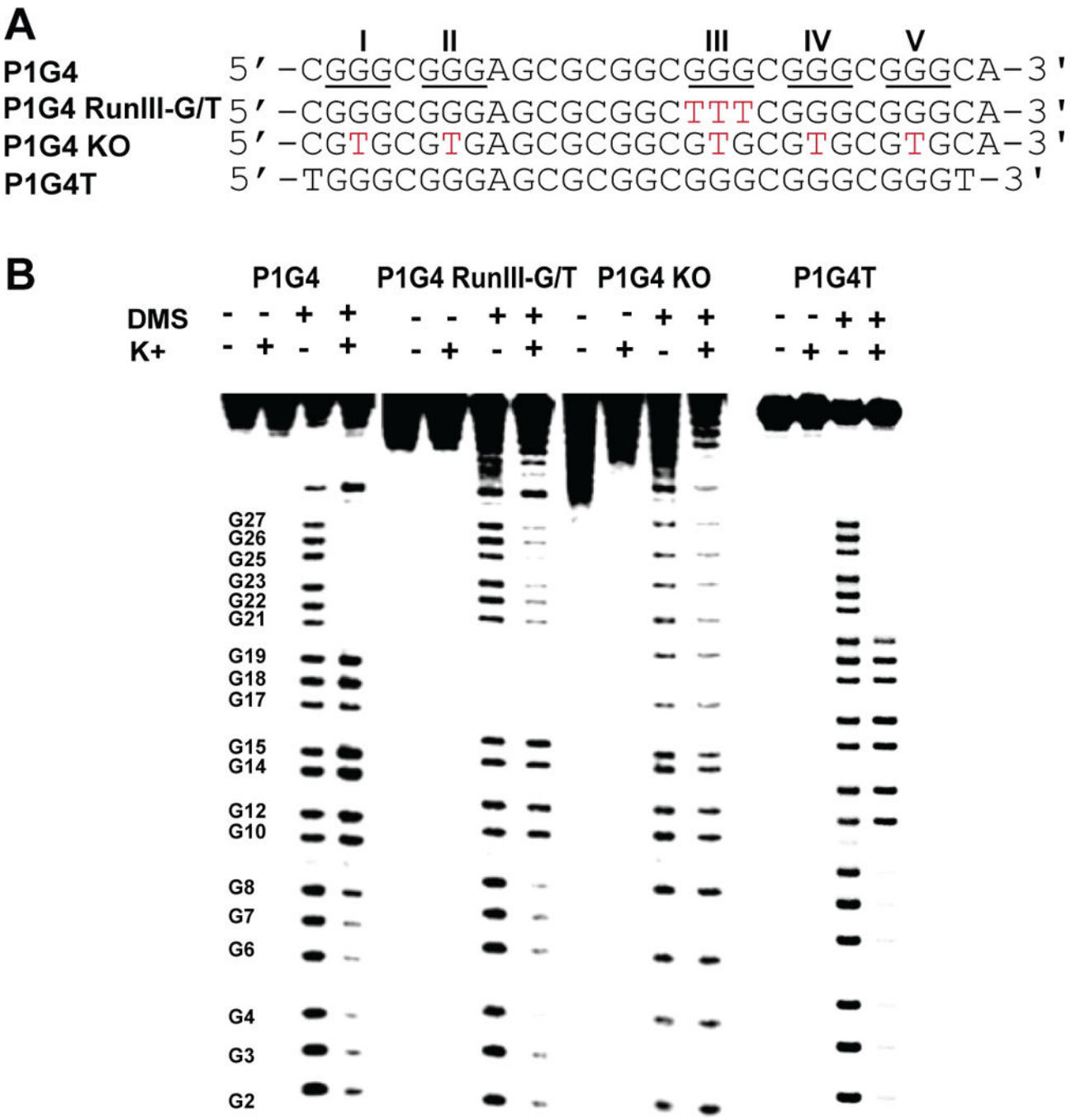


Figure 2.

(A) The promoter sequence of the BCL2 P1G4 and its modifications. The top sequence is the 29-mer wild-type G-rich sequence (P1G4). The five G-runs are underlined and numbered. P1G4 RunIII-G/T and P1G4 KO are modified P1G4 sequences with G to T mutations shown in red. P1G4T is P1G4 sequence with a flanking T at both ends. (B) DMS footprinting of P1G4, P1G4 RunIII-G/T, P1G4 KO, and P1G4T.

