

G-Quadruplex

Beyond Structure: Methylation Fine-Tunes Stability and Folding Kinetics of bcl2Mid G-Quadruplex

Nataša Medved, Mirko Cevce, Uroš Javornik, Jurij Lah, San Hadži, and Janez Plavec*

Abstract: Cytosine methylation, a key epigenetic modification in the regulation of gene expression, raises intriguing questions about its role in the formation and thermodynamic stability of G-quadruplex (G4) structures. We investigated the impact of the 5-methylcytosine residue (C^m) on the well-characterized bcl2Mid G4 structure that forms in a GC-rich region of the B-cell lymphoma 2 (BCL2) gene promoter, which influences its expression. Using solution-state NMR and biophysical techniques, we discovered an unexpected sequence-specific effect of C^m on the folding kinetics of bcl2Mid G4. Specifically, substituting cytosine at position C6 with C^m slows down G4 folding kinetics and influences the equilibrium between major and minor structures in the presence of K^+ ions. Notably, the increased population of the minor structure enabled the characterization of its previously unidentified topology. Additionally, the presence of a single C^m residue induces local structural rearrangements in the major G4 structure and decreases its thermodynamic stability. Furthermore, we found that the zinc finger 3 motif of the Sp1 transcription factor preferentially binds to the minor G4 structure. These results suggest that C^m not only influences G4 polymorphism but may also regulate interactions with transcription factors, potentially affecting the regulation of gene expression.

Introduction

DNA cytosine methylation is a crucial epigenetic modification that regulates gene expression without changing the DNA sequence, allowing cells to respond dynamically to environmental signals.^[1–3] The addition of a methyl group to the C5 atom of cytosine in CpG dinucleotides is catalyzed by enzymes DNA methyltransferases (DNMTs), including DNMT1, DNMT3A, and DNMT3B, which act as methylation “writers”, establishing and maintaining methylation patterns throughout development and differentiation.^[4,5] The methylation patterns are interpreted by “readers”, such as the methyl-CpG-binding domain (MBD) proteins, linking DNA methylation to histone modifications and influencing chromatin structure and gene activity.^[6] “Erasers” like

the ten-eleven translocation (TET) enzymes convert 5-methylcytosine (C^m) into 5-hydroxymethylcytosine.^[7] Most CpG dinucleotides in genic and intergenic regions are typically methylated, whereas CpG islands (CGIs) in GC-rich promoter regions of genes are usually unmethylated, indicating active promoters.^[8,9] However, when CGIs are methylated, they play crucial roles in gene silencing and are involved in processes such as X-chromosome inactivation, genomic imprinting, and cell differentiation.^[10–12] Aberrant methylation patterns in CGIs of promoter regions of tumor suppressor genes and proto-oncogenes can disrupt normal gene expression, leading to abnormal cell proliferation and tumorigenesis.^[13–16]


Recent studies suggest that G-rich sequences in the genome are hotspots for methylation instability and actively influence DNA methylation.^[17,18] It is well established that G-rich regions can fold into noncanonical DNA structures known as G-quadruplexes (G4s). These four-stranded structures consist of at least two G-quartets, where four guanine residues are connected by Hoogsteen-type hydrogen bonds and assembled into a planar motif.^[19–21] Factors such as loop length and nucleotide base composition significantly influence the G4 structure and its stability. Shorter central loops often promote parallel G4 conformations, while longer ones tend to favor antiparallel or hybrid structures.^[22] Additionally, the presence of multiple short loops can predispose a G4 toward a parallel topology.^[23] Notably, water molecules stabilize G4 structures by forming hydrogen bonds with both the G-quartets and the loops, supporting the stability of the G-quartet core and influencing loop conformation, though hydration patterns vary with G4 topology.^[24]


Although factors affecting G4 formation and stability – such as sequence composition, loop length, and the presence

[*] N. Medved, M. Cevce, U. Javornik, J. Plavec
Slovenian NMR Centre, National Institute of Chemistry, Hajdrihova
19, Ljubljana, Slovenia
E-mail: janez.plavec@ki.si

J. Lah, S. Hadži, J. Plavec
Faculty of Chemistry and Chemical Technology, Večna pot 113,
Ljubljana, Slovenia

J. Plavec
EN- FIST Centre of Excellence, Trg OF 13, Ljubljana, Slovenia

 Additional supporting information can be found online in the Supporting Information section

 © 2025 The Author(s). Angewandte Chemie International Edition published by Wiley-VCH GmbH. This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

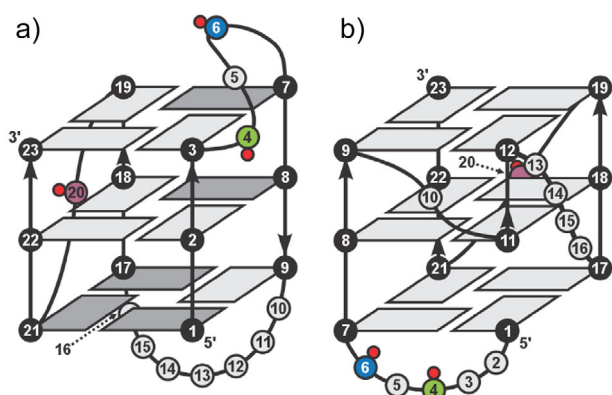


Figure 1. Schematic representation of the G4 structures adopted by the G-rich sequence originating from the BCL2 promoter region. The positions of the three cytosine residues individually substituted with 5-methylcytosine (C^m) are indicated by colored spheres (green, blue, and violet) with an additional red sphere representing the C5-Me group. a) The major G4 structure with (3 + 1) hybrid topology adopted by unmethylated (unmC) and methylated (mC4, mC6, and mC20) oligonucleotides. b) The minor G4 structure with parallel strand orientations and a vacancy in its outer G-quartet that is filled by a snapback element. *Syn* and *anti* glycosidic conformations are colored dark and light grey, respectively.

of various cations and cosolutes – have been extensively studied,^[25–29] the effects of cytosine methylation remain poorly understood. Conflicting findings exist regarding its impact on G4 structures. Although cytosine methylation did not alter the G4 structure of DAT,^[30] it significantly reduced the binding affinity of the G4 structure in the VEGF promoter to the VEGF165 protein, suggesting induced changes in G4 structure.^[31] Interestingly, cytosine methylation has been shown to increase the stability of G4 structures in BCL2 and VEGF, while decreased stability was observed for the MEST-G4 in *in vitro* assays.^[32–34] In the sequence originating from WNT1 promoter, cytosine methylation has been shown to modulate G4 structure.^[35] This variability indicates a complex interplay between DNA methylation and G4 properties indicating that the effects of methylation may differ depending on the specific genomic context. Additionally, recent studies highlight that G4-forming DNA sequences can significantly influence liquid–liquid phase separation (LLPS), a process closely linked to chromatin organization and gene regulation.^[36–39] Understanding how cytosine methylation modifies the formation of G4s is therefore essential, as it may shed light on the mechanisms by which chromatin structure and gene expression are regulated.

In this work, we investigate the effect of introducing individual C^m residues on the folding of bcl2Mid G4 structure adopted by the GC-rich region upstream of the P1 promoter – an important site for interactions with transcription factors that regulate the expression of *BCL2*. The major G4 structure of bcl2Mid adopts a 3 + 1 hybrid topology, featuring one propeller and two lateral loops (Figure 1a).^[40] The 3-nt lateral loop contains two cytosine residues, while the propeller loop comprises a single C20 residue. We propose that substituting cytosine residues at specific CpG sites (namely, C4, C6, and C20) with C^m will induce structural changes and alter the

thermodynamic properties of the resulting G4 structures, with the extent of changes depending on the exact position of C^m within the loops that define and stabilize the major bcl2Mid G4.

Given the polymorphic nature of the GC-rich sequence upstream of the P1 promoter, which can adopt multiple structural conformations,^[41] introducing C^m residues may also influence the formation of less-studied minor structures. By strategically introducing C^m residues within the bcl2Mid sequence, we may expect the formation of minor G4 structures, thereby enhancing our understanding of their role in the regulation of genes.

