PERIODICUM BIOLOGORUM
VOL. 127, No 3–4, 193–206, 2025
DOI: 10.18054/pb.v127i3-4.39372

UDC 57:61
CODEN PDBIAD
ISSN 0031-5362

Original research article

Potential long-looped G-quadruplexes in the BCL-2 promoter and their implications in gene regulation

LAN TANKO[1,#]
DAVID KRANJC[1,#]
JANEZ PLAVEC[1,2,3,*]

[1] University of Ljubljana, Faculty of Chemistry and Chemical Technology, Ljubljana, Slovenia

[2] National Institute of Chemistry, Ljubljana, Slovenia

[3] EN→FIST Centre of Excellence, Ljubljana, Slovenia

[#] Contributed equally.

[*]**Correspondence:**
Janez Plavec
E-mail address: janez.plavec@ki.si

Keywords: DNA; G-quadruplex; long-looped G4s; BCL-2; cancer

Abbreviations:
A-DNA – right-handed DNA structure
BCL-2 – B-cell lymphoma 2
B-DNA – canonical right-handed DNA structure
CRE – cAMP response element
EPD – Eukaryotic Promoter Database
G4 – G-quadruplex
MFE – minimum free energy
nt – nucleotide
TSS – transcription start site
Z-DNA – non-canonical left-handed DNA structure

Received November 12, 2025
Revised December 14, 2025
Accepted December 15, 2025

Abstract

DNA is a highly dynamic polymer composed of nucleotides that typically forms canonical double helix but can also fold into a variety of non-canonical secondary structures under specific conditions. G-quadruplexes (G4s) are an example of such structures that arise from guanine-rich regions and increasingly emerge as key regulators of gene expression. While G4s usually follow a well-established consensus sequence and typically contain short loops, recent evidence indicates that non-standard G4s with long loops may also form and impact gene regulation.

In this article, we examine the potential roles of long-looped G4s in the BCL-2 promoter. We summarize known G4-forming regions, highlighting their distinct topologies and regulatory outcomes, and present structural features that suggest extended loops may cause specific molecular recognition. Using custom computational analysis, we identified putative long-looped G4 motifs that conventional algorithms, such as G4Hunter, often overlook due to loop length penalties. Predictions with RNAfold further suggest that some long-looped candidates exhibit stable hairpin-like loops, supporting their potential to fold under physiological conditions.

We discuss how long loops may enable unique promoter–enhancer interactions, act as additional interaction sites for transcription factors or ligands and expand opportunities for selective targeting by dual-specific ligands. Altogether, we propose that long-looped G4s represent an underexplored layer of gene regulation with the BCL-2 gene as one example of genes potentially including such G4s.

INTRODUCTION

The discovery of the DNA double helix represents one of the most significant scientific achievements of the 20th century. Over 70 years have passed since James Watson and Francis Crick proposed their model founded on the complementary pairing of four different nucleotide bases, while Rosalind Franklin's X-ray diffraction images provided crucial information for this discovery *(1)*. Although there have been some changes to the model in the subsequent years, its fundamental structural features remain the same.

The primary structural components of DNA are deoxyribonucleotides, composed of a phosphate group, a 2'-deoxyribose sugar and one of four nucleobases – adenine, thymine, guanine or cytosine (Figure 1A) *(2)*. Canonical B-DNA consists of two polynucleotide strands stabilized by van der Waals forces and other hydrophobic interactions between the stacked bases, while hydrogen bonds between the nucleobase moieties