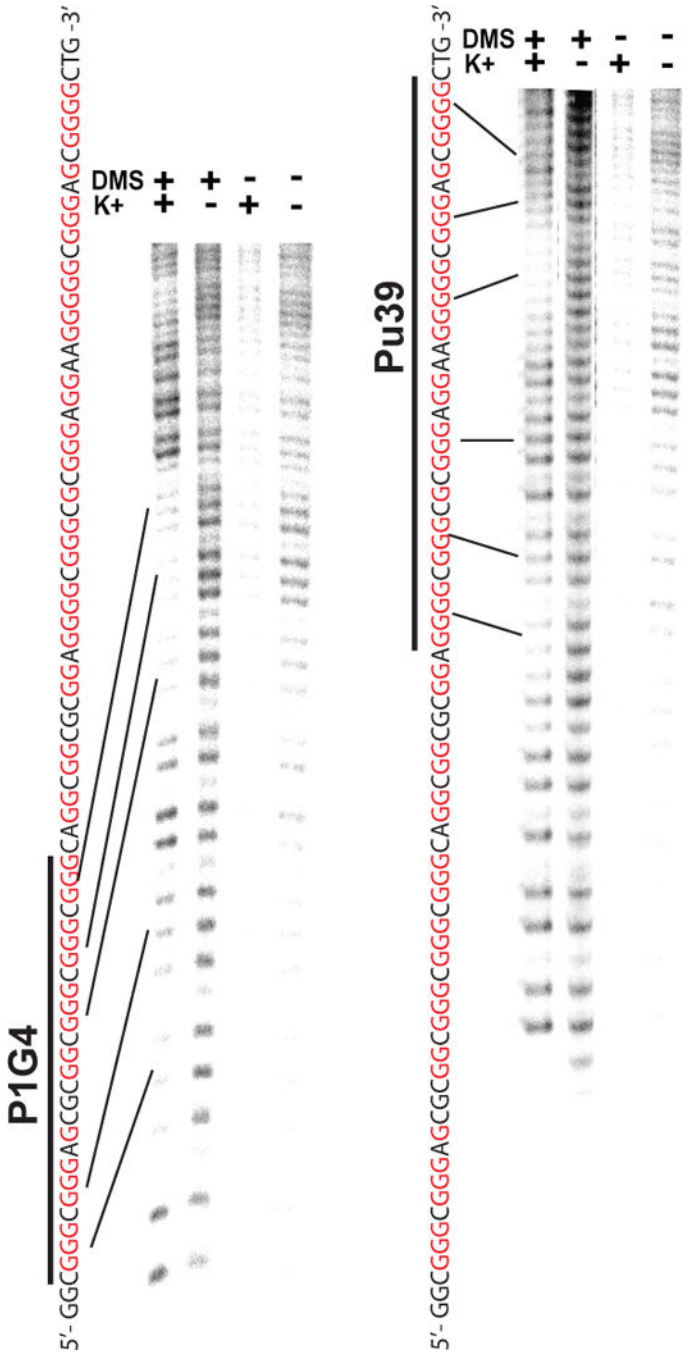


Figure 3.

DMS footprinting of the 94-nt extended BCL2 promoter sequence that contains both P1G4 and Pu39. Due to the length of the 94-nt sequence, the same DMS footprinting sample was loaded on the same gel at two different times to allow the best separation of 5′ (P1G4, left) and 3′ (Pu39, right) bands. Shown on the left of the two footprinting results is the 94-nt BCL2 promoter sequence with P1G4 and Pu39 labeled.