Our findings demonstrate that introducing a single C^m residue into the bcl2Mid sequence modulates G4 formation in a sequence-dependent manner. Specifically, the presence of a C^m residue at position C6 increases the population of a previously uncharacterized minor G4 structure and slows the kinetics of (re)folding. The shift in equilibrium preferences from the major to minor G4 structure enabled the determination of the folding topology of the minor bcl2Mid structure (Figure 1b). Additionally, the presence of the C^m residue causes structural rearrangements of the major G4 structure limited to the vicinity of the methylation site, which leads to a decrease in thermodynamic stability. These results suggest that cytosine methylation influences G4 polymorphism and possibly its functional role in cellular processes originating from the selectivity of interaction with the zinc finger motif of the DNA-binding domain of Sp1 transcription factor. This indicates that epigenetic modifications may influence gene regulation by fine-tuning G4 structural equilibrium.

Results

1D ^1H NMR Reveals the Presence of Two Distinct G4 Structures

At the outset, we prepared four oligonucleotides, an unmethylated parent (unmC) originating from the promoter of *BCL2* and its three methylated analogs in which the cytosine residues at positions 4, 6, and 20 were individually substituted with C^m (Table 1). Using 1D ^1H NMR experiments, we examined their ability to fold into a G4 in the presence of 110 mM K^+ ions, a concentration that closely resembles intracellular K^+ ion levels. The 12 well-resolved imino proton resonances in the spectrum of unmC (Figure 2a, top panel) indicate G4 formation as a major component in solution, which is consistent with the results of the Dai group.^[40] Likewise, 12 well-resolved imino proton resonances were detected in the spectra of methylated analogs suggesting the formation of a major G4 structure with a chemical shift fingerprint similar to the parent oligonucleotide (Figure 2). Furthermore, the appearance of weaker but well-resolved resonances indicates the coexistence of minor G4 structures in both unmC and its methylated analogs, illustrated in Figures 2, S1. The relative signal intensities of the two G4 structures for mC4 and mC20 are comparable to those of unmC. On the other hand, the comparison of relative intensities of the two

Table 1: List of oligonucleotide sequences used in this study and the corresponding mole fractions of the major and minor G4 structures.

#	Oligonucleotide ^{a)}	Sequence d(5'-3')	Mole fractions of major:minor G4 ^{d)}	
			15 min ^{b)}	22 h ^{b)}
1	unmC (bcl2Mid)	<u>GGG</u> CGC <u>GGG</u> AGGAATT <u>GGG</u> C <u>GGG</u>	0.66:0.34	0.89:0.11
2	mC4	<u>GGG</u> C ^m GC <u>GGG</u> AGGAATT <u>GGG</u> C <u>GGG</u>	0.63:0.37	0.86:0.14
3	mC6	<u>GGG</u> CGC ^m <u>GGG</u> AGGAATT <u>GGG</u> C <u>GGG</u>	0.56:0.44	0.79:0.21
4	mC20	<u>GGG</u> CGC <u>GGG</u> AGGAATT <u>GGG</u> C ^m <u>GGG</u>	0.68:0.32	0.86:0.14
5	unmC(T2T3) ^{c)}	<u>GTT</u> CGC <u>GGG</u> AGGAATT <u>GGG</u> C <u>GGG</u>	–	–
6	mC4(T2T3) ^{c)}	<u>GTT</u> C ^m GC <u>GGG</u> AGGAATT <u>GGG</u> C <u>GGG</u>	–	–
7	mC6(T2T3) ^{c)}	<u>GTT</u> CGC ^m <u>GGG</u> AGGAATT <u>GGG</u> C <u>GGG</u>	–	–
8	mC20(T2T3) ^{c)}	<u>GTT</u> CGC <u>GGG</u> AGGAATT <u>GGG</u> C ^m <u>GGG</u>	–	–
9	unmC(T11T12) ^{c)}	<u>GGG</u> CGC <u>GGG</u> ATTAATT <u>GGG</u> C <u>GGG</u>	–	–
10	mC4(T11T12) ^{c)}	<u>GGG</u> C ^m GC <u>GGG</u> ATTAATT <u>GGG</u> C <u>GGG</u>	–	–
11	mC6(T11T12) ^{c)}	<u>GGG</u> CGC ^m <u>GGG</u> ATTAATT <u>GGG</u> C <u>GGG</u>	–	–
12	mC20(T11T12) ^{c)}	<u>GGG</u> CGC <u>GGG</u> ATTAATT <u>GGG</u> C ^m <u>GGG</u>	–	–

^{a)} Underlined residues denote those that contribute to the formation of G-quartets. C^m indicates the 5-methylcytosine modification. ^{b)} Time after K⁺ ion concentration was increased to 110 mM. ^{c)} The G-to-T substituted sequences form a single structure. ^{d)} Error estimates of the individual fractions amount to ± 0.03 .

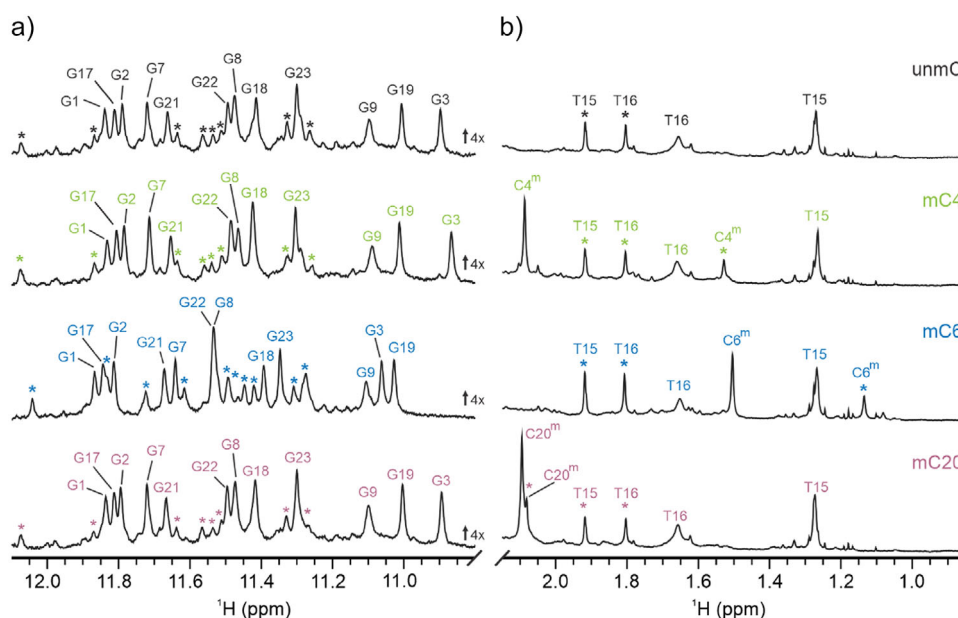


Figure 2. Comparison of the 1D ^1H NMR spectra of unimC and its methylated analogs. a) The imino and b) methyl spectral regions of the oligonucleotides with their names indicated on the right side of each spectrum. The assignments of the imino and methyl proton resonances of the major G4 structures are colored and indicated above each spectrum using the following colors: black for unimC, green for mC4, blue for mC6, and violet for mC20. The resonances marked with an asterisk correspond to the minor G4 structure. The spectra were recorded in 90% H_2O , 10% D_2O , 90 mM KCl, and 20 mM KPi (pH 7.0) at 25 °C on a 600 MHz NMR spectrometer 15 min after the K⁺ ion concentration was increased to 110 mM. The concentration of oligonucleotides was 0.65 mM per strand.

sets of signals in ^1H NMR spectra for mC6 shows a higher population of the minor G4 structure. The methyl proton region (Figure 2b) provides valuable insights, particularly through the clear differentiation of well-resolved methyl group signals of the major and minor G4s. Specifically, the methyl groups corresponding to thymine residues T15 and T16 of the major structure of unimC resonate at δ 1.27 and 1.66 ppm, respectively. For the minor structure of unimC, the two signals at δ 1.80 and 1.92 ppm were assigned to T16 and T15, respectively, based on the characteristic ^{13}C chemical

shifts indicative of methyl groups using 2D ^1H - ^{13}C HSQC in combination with NOESY spectra. Similar observations were made for the methylated analogs. Notably, the ^1H NMR chemical shifts of T15 and T16 for both major and minor structures for unimC and methylated analogs are very similar ($\Delta\delta < 0.01$ ppm) suggesting equivalent topologies.

Moreover, methylated oligonucleotides exhibit two additional signals in the methyl region of the ^1H NMR spectra (Figure 2b). Specifically, the signals belonging to the C5-Me group in the major and minor structure of mC4 resonate at δ

2.09 and 1.53 ppm, respectively. Interestingly, both signals for the C5-Me group in the major and minor structure of mC6 are upfield shifted to δ 1.51 and 1.14 ppm, respectively. For mC20, the chemical shifts of the C5-Me groups of the major and minor structure are δ 2.09 and 2.08 ppm, respectively. The observed variations in chemical shift for the C5-Me groups as well as for other protons and heteronuclei indicate their distinct positions within the folded environment of both the major and minor G4 structures.