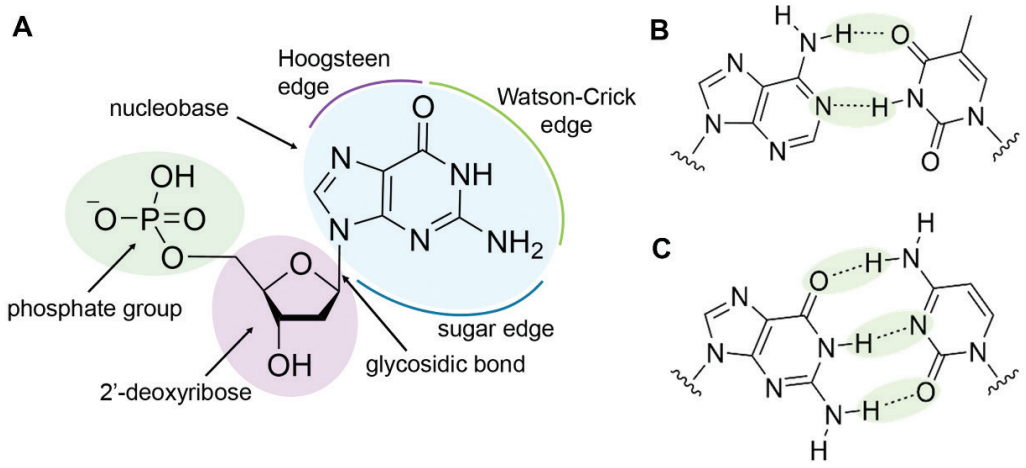

Figure 1. *Nucleotide structure and base pairing. (**A**) Structural formula of 2'-deoxyriboguanosine monophosphate with labeled important structural elements and the Hoogsteen, Watson-Crick and sugar hydrogen-bonding edges. (**B**) Adenine-thymine base pair in the Watson-Crick geometry present in double stranded DNA. (**C**) Guanine-cytosine base pair. Hydrogen bonds in panels B and C are highlighted in green.*

ensure pairing specificity. Adenine pairs with thymine and guanine pairs with cytosine (Figure 1B, C).

The two strands are antiparallel, meaning they run in opposite directions to form a right-handed double helix (Figure 2A). The alternating phosphate and deoxyribose groups are exposed on the outside of the DNA helix, while the hydrophobic, nonpolar nucleobases remain tightly stacked inside, minimizing unfavorable interactions with water (Figure 2B) *(3)*. This arrangement not only thermodynamically stabilizes the double helix but also creates ridges defined by the sugar-phosphate backbone. Between these ridges lie the minor and major grooves, whose dimensions and chemical properties provide the primary interface for DNA-binding proteins, thereby determining interaction specificity (Figure 2C) *(4, 5)*.

Further studies of structural dynamics revealed that the DNA molecule is far more complex and dynamic than previously assumed. Beyond the canonical double helix, studies have not only confirmed alternative helical forms such as A-DNA, left-handed Z-DNA, and triple helices but have also identified non-helical conformations, most notably i-motifs and G-quadruplexes *(7–11)*.

G-QUADRUPLEXES

Advances in genome-wide technologies have enabled in-depth genome analysis, revealing that DNA contains numerous G-rich regions. These sequences can form G-quadruplexes (G4s), a secondary structure of nucleic acids where guanine bases are connected via hydrogen bonding in planar G-quartets along their Hoogsteen edge

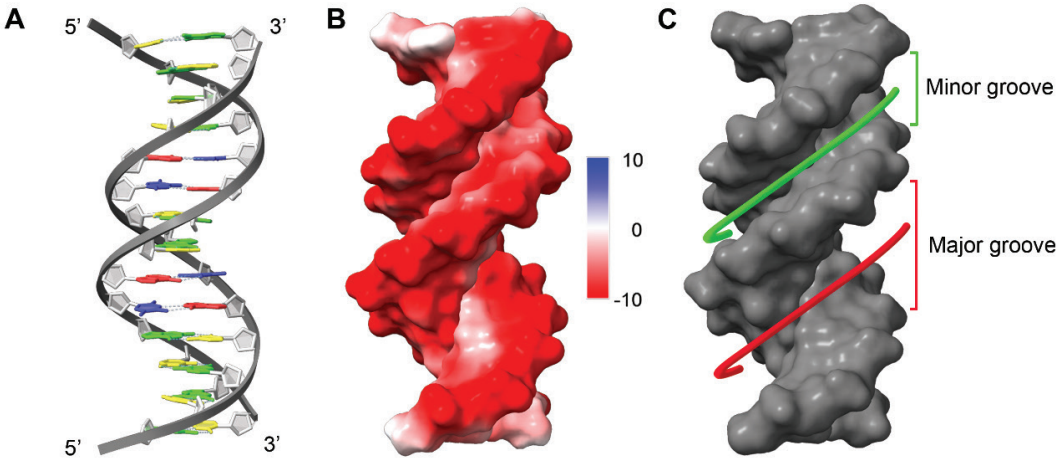

Figure 2. *(**A**) Cartoon representation of the DNA double helix with the nucleobases represented in distinct colors: adenine – red, thymine – blue, guanine – green and cytosine – yellow. The 5' and 3' ends are labeled (PDB: 2M2C; Ref. 6). (**B**) Molecular surface of the double helix colored according to the molecule's electrostatic potential. The negatively charged backbone results in negative electrostatic potential shown in red. (**C**) Visualization of the minor and major grooves shown in green and red, respectively.*