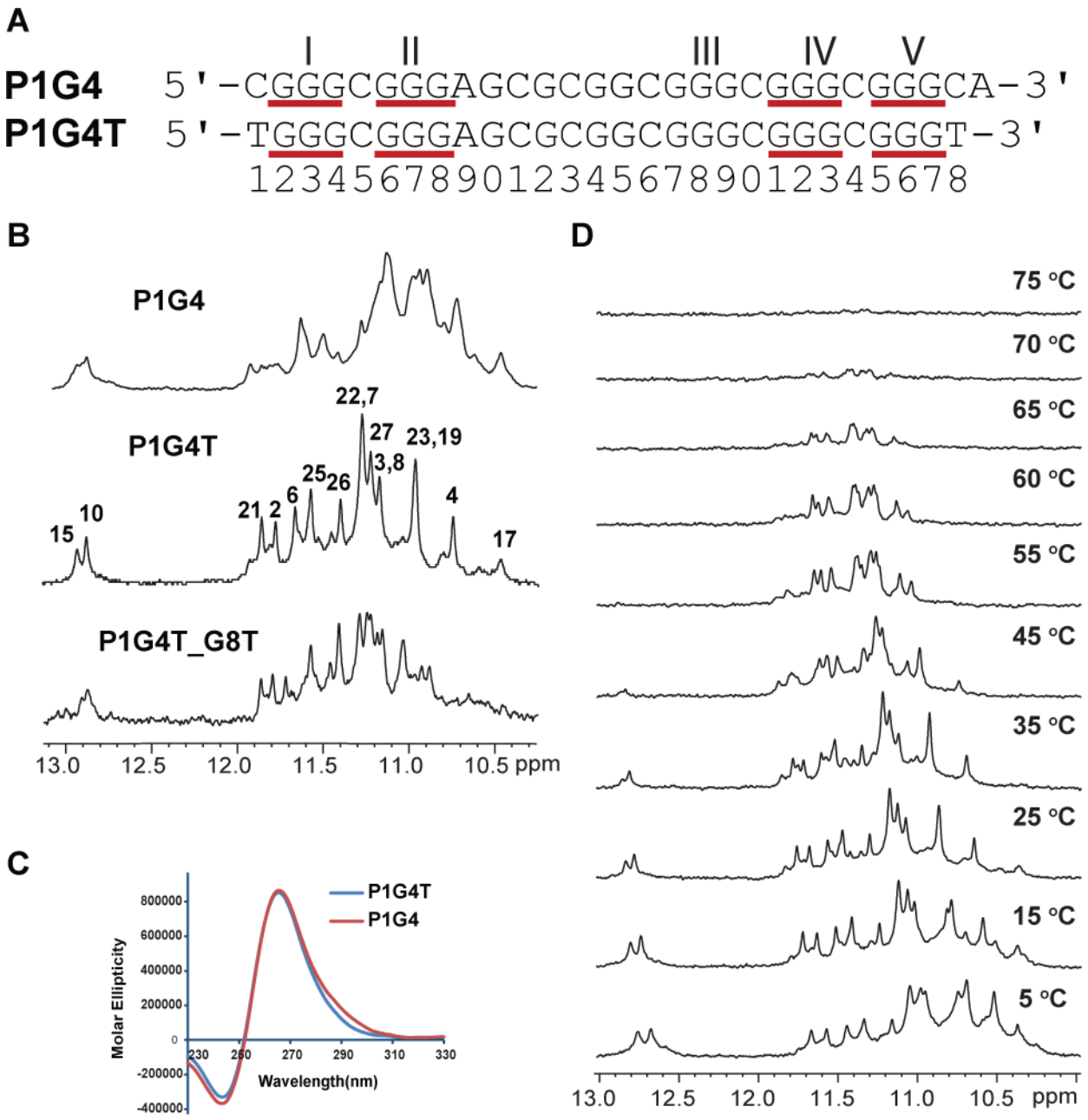


Figure 4.

(A) Sequences of P1G4 and P1G4T. The four G-runs are underlined and numbered. The numbering used in this study is shown for P1G4T. (B) Imino regions of 1D ^{1}H NMR spectra of BCL2 P1G4 sequences at 25 °C in 25 mM K^{+}, pH 7.0. (C) CD spectra of BCL2 P1G4 and P1G4T sequences in 50 mM K^{+}, pH 7.0. (D) NMR variable temperature study of P1G4T in 50 mM K^{+}, pH 7.0.