Effect on Folding Kinetics and Populations

The visual inspection of changes in ^1H NMR spectra shows that populations of the major and minor structures vary between the parent and methylated analogs. After the addition of K^+ ions, the intensity of the respective signals of the major structure increases with time, while the intensity of the signals of the minor structure decreases. A stationary state was reached after approximately one day (Figure S2).

The mole fractions of the major and minor structure determined by integration of methyl resonances of T15 and T16 for unmC after 22 h are 0.89 and 0.11 (± 0.03), respectively and remain unchanged over several days. Very similar fractions (within the error limits) are observed for mC4 and mC20. Interestingly, the decrease in the fraction of the minor structure of mC6 flattens out at 0.21 ± 0.03 after 22 h, while the major structure reaches a plateau at 0.79 ± 0.03 . Table S1 shows that the rate constant for the B \rightarrow A transition (k_1) is about the same for unmC, mC4, and mC20 with an average $k_1 = 0.15 \pm 0.02 \text{ h}^{-1}$, whereas for mC6 it is 25% lower suggesting that the C5-Me group of residue C6 represents a kinetic barrier that slows down the G4 folding in mC6. The rate constant for the backward A \rightarrow B step (k_{-1}) is about the same for all the methylated analogs with an average $k_{-1} = 0.026 \pm 0.005 \text{ h}^{-1}$ and is 45% higher than k_1 for unmC, indicating that the methylation increases the rate of G4 unfolding. Inspection of Table S1 further indicates that the difference in thermodynamic stability at 25 °C between the major structure (A) and the minor structure (B) is small, but significant ($\Delta G_{AB} = 4.4 \pm 0.7 \text{ kJ mol}^{-1}$) with mC6 exhibiting the lowest thermodynamic stability of the structure A relative to the structure B.

The Minor Structure is a Parallel G4 Featuring a Snapback Element at the 5'-end

The minor structure of mC6, which has the highest population among the methylated analogs, was selected for a more detailed structural characterization. To identify guanine residues involved in G-quartet formation in both the major and minor structures and to determine the folding topology of the minor G4 structure, we performed 1D ^{15}N -edited HSQC experiments acquired on 20% ^{15}N - and ^{13}C -residue specific isotopically labeled guanine residues (Figure S3). Analysis of the spectra identified 12 imino proton signals characteristic of the major structure and 12 imino proton signals for the minor structure (δ 11.2–12.0 ppm). Their number and respective

chemical shifts suggest that the minor G4 contains three G-quartets. Interestingly, G2 and G3 in the first G-tract are not involved in the formation of G-quartets in the minor structure, as evidenced by the absence of their characteristic imino proton resonances. In contrast, they are involved in G-quartet formation in the major structure of mC6. G5 does not participate in the formation of G-quartets in either structure. Notably, G11 and G12 are part of G-quartets in the minor structure, whereas in the major structure, they are part of a 7-nt lateral loop and are not involved in G-quartet formation. For all other guanine residues, imino signals were detected in both the minor and major structures, indicating their involvement in hydrogen bonding within G-quartets.

Given the mole fractions determined directly after the addition of K^+ ions (vide supra) for the major and minor structures in mC6 as 0.56 and 0.44, respectively, the 2D NOESY spectra showed considerable overlap within the imino and aromatic regions, making spectral assignment ambiguous. To prevent the formation of the major structure and increase the population of the minor structure of mC6, we substituted G-to-T at positions G2 and G3. The imino region of mC6(T2T3) spectrum displayed one set of 12 well-resolved signals indicating the formation of a single G4 structure (Figure S4). In contrast, G-to-T substitutions at positions G11 and G12 in mC6 were expected to promote solely the formation of the major G4 structure in mC6(T11T12). Accordingly, the spectrum of mC6(T11T12) showed one set of signals corresponding to the major structure in mC6, considering the chemical shift changes due to G-to-T substitutions. Additionally, we assessed the impact of the C^m residue on the structural features of G-to-T modified oligonucleotides at positions 2 and 3 by comparing their NMR and CD spectra. Very similar 1D ^1H NMR spectra ($\Delta\delta < 0.02 \text{ ppm}$) of unmC(T2T3), mC4(T2T3), mC6(T2T3), and mC20(T2T3) suggest minimal impact of the C^m residue on their structure (Figure S5). CD spectra further confirmed the preserved G4 topology of all four G-to-T analogs by exhibiting a positive peak at 265 nm and a negative peak at 245 nm (Figure S6).

The reduced spectral overlap in the H1-H1 and H1-H8 regions of the 2D NOESY spectrum (Figure 3b) enabled the assignment of guanine residues within G-quartets (Figure 3a), which facilitated the determination of the “minor” (in fact the only structure) G4 topology adopted by mC6(T2T3). Analysis of NOE cross-peaks between imino and H8 protons established three quartets: G1 \rightarrow G17 \rightarrow G21 \rightarrow G7, G11 \rightarrow G18 \rightarrow G22 \rightarrow G8 and G12 \rightarrow G19 \rightarrow G23 \rightarrow G9, with arrows indicating hydrogen-bond directionality. The weak intra-residual H1'-H8 NOE intensities observed in the NOESY spectrum (Figure S7) and the downfield chemical shifts of $\delta_{\text{C}} \approx 137 \text{ ppm}$ for C8 carbon atoms indicate that all guanine residues forming the G-quartets adopt the *anti* conformation around the glycosidic bonds. Notably, within all three G-quartets in the G4 formed by mC6(T2T3) the hydrogen-bond directionality from donor to acceptor was confirmed to be anticlockwise (Figure 3c), indicating a parallel topology. Accordingly, the CD spectrum of mC6(T2T3) shows a positive peak at 265 nm and a negative peak at 245 nm (Figure 3d). In contrast, mC6(T11T12) exhibited a CD spectrum characteristic of a (3 + 1) hybrid topology featuring

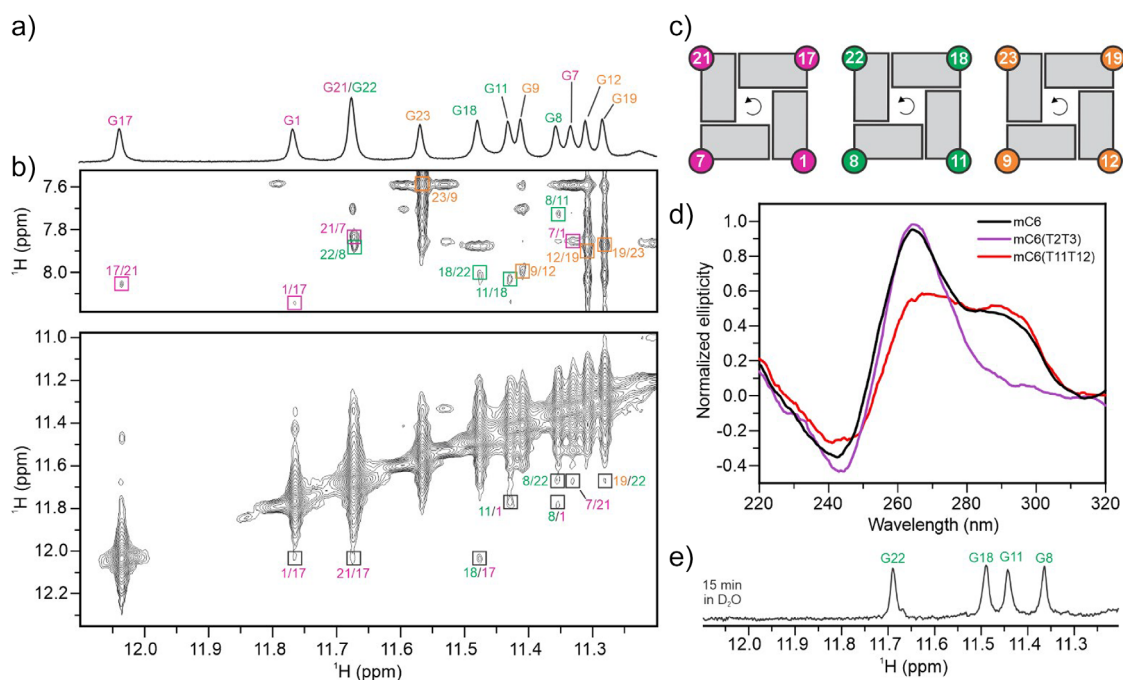


Figure 3. Determination of the G4 topology adopted by mC6(T2T3). a) Imino region of the ^1H NMR spectrum with the assignment. b) Imino-aromatic and imino-imino regions of the 2D NOESY spectrum (mixing time 300 ms). Imino-aromatic cross-peaks are colored and correspond to intraquartet interactions within G1→G17→G21→G7, G11→G18→G22→G8, and G12→G19→G23→G9 quartets. NOESY spectrum was recorded in 90% H_2O , 10% D_2O at 25 °C, 90 mM KCl, and 20 mM KPi buffer with pH 7.0 on a 600 MHz spectrometer. The concentration of oligonucleotide was 1 mM per strand. c) G-quartets with the corresponding donor-to-acceptor hydrogen bond directionalities. All residues within the G-quartet core adopt *anti* glycosidic conformation. d) CD spectra of mC6, mC6(T2T3), and mC6(T1T1T2). Spectra were recorded in 20 mM KPi buffer (pH 7.0) containing 90 mM KCl at 25 °C. The concentration of oligonucleotides was 30 μM . e) Imino proton spectral region of NMR hydrogen–deuterium exchange experiment. The sample was lyophilized in the presence of 110 mM K^+ ions and redissolved in 100% D_2O before the NMR measurement was taken 15 min later. The assignment of the nonexchanging imino protons is listed above the spectrum.

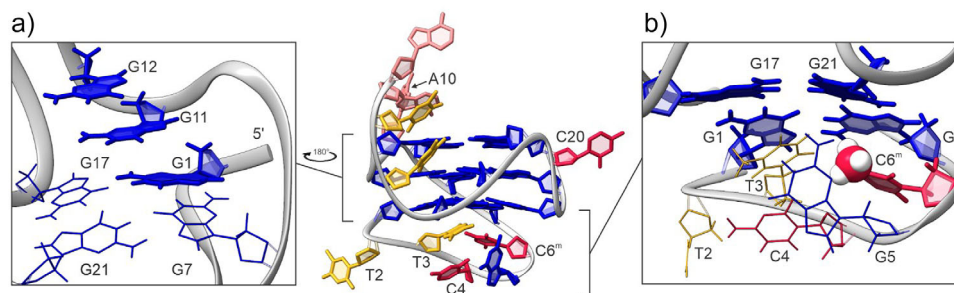


Figure 4. Structure of parallel G4 formed by mC6(T2T3) in the presence of K^+ ions. a) Side view of the snapback element at the 5'-end, where the G1 residue fills the vacancy in the G-quartet. b) Position of C6^m in the lateral loop below G1-G17-G21-G7 quartet. Guanine residues are shown in blue, adenine in salmon pink, cytosine in red, and thymine in yellow.

a negative peak at 245 nm and two positive peaks at 265 and 295 nm. The CD spectrum of mC6 exhibits a negative peak at 245 nm and two positive peaks at 265 and 295 nm corresponding to a structural equilibrium between the major and minor structures.

A parallel G4 structure adopted by mC6(T2T3) with three G-quartets exhibits an intriguing feature: a vacant site at the 5'-end G-quartet is filled by G1 residue, which acts as a snapback element (Figure 4a). This position of G1 is supported by G1H1–G11H1 and G11H1–G12H1 inter-residual NOEs along the G1–G11–G12 tract. Additionally,

the G4 structure adopted by mC6(T2T3) exhibits three propeller loops (A10, A13–A14–T15–T16, and C20) and a lateral snapback loop T2–T3–C4–G5–C6^m (cf. Figure 1b and Figure 4). The snapback loop is positioned under the G1–G17–G21–G7 quartet with C6^m stacked with G7 (Figure 4b). This structural arrangement is consistent with the observed upfield chemical shift of the protons of C5–Me group as well as its NOE contact with the imino proton of G1 in mC6(T2T3). Residues A10 and C20, which form the two 1-nt propeller loops, are oriented outwards and are fully exposed to the solvent (Figure 4). This orientation is supported by an

assessment of water localization using a 2D jump-and-return NOESY experiment at 0 °C (Figure S8). This semiquantitative analysis revealed the effect of the hydrophobic C5-Me group of the cytosine residues in the minor structure. Specifically, the C5-Me groups of C4^m and C6^m in the lateral snapback loop of mC4(T2T3) and mC6(T2T3) show no cross-peaks with water protons, suggesting that these residues at the 5'-end are more protected from solvent exchange by surrounding loop residues, limiting solvent accessibility. In contrast, in mC20(T2T3), the C5-Me group of C20^m shows cross-peaks with water protons, confirming its high solvent exposure.

Furthermore, guanine residues G11, G8, G22, and G18 of mC6(T2T3) show imino proton resonances after 15 min incubation in D₂O, indicating protection from exchange with bulk solvent and their involvement in the central G-quartet (Figures 3e, S9).

Local Structural Changes Induced by Cytosine Methylation in the Major G4 Structure

Detailed analysis of the 1D ¹H and 2D NOESY spectra revealed a high degree of structural similarity between the major G4 structures formed by unmethylated and all three methylated analogs (Figures S10, S11). The major structure of unmC exhibits an interesting conformation of the C4 residue, which is part of the C4-G5-C6 lateral loop. This residue extends across the wide groove and is perpendicular to the G-quartets. Its position close to the G3 residue within the G3-G23-G19-G7 quartet (Figure 5a) is clearly evidenced by medium-to-strong NOE cross-peaks (C4H5-G3H1', C4H6-G3H1', C4H5-G3H4', and C4H6-G3H4'). Additionally, NOE interactions between G5H8 and the sugar protons of C4 support the perpendicular orientation of the C4 residue relative to the G-quartets in unmC.

When the C4 residue is substituted with C4^m, very small changes in chemical shifts with $\Delta\delta$ values of less than 0.01 ppm are observed for imino protons, except for G3H1, which as a neighboring residue shows $\Delta\delta$ of 0.03 ppm compared to unmC. Among the aromatic protons of mC4, the largest $\Delta\delta$ value is observed for its C4^mH6 (0.21 ppm), while G5H8 exhibits $\Delta\delta$ of 0.03 ppm. Interestingly, the presence of C4^m residue weakens the interactions between its C5-Me group and H6 with the sugar protons of G3, as reflected by the NOEs of weaker intensity for C4^mMe-G3H1', C4^mH6-G3H1', C4^mMe-G3H4', and C4^mH6-G3H4', compared to the corresponding NOEs in unmC. In mC4, the signal intensities of C4^mH2''-G5H5'' and C4^mH1''-G5H8 NOEs are stronger compared to C4H2''-G5H5'' and C4H1''-G5H8 in unmC, suggesting more dynamic positioning of the C4^m residue (Figure 5b). This difference likely contributes to changed thermodynamic stability of mC4 compared to unmC ($\Delta\Delta H = 22 \pm 3$ kJ mol⁻¹), as shown in Figures 5e, S12. The unfavorable enthalpic contribution is largely compensated by an increase in entropy ($T\Delta\Delta S = 18 \pm 3$ kJ mol⁻¹), possibly due to the increased dynamics of the C4^m residue compared to C4 in unmC. Interestingly, among the three methylated analogs, the presence of the C5-Me group significantly