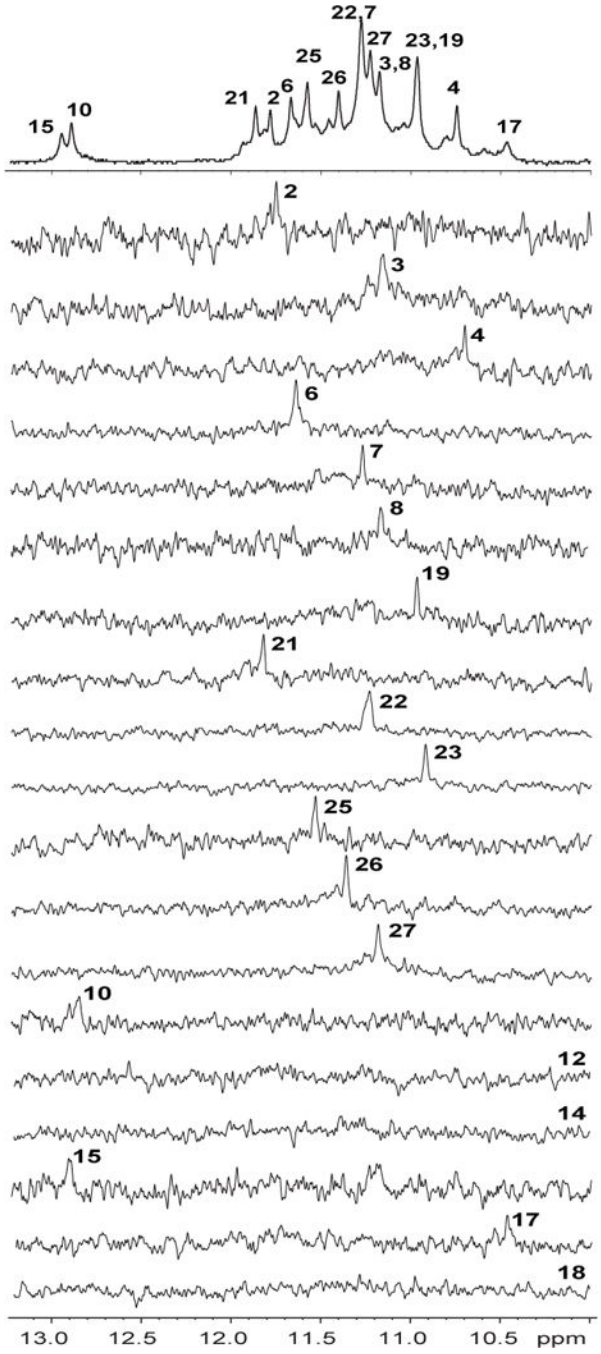


Figure 5.

Imino H1 proton assignments of P1G4T by 1D ^{15}N-filtered experiments using site-specifically labeled oligonucleotides at 25 °C, in 25 mM K^+, pH 7.0.

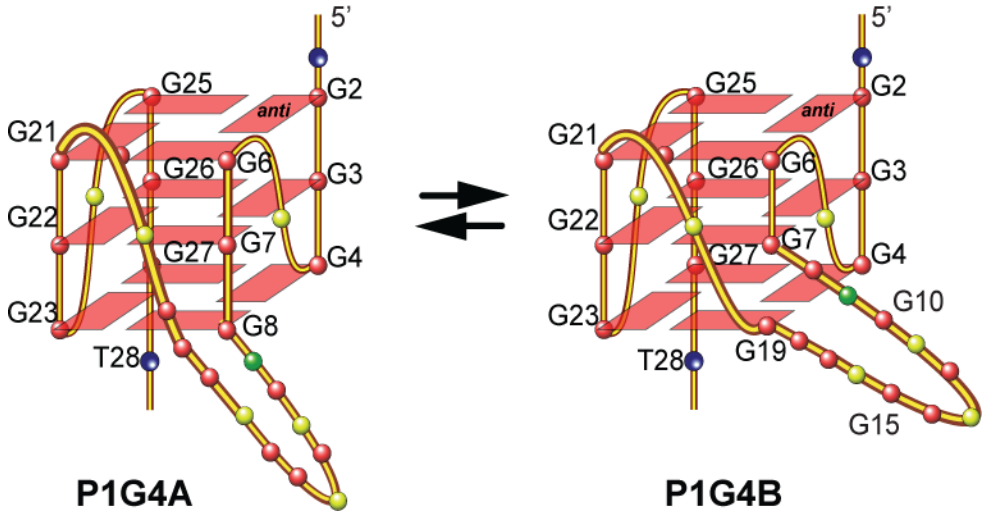


Figure 6.

Schematic drawing of two parallel-stranded P1G4 G-quadruplexes in dynamic equilibrium (G=red, A=green, C=yellow, T=blue).

Table 1

The melting temperatures (Tm) of P1G4T at different DNA concentrations in 50 mM K^+ solution at pH 7 as determined by CD.

[P1G4T]	Tm (°C)
1 μM	68.2
2 μM	68.8
5 μM	68.5
10 μM	68.2
20 μM	68.5