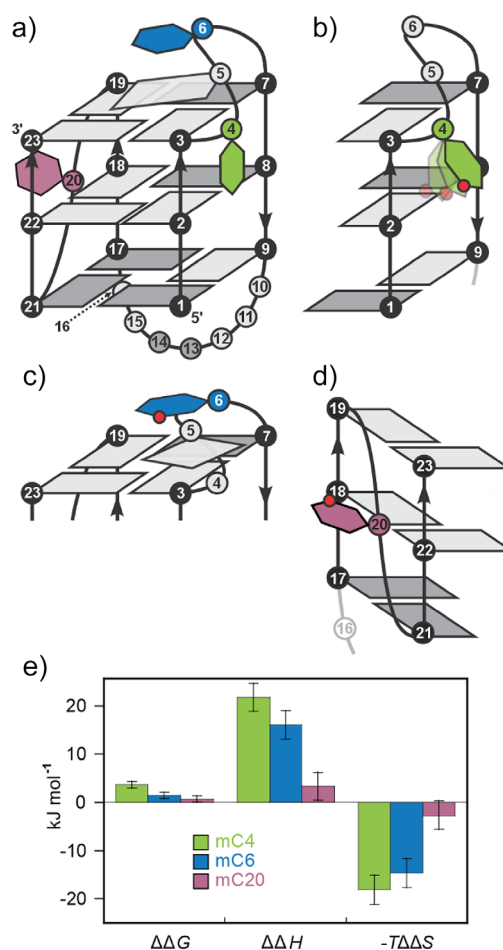


Figure 5. A schematic representation of the local structural changes in the major G4 structure induced by the introduction of Cm in the sequence from the BCL2 promoter. a) The topology adopted by the parent, unmethylated oligonucleotide unmC, with the positions of cytosine residues indicated in green, blue, and violet. b) A view of the wide groove in the G4 structure adopted by mC4 with the positioning and dynamics of the C4^m residue. c) A perspective view on the C4-G5-C6m lateral loop over the 3'-end G-quartet in the G4 adopted by mC6. d) The positioning of C20^m within the 1-nt propeller loop turned away from the medium-sized groove of the G4 structure formed by mC20. The presence of the C5-Me group is indicated by a red sphere in each case. e) Difference in standard thermodynamic parameters of G4 folding ($\Delta\Delta F = \Delta F(\text{mCX}) - \Delta F(\text{unmC})$, F = G, H, and -TS) given for $T = 25$ °C. Positive values correspond to destabilizing contributions and negative to stabilizing ones. X indicates the position of the Cm residue in bcl2Mid oligonucleotide sequence.

destabilizes the major G4 structure only in the case of mC4 ($\Delta\Delta G = 4 \pm 1$ kJ mol⁻¹).

In mC6, C6^m residue exhibits a significant downfield shift of the C6^mH6 proton by $\Delta\delta$ of 0.23 ppm compared to unmC. The $\Delta\delta$ values for imino protons of mC6 are smaller, ranging from 0.01 to 0.15 ppm in both up- and downfield directions. The resonances of G3H1, G7H1, and G23H1, located within the G-quartet at the 3'-end of the major G4 structure adopted by mC6, exhibit $\Delta\delta$ values of 0.15, 0.08, and 0.05 ppm, respectively, in comparison to unmC. In the aromatic region of the ¹H spectrum of mC6 $\Delta\delta$ of 0.07 ppm

was observed for G5H8 located in the lateral loop adjacent to the C6^m residue, along with $\Delta\delta$ of 0.02 ppm for G7H8, which participates in the G3-G7-G19-G23 quartet. In the G4 structure of unmC, G5 and C6 are stacked above the 3'-end G-quartet, with G5 stacked over G3, while C6 is positioned between G7 and G19. The substitution of H5 in C6 residue with a C5-Me group in mC6 results in the appearance of NOE cross-peak between the methyl group of C6^m and the imino proton of G3. Moreover, NOEs between the C5-Me group of C6^m and imino protons of G7 and G19 in mC6 were of stronger intensities compared to those between C6H5 and the corresponding imino protons in unmC. In contrast, the NOE intensity between the C5-Me group of C6^m and the imino proton of G23 suggests positioning of C6^m above G3, G19, and G7 (Figure 5c). This orientation brings sugar H1' and H5' protons of C6^m closer to the pyrimidine ring of G7.

In addition, stronger G5H5''-G7H8 and G5H1''-G23H1 NOEs were detected in mC6, compared to the corresponding atoms in unmC, indicating a repositioning of G5 in the loop that brings its sugar moiety closer to G7 residue. This rearrangement is further substantiated by significantly stronger NOEs between G5H8 and C4H1'/H4'/H5'' compared to unmC, indicating that C6^m influences the entire loop region. The perturbed stacking interactions that probably cause less favorable enthalpic contribution in mC6 compared to unmC (Figure 5e, $\Delta\Delta H = 16 \pm 3$ kJ mol⁻¹) is fully compensated by an increase in entropy ($T\Delta\Delta S = 15 \pm 3$ kJ mol⁻¹).

Negligible chemical shift differences ($\Delta\delta \leq 0.01$ ppm) were observed for mC20 in comparison to unmC (Figure S1). The only exception was the C20^mH6 proton, which showed $\Delta\delta$ of 0.12 ppm, potentially due to the presence of a methyl group at the nearby C5 atom. The minimal influence of C20^m on the structure of mC20 (Figure 5d) was further supported by comparable NOE cross-peak intensities between the sugar protons of G19 and C20^m resembling those observed in unmC between the corresponding residues. This observation is consistent with a poorly defined and dynamic position of the C20 residue in unmC facing away from the medium-sized groove with only two NOEs (G19H1'-C20H5' and G19H4'-C20H3') within the 1-nt propeller loop. Interestingly, in contrast to the lack of significant changes in the ¹H spectrum of mC20, the ³¹P spectrum showed a notable change that was not observed in other methylated analogs. In particular, the C20^m phosphate group connecting C20^m and G21 showed a significantly upfield shifted resonance at δ_P -1.88 ppm compared to the corresponding signal in unmC at δ_P -1.52 ppm (Figure S13), indicating a reorientation of the sugar-phosphate backbone. Thermodynamic stability parameters of mC20 and unmC (Figure 5e) agree within experimental error thus fully supporting the structural observations.

Zinc Finger 3 Motif of Sp1 Transcription Factor Binds Preferentially to Parallel Minor G4

Considering that transcription factors can selectively recognize different G4 topologies,^[42,43] we investigated whether the methylation-induced increase of the minor G4 population influences the binding of Sp1 to G4 structures formed by

bcl2Mid. To assess the interaction of the zinc finger 3 motif (Fin3) with different G4 topologies, we performed ¹H NMR-monitored titration experiments (Figures S14–S17).

Initially, we examined the interaction of Fin3 with the parallel topology of mC6(T2T3). Upon addition of 1 molar equivalent of Fin3, the imino proton signals of G9, G19, and G17 from the outer G-quartets exhibited the most pronounced chemical shift perturbations, with $\Delta\delta$ values of 0.13, 0.07, and 0.07 ppm, respectively (Figure 6a). In contrast, the addition of 1 molar equivalent of Fin3 to the hybrid-type topology of mC6(T1T12) resulted in no significant changes in the chemical shifts with $\Delta\delta < 0.01$ ppm of the guanine imino protons involved in G-quartet formation (Figure 6b). Interestingly, the analysis of mC6, which consists of an equilibrium between 56% major and 44% minor G4 structures, showed that the addition of 1 molar equivalent of Fin3 induced the largest perturbations in the chemical shifts of the imino protons of G9, G19, and G23 with $\Delta\delta$ values of 0.03, 0.05, and 0.03 ppm, respectively, corresponding to the 3'-end G-quartet of the minor G4 structure. In contrast, the chemical shifts of the major G4 structure remained largely unaffected with $\Delta\delta < 0.01$ ppm (Figure 6c). Surprisingly, when 1 equiv of Fin3 was added to unmethylated unmC, which also contains both minor and major G4 structures but has 10 units % less minor G4 structure compared to mC6, we observed a marked decrease in the intensity of the signals corresponding to the imino protons of the minor structure (Figure 6d), suggesting an interaction albeit with specific dynamics on the NMR chemical shift time-scale. However, the chemical shifts of the major G4 structure remained largely unaffected, with $\Delta\delta < 0.01$ ppm, which points to the absence of an interaction.

Discussion

DNA methylation sequencing analysis and clinical data have revealed increased hypermethylation of CpG dinucleotides in gene promoter regions,^[44–46] often associated with down-regulation of gene expression and cancer progression.^[47,48] Given the G-rich nature of gene promoter regions and their propensity to form G4 structures,^[49–51] understanding the relationship between cytosine methylation and G4 formation is crucial. The formation of G4 can protect certain CpG islands from methylation by sequestering DNMT1; however, their interplay is complex, influenced by genomic context, cellular environment, and regulatory elements.^[17,52] This relationship highlights G4s as potential therapeutic targets in cancer treatment. Additionally, recent studies indicate that aberrant DNA methylation in intragenic regions is associated with metabolic diseases and hepatocellular carcinoma, suggesting these regions may serve as alternative regulatory elements influenced by G4 formation.^[53]

Our findings presented here provide new insights into the role of the C^m residue in modulating the polymorphism and folding kinetics of G4 structures, using the bcl2Mid oligonucleotide as a model system. The bcl2Mid sequence was selected for its well-defined structural features and its critical role in the regulation of BCL2 transcription. The GC-rich promoter region acts as a key regulatory element and

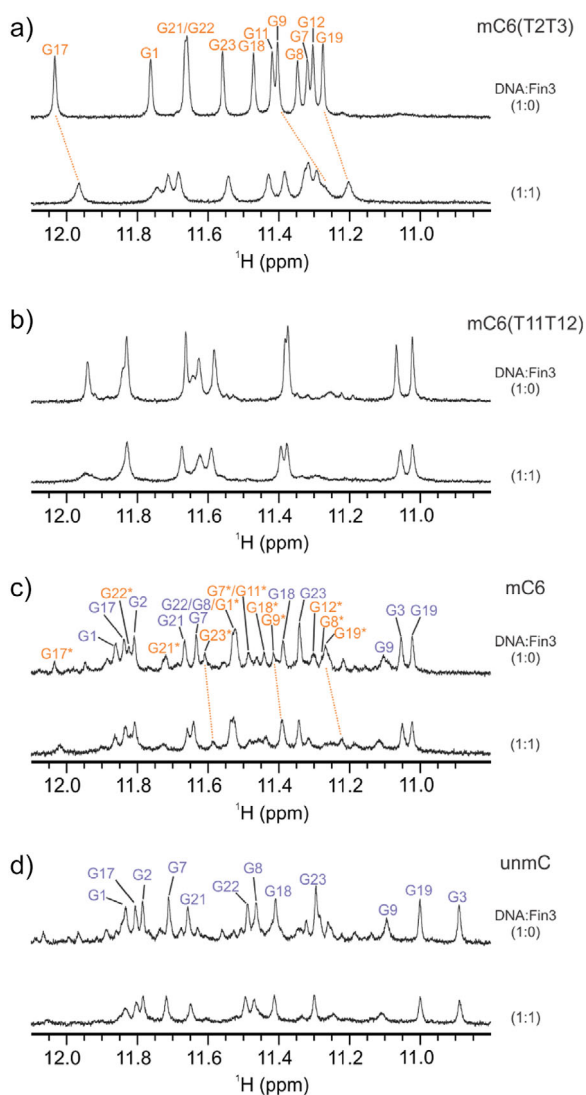


Figure 6. Comparison of the imino proton regions in 1D ^1H NMR spectra of G4 structures in the absence and presence of an equimolar amount of Fin3 peptide, shown on the right side of each spectrum. The signals of the minor G4 structure are denoted by orange assignments in panels A and C, while the signals of the major G4 structure are marked in purple in panels C and D. The dashed lines in panels A and C highlight the imino protons with the largest changes in chemical shifts. a) Assigned imino proton region of mC6(T2T3) showing the spectrum of the DNA alone (top) and after addition of Fin3 peptide in a molar ratio of 1:1 (bottom). b) Imino proton region of mC6(T11T12) in the absence and presence of Fin3 peptide. c) Spectrum of mC6 illustrating the equilibrium between 56% major and 44% minor G4 structures. The assigned guanine imino protons for both structures are shown in the absence of Fin3, with the minor structure indicated by asterisks and the most significant chemical shift changes upon peptide addition marked by orange dashed lines. d) Spectrum of the unmethylated parent unmC showing the equilibrium between 66% major and 34% minor G4 structures, with assigned guanine imino proton signals of the major G4 structure. All DNA spectra with 100 μM concentration per strand were recorded in a 20 mM potassium phosphate buffer at pH 7.0 with 90 mM potassium chloride (KCl) at 25 $^{\circ}\text{C}$. The Fin3 peptide, dissolved in 90% H_2O and 10% D_2O (v/v), containing 10 mM Tris-HCl buffer (pH 7.0), 1 mM DTT and ZnCl_2 at a 10% molar excess, was added to the DNA samples at a molar ratio of 1:1.

interacts with transcription factors such as Sp1, CREB, and WT1. These interactions, which are mediated by sequence-specific binding, can down- or up-regulate the expression of the *BCL2* gene. Previous studies demonstrated the formation of a major G4 structure by bcl2Mid oligonucleotide in the presence of 60 mM K^+ ions, accompanied by a minor species comprising less than 5% of the total population.^[40] Intriguingly, our results reveal that substituting cytosine with 5-methylcytosine in all three bcl2Mid analogs leads to the coexistence of major and minor G4 structures at a more biologically relevant concentration of 110 mM K^+ ions. Over time, the kinetically favored minor G4 undergoes a refolding process and transitions into the thermodynamically stable major G4, reaching equilibrium after approximately 24 h.

Interestingly, the populations of the major and minor G4 structures are related to the position of the methylated cytosine residue in the oligonucleotide sequence. The most pronounced differences are observed in mC6 which exhibits a higher population of the minor G4 (21%) compared to unmC (11%). Interestingly, the initial population of the minor G4 in mC6 reaches up to 44% after an increase of K^+ ion concentration to 110 mM. This corresponds to an approximately 10 unit % higher population of the minor G4 structure compared to unmC and the other two methylated analogs, with this increased population persisting even when the system reaches equilibrium.

Gene silencing is often associated with the hypermethylation of gene promoter regions.^[54,55] However, our results suggest that the presence of two or three C^{m} residues in the bcl2Mid sequence does not have a significant additive effect on the structural equilibrium of G4s (Figure S18). Sequences containing C6^{m} in combination with C4^{m} and/or C20^{m} residues exhibit a similar ratio of minor to major G4 structures as observed for sequences with a single C^{m} at position C6. This implies that additional C^{m} residues do not significantly affect the G4 structural equilibrium, indicating that methylation at a single site, particularly at C6 in the case of bcl2Mid, is sufficient to affect G4 polymorphism and subsequently interaction with Fin3.

To elucidate the factors contributing to the increased minor G4 population in mC6, we characterized a modified variant mC6(T2T3) which adopts a structure similar to the minor G4 as the only structure. Structural analysis indicates that mC6(T2T3) forms a three-quartet G4 with parallel strand orientations and a vacancy in its outer G-quartet. A vacant site is filled by the G1 residue as part of a snapback element at the 5'-end, forming a G4 with three intact G-quartets. The mC6(T2T3) features two 1-nt propeller loops spanning medium-sized grooves. The A10 loop bridges two G-quartets by connecting the G-quartets at the middle and the 3'-end, while the C20 loop bridges three G-quartets by connecting the G-quartets at the 3'- and 5'-end.

The destabilizing effect of an adenine residue in 1-nt propeller loops is well-documented.^[56] In the mC6(T2T3) structure, the presence of the A10 residue within the 1-nt loop may promote a transition to the major G4 structure, possibly forming a base pair with T15 and further stabilizing the major structure through stacking interactions with the G-quartet at the 5'-end. A similar adenine-driven structural switch from

a two-quartet to a three-quartet G4 has been observed in Ran4,^[57] further supporting the role of adenine in facilitating transitions in G4 structures.

In the structural equilibrium of major and minor structures of unmC, mC4, mC6, and mC20, the predominant conformation is the 3 + 1 hybrid topology of the major G4 structure, likely due to more favorable characteristic glycosidic conformations along the G-tracts.^[58] The major structure features three G-tracts arranged in a 5'-*syn-anti-anti* pattern and one G-tract in a 5'-*syn-syn-anti* pattern. In contrast, the minor G4 structure has a less favorable 5'-*anti-anti-anti* arrangement along all four G-tracts. The favorable structural properties of the major G4 structure prevent a transition back to the minor G4, thus maintaining a stable ratio between the two structures in the stationary state.

The transition from minor to major G4 structure may also be influenced by the dynamic 5-nt lateral snapback loop. The absence of NOE contacts between the loop residues in the NOESY spectra suggests a high degree of loop flexibility. However, in the mC6(T2T3) structure, the specific interaction between the hydrophobic C5-Me group of the C6^m residue and the imino proton of G1 within the G-quartet at the 5'-end suggests the formation of a localized hydrophobic core at the bottom of the medium-sized groove. This interaction likely stabilizes the minor structure by reducing loop flexibility and creating a kinetic barrier that impedes the transition to the major G4 structure. This is further supported by the observed 25% decrease in the rate constant (k_1) for mC6 compared to unmC, mC4, and mC20. Interestingly, the C4^m residue in mC4 does not affect the rate constant k_1 despite the proximity of mC4 and mC6 in the bcl2Mid oligonucleotide sequence suggesting similar effects. The central position of the C4^m residue in the flexible 5-nt loop allows it to adopt a favorable orientation, resulting in minimal impact on transition kinetics.

Furthermore, the sequence-specific influence of the C^m residue is reflected in the structural and thermodynamic properties of the major G4 structure. UV melting curves of unmC and its methylated analogs showed no hysteresis and highly similar melting temperatures (T_m) differing by approximately ± 1 °C (Figure S19). Although the major G4 structure is preserved, subtle local structural rearrangements were observed in the methylated analogs compared to unmC. The altered positioning of the C4^m residue within the wide groove of the major G4, relative to C4 in unmC, causes considerable thermodynamic destabilization in mC4. Conversely, the C6^m residue in mC6 induces small local changes around the C5-Me group in the major G4, resulting in only a minimal decrease in its thermodynamic stability.

Despite minimal structural changes in the major G4, we observed differences in water localization patterns. In unmC, correlations were observed between water protons and imino protons of the G-quartets at 3'- and 5'-end (Figure S8b). However, the introduction of C4^m and C6^m into the 3-nt lateral loop disrupted correlation with G19, suggesting that cytosine methylation may affect water-mediated interactions that are important for structural stability and possibly also for stabilization of the DNA-protein complex in the regulation of gene expression.

Additionally, the remarkable tolerance to the presence of C20^m in mC20 is reflected in its minimal impact on the major G4 and its thermodynamic stability compared to unmC. The C20 residue forms a 1-nt loop that connects G17-G18-G19 and G21-G22-G23 parallel G-tracts, representing a conserved motif in both major and minor G4 structures. This suggests that the formation of a G₃NG₃ hairpin-like motif may represent an early step in the (re)folding pathway from the minor to major G4 structure. The G₃NG₃ motif was also observed in the PDGFR- β gene promoter region, where it is thought to provide a stable structural scaffold for G4 formation.^[59]

In addition to these structural considerations, G4-protein interactions have attracted increasing interest in recent years, particularly in the context of gene regulation.^[60–62] Among the proteins involved, the transcription factor Sp1 has received considerable attention due to its zinc finger motifs that mediate sequence-specific DNA binding with an affinity for GC-rich sequences. Notably, the second and third zinc fingers of Sp1 are essential for high-affinity interactions, with the third zinc finger playing a crucial role in stabilizing the DNA-protein complex through base-specific contacts.^[63,64] Our results indicate that nucleotide-dependent methylation of the bcl2Mid sequence alters the structural equilibrium of G4s and shifts it toward the minor G4 form, which is preferentially bound by the third zinc finger motif of Sp1. Our results suggest that by slowing down G4 folding, cytosine methylation may facilitate Sp1 binding to otherwise inaccessible regions, potentially enhancing transcriptional activation. This finding bridges the apparent contrast between the role of methylation in gene silencing and G4 structures as transcriptional activators and highlights a complex regulatory mechanism that requires further investigation to fully elucidate its broader effects on transcriptional control.

Conclusion

Cytosine methylation of the GC-rich region upstream of the P1 promoter, which regulates the expression of BCL2, influences its folding into G4 structures, the major, 3 + 1 hybrid and the minor, parallel G4. The increased population of the minor G4 structure induced by 5-methylcytosine allowed us to characterize its previously unknown structure. Our findings on residue-specific stabilization and the kinetics of formation of one or the other G4 structure highlight the importance of considering noncanonical structures in biological processes from the physicochemical perspective, beyond changes in the pKa value of the (un)methylated nucleobase, its van der Waals interactions that affect nucleobase stacking, or hydrophobic properties that may influence the early stages of folding of a G-rich sequence by local (de)hydration. Our NMR study of the minor species has revealed a 3D structure consisting of three G-quartets connected with three propeller-type loops. The most exciting and intriguing feature is that a vacancy at the 5' end of the G-quartet is occupied by a G1 residue that acts as a snapback element. Insights into how 5-methylcytosine alters the folding kinetics of G-rich DNA regions and thus affects structural equilibria can have

profound implications for the regulation of gene expression and drug targeting.

Supporting Information

The data that support the findings of this study are available in the Supporting Information of this article. The authors have cited additional references within the Supporting Information.^[65–69]

Acknowledgements

The authors acknowledge the financial support from Slovenian Research and Innovation Agency [ARIS, grants P1-0242, P1-0201 and J1-60019] and CERIC – ERIC consortium for access to experimental facilities and financial support.

Conflict of Interests

The authors declare no conflict of interest.

Data Availability Statement

The NMR spectra supporting this study are available at <https://doi.org/10.5281/zenodo.13285005>.

Keywords: Differential scanning calorimetry • DNA methylation • G-quadruplexes • Kinetics • NMR spectroscopy

- [1] A. Bird, *Nature* **2007**, *447*, 396–398.
- [2] T. Sexton, G. Cavalli, *Cell* **2015**, *160*, 1049–1059.
- [3] Z. D. Smith, A. Meissner, *Nat. Rev. Genet.* **2013**, *14*, 204–220.
- [4] Z. Chen, Y. Zhang, *Annu. Rev. Biochem.* **2020**, *89*, 135–158.
- [5] H. Hashimoto, P. M. Vertino, X. Cheng, *Epigenomics* **2010**, *2*, 657–669.
- [6] Q. Du, P. L. Luu, C. Stirzaker, S. J. Clark, *Epigenomics* **2015**, *7*, 1051–1073.
- [7] S. Ito, A. C. D'Alessio, O. V. Taranova, K. Hong, L. C. Sowers, Y. Zhang, *Nature* **2010**, *466*, 1129–1133.
- [8] R. S. Illingworth, A. P. Bird, *FEBS Lett.* **2009**, *583*, 1713–1720.
- [9] A. M. Deaton, A. Bird, *Genes Dev.* **2011**, *25*, 1010–1022.
- [10] A. Fedoriw, J. Mugford, T. Magnuson, *Cold Spring Harb. Perspect. Biol.* **2012**, *4*, a008136.
- [11] C. G. Duncan, S. A. Grimm, D. L. Morgan, P. R. Bushel, B. D. Bennett, J. D. Roberts, F. L. Tyson, B. A. Merrick, P. A. Waude, *Sci. Rep.* **2018**, *8*, 10138.
- [12] S. B. Peeters, B. J. Posynick, C. J. Brown, *Epigenomes* **2023**, *7*, 29.
- [13] B. Vogelstein, K. W. Kinzler, *Nat. Med.* **2004**, *10*, 789–799.
- [14] J. Shu, J. Jelinek, H. Chen, Y. Zhang, T. Qin, M. Li, L. Liu, J.-P. J. Issa, *Nucleic Acids Res.* **2024**, *52*, 4857–4871.
- [15] M. Esteller, J. M. Silva, G. Dominguez, F. Bonilla, X. Matias-Guiu, E. Lerma, E. Bussaglia, J. Prat, I. C. Harkes, E. A. Repasky, E. Gabrielson, M. Schutte, S. B. Baylin, J. G. Herman, *J. Natl. Cancer Inst.* **2000**, *92*, 564–569.
- [16] J. G. Herman, A. Umar, K. Polyak, J. R. Graff, N. Ahuja, J. P. Issa, S. Markowitz, J. K. Willson, S. R. Hamilton, K. W. Kinzler, M. F. Kane, R. D. Kolodner, B. Vogelstein, T. A. Kunkel, S. B. Baylin, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 6870–6875.
- [17] S.-Q. Mao, A. T. Ghanbarian, J. Spiegel, S. Martínez Cuesta, D. Beraldi, M. Di Antonio, G. Marsico, R. Hänsel-Hertsch, D. Tannahill, S. Balasubramanian, *Nat. Struct. Mol. Biol.* **2018**, *25*, 951–957.
- [18] J. Rauchhaus, J. Robinson, L. Monti, M. Di Antonio, *Genes* **2022**, *13*, 1665.
- [19] J. Spiegel, S. Adhikari, S. Balasubramanian, *Trends Chem* **2020**, *2*, 123–136.
- [20] A. T. Phan, V. Kuryavyi, D. J. Patel, *Curr. Opin. Struct. Biol.* **2006**, *16*, 288–298.
- [21] S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd, S. Neidle, *Nucleic Acids Res.* **2006**, *34*, 5402–5415.
- [22] A. Arora, S. Maiti, *J. Phys. Chem. B* **2009**, *113*, 8784–8792.
- [23] A. Y. Q. Zhang, A. Bugaut, S. Balasubramanian, *Biochemistry* **2011**, *50*, 7251–7258.
- [24] L. Kevin, L. Yatsunyk, S. Neidle, *Nucleic Acids Res.* **2021**, *49*, 519–528.
- [25] N. V. Hud, J. Plavec, in *Quadruplex Nucleic Acids* (Eds.: S. Neidle, S. Balasubramanian), The Royal Society of Chemistry, RSC Publishing, Cambridge, **2006**.
- [26] T. Fujii, P. Podbevšek, J. Plavec, N. Sugimoto, *J. Inorg. Biochem.* **2017**, *166*, 190–198.
- [27] P. Šket, J. Plavec, *J. Am. Chem. Soc.* **2010**, *132*, 12724–12732.
- [28] M. Marušič, J. Plavec, *Angew. Chem. Int. Ed.* **2015**, *54*, 11716–11719.
- [29] M. Trajkovski, T. Endoh, H. Tateishi-Karimata, T. Ohya, S. Tanaka, J. Plavec, N. Sugimoto, *Nucleic Acids Res.* **2018**, *46*, 4301–4315.
- [30] N. Nain, A. Singh, S. Khan, S. Kukreti, *Biochem. Biophys. Res. Rep.* **2023**, *34*, 10146.
- [31] K. Tsukakoshi, S. Saito, W. Yoshida, S. Goto, K. Ikebukuro, *Molecules* **2018**, *23*, 944.
- [32] J. Lin, J.-Q. Hou, H.-D. Xiang, Y.-Y. Yan, Y.-C. Gu, J.-H. Tan, D. Li, L.-Q. Gu, T.-M. Ou, Z.-S. Huang, *Biochem. Biophys. Res. Commun.* **2013**, *433*, 368–373.
- [33] W. Yoshida, H. Yoshioka, D. H. Bay, K. Iida, K. Ikebukuro, K. Nagasawa, I. Karube, *Anal. Chem.* **2016**, *88*, 7101–7107.
- [34] A. J. Stevens, S. Stoffrein-Roberts, S. L. Cree, A. Gibb, A. L. Miller, K. Doudney, A. Aitchison, M. R. Eccles, P. R. Joyce, V. V. Filichev, M. A. Kennedy, *PLoS One* **2014**, *9*, e113955.
- [35] Z.-F. Wang, M.-H. Li, I.-T. Chu, F. R. Winnerdy, A. T. Phan, T.-C. Chang, *Nucleic Acids Res.* **2020**, *48*, 1120–1130.
- [36] M. Mimura, S. Tomita, Y. Shinkai, T. Hosokai, H. Kumeta, T. Saio, K. Shiraki, R. Kurita, *J. Am. Chem. Soc.* **2021**, *143*, 9849–9857.
- [37] N. F. Antarksa, M. Di Antonio, *Acc. Chem. Res.* **2024**, *57*, 3397–3406.
- [38] G. Giraud, M. Rodà, P. Huchon, M. Michelet, S. Maadadi, D. Jutzi, R. Montserret, M.-D. Ruepp, R. Parent, C. Combet, F. Zoulim, B. Testoni, *Nucleic Acids Res.* **2024**, *52*, 2290–2305.
- [39] W. Wang, D. Li, Q. Xu, J. Cheng, Z. Yu, G. Li, S. Qiao, J. Pan, H. Wang, J. Shi, T. Zheng, G. Sui, *Nat. Commun.* **2024**, *15*, 1045.
- [40] J. Dai, D. Chen, R. A. Jones, L. H. Hurley, D. Yang, *Nucleic Acids Res.* **2006**, *34*, 5133–5144.
- [41] S. Bielskuté, J. Plavec, P. Podbevšek, *Nucleic Acids Res.* **2021**, *49*, 2346–2356.
- [42] S. Marzano, G. Pinto, A. Di Porzio, J. Amato, A. Randazzo, A. Amoresano, B. Pagano, *Commun Chem* **2025**, *8*, 64.
- [43] J. Spiegel, S. M. Cuesta, S. Adhikari, R. Hänsel-Hertsch, D. Tannahill, S. Balasubramanian, *Genome Biol.* **2021**, *22*, 117.
- [44] L. Zhao, X. Wu, J. Zheng, D. Dong, *Oncogene* **2021**, *40*, 1884–1895.
- [45] J. W. Chiang, B. Y. Karlan, L. Cass, R. L. Baldwin, *Gynecol. Oncol.* **2006**, *101*, 403–410.
- [46] V. V. Strel'nikov, E. B. Kuznetsova, A. S. Tanas, V. V. Rudenko, A. I. Kalinkin, E. V. Poddubskaya, T. V. Kekeeva, G. G. Chesnokova, I. D. Trotsenko, S. S. Larin, S. I. Kutsev, D. V.

- Zaletaev, M. V. Nemtsova, O. A. Simonova, *Sci. Rep.* **2021**, *11*, 2264.
- [47] R. Lakshminarasimhan, G. Liang, *Adv. Exp. Med. Biol.* **2016**, *945*, 151–172.
- [48] G. L. Gregory, I. M. Copple, *Mol. Ther. Nucleic Acids* **2023**, *31*, 211–223.
- [49] J. Brčić, J. Plavec, *Front. Chem. Sci. Eng.* **2016**, *10*, 222–237.
- [50] A. Kotar, R. Rigo, C. Sissi, J. Plavec, *Nucleic Acids Res.* **2019**, *47*, 2641–2653.
- [51] J. L. Huppert, S. Balasubramanian, *Nucleic Acids Res.* **2007**, *35*, 406–413.
- [52] M. Jara-Espejo, S. R. Peres Line, *FEBS J.* **2020**, *287*, 483–495.
- [53] A. Lahnsteiner, V. Ellmer, A. Oberlercher, Z. Liutkevičiūtė, E. Schönauer, B. Paulweber, E. Aigner, A. Risch, *Sci. Rep.* **2024**, *14*, 20215.
- [54] K. C. Ehrlich, M. Lacey, C. Baribault, S. Sen, P. O. Esteve, S. Pradhan, M. Ehrlich, *Epigenomes* **2022**, *6*, 43.
- [55] J. S. Fain, A. Lorient, A. Diacofotaki, A. Van Tongelen, C. de Smet, *Sci. Rep.* **2021**, *11*, 17346.
- [56] J. Jana, Y. M. Vianney, N. Schröder, K. Weisz, *Nucleic Acids Res.* **2022**, *50*, 7161–7175.
- [57] M. L. Živković, J. Rozman, J. Plavec, *Angew. Chem.* **2018**, *130*, 15621–15625.
- [58] X. Cang, J. Šponer, T. E. Cheatham, *Nucleic Acids Res.* **2011**, *39*, 4499–4512.
- [59] Y. Chen, P. Agrawal, R. V. Brown, E. Hatzakis, L. Hurley, D. Yang, *J. Am. Chem. Soc.* **2012**, *134*, 13220–13223.
- [60] F. Merlino, S. Marzano, P. Zizza, F. D'Aria, N. Grasso, A. Carachino, S. Iachettini, A. Biroccio, S. Di Fonzo, P. Grieco, A. Randazzo, J. Amato, B. Pagano, *Nucleic Acids Res.* **2024**, *52*, 6748–6762.
- [61] A. de Rache, J. Marquieville, S. Bouaziz, B. Vialet, M.-L. Andreola, J.-L. Mergny, S. Amrane, *J. Mol. Biol.* **2024**, *436*, 168359.
- [62] S. Da Ros, G. Nicoletto, R. Rigo, S. Ceschi, E. Zorzan, M. Dacasto, M. Giantin, C. Sissi, *Methods Mol. Biol.* **2021**, *22*, 329.
- [63] J. Song, H. Ugai, K. Ogawa, Y. Wang, A. Sarai, Y. Obata, I. Kanazawa, K. Sun, K. Itakura, K. K. Yokoyama, *J. Biol. Chem.* **2001**, *276*, 30429–30434.
- [64] E.-A. Raiber, R. Kranaster, E. Lam, M. Nikan, S. Balasubramanian, *Nucleic Acids Res.* **2012**, *40*, 1499–1508.
- [65] F. Lankaš, T. E. Cheatham, III, N. Špačáková, P. Hobza, J. Langowski, J. Šponer, *Biophys. J.* **2002**, *82*, 2592–2609.
- [66] M. Bončina, J. Lah, I. Prislán, G. Vesnaver, *J. Am. Chem. Soc.* **2012**, *134*, 9657–9663.
- [67] S. Hadži, M. Bončina, J. Lah, *Methods Mol. Biol.* **2019**, *2035*, 117–130.
- [68] B. Lee, F. M. Richards, *J. Mol. Biol.* **1971**, *55*, 379–IN4.
- [69] S. Hadži, J. Lah, *Biochim. Biophys. Acta Gen. Subj.* **2021**, *1865*, 129774.

Manuscript received: April 03, 2025

Revised manuscript received: April 04, 2025

Accepted manuscript online: April 07, 2025

Version of record online: April 16, 2025