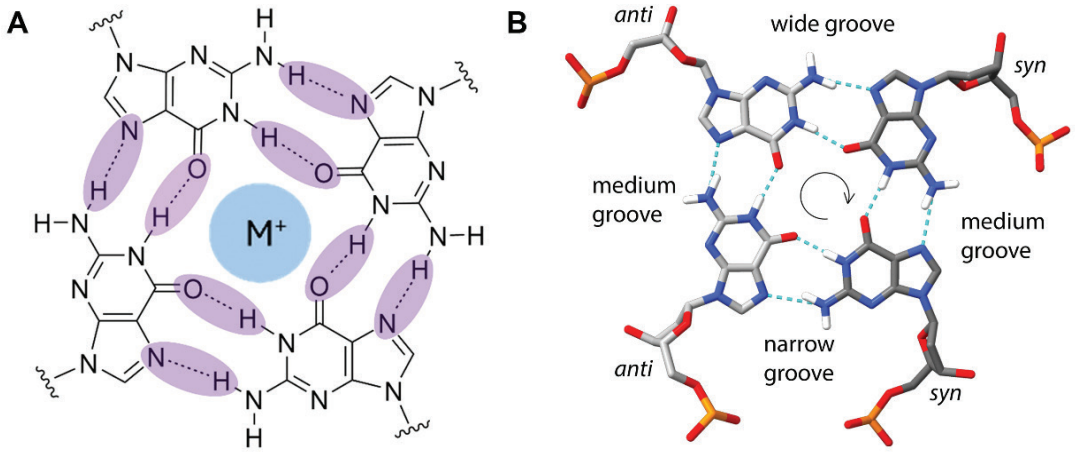

Figure 3. (**A**) *Schematic structure of a G-quartet with Hoogsteen base pairing geometry. Hydrogen bonds are highlighted in purple; metallic ion is highlighted in blue.* (***B***) *Groove size in relation to conformation across adjacent glycosidic bonds. Anti-anti and syn-syn transitions within a G-quartet result in medium grooves, anti-syn in wide grooves and syn-anti in narrow grooves (PDB: 6ZX7; Ref. 17).*

(Figure 3A) *(7, 10, 11)*. Multiple G-quartets can stack together to form a G4 core. This motif is further stabilized by the π-π stacking interactions and the presence of monovalent or seldomly divalent cations in the central cavity. The most prominent are potassium ions (K^+), followed by sodium (Na^+) or some other cations *(12, 13)*. In 2005, approximately 370 000 sequences within the human genome were reported to have the potential to form G4 structures *(14, 15)*. Recent studies however suggest that this number exceeds 700 000, making them an important field of research *(16)*.

Being built from at least four G-tracts and interconnecting sequences called loops of varying lengths and shapes, the generally accepted consensus sequence for a unimolecular G4 is $G_{\geq 2}N_{1\text{-}7}G_{\geq 2}N_{1\text{-}7}G_{\geq 2}N_{1\text{-}7}G_{\geq 2}$, where *N* stands for any nucleotide *(18)*. Additionally, intermolecular G4s with the G-tracts originating from more than one polynucleotide chain also exist. Although most G4s exhibit high levels of apparent similarity in their primary sequence, they can adopt diverse topologies, dictated by the pattern of strand polarities and the orientation of the loops, resulting in three main forms: parallel (4+0), hybrid (3+1) and antiparallel (2+2) *(19, 20)*. Loops can be classified as propeller, lateral and diagonal. Propeller loops connect two adjacent parallel strands, whereas lateral loops link neighboring antiparallel strands. On the other hand, the diagonal loops join two nonadjacent strands on the diagonal of the G4. Parallel G4s are therefore generally defined by three propeller loops connecting the upper and lower G-quartets, with all guanines adopting the *anti*-conformation, while hybrid G4s can contain all three loop types. Additionally, antiparallel topology can be further divided into chair conformations, which contain only lateral loops, and basket conformations, which have both lateral and diagonal loops (Figure 4) *(21)*. However, other less common loop types also exist, such as the V-shaped loop, a propeller loop that directly connects + two strands without any additional residues *(22)*. Furthermore, several alternative G4 assemblies have been identified, expanding the known structural diversity of these motifs. Examples include bulges that can form in the G-tracts, vacant guanine positions in G-quartets that may be compensated by nucleotides from snapback loops and other atypical arrangements. Moreover, while G4s most often adopt a right-handed helical twist, other variants such as left-handed and hybrid conformations have also been described *(23)*.

Another key feature of the G4 structure is the presence of grooves between two adjacent strands, which are important for potential G4 recognition or targeting. Their size is defined by the *anti* or *syn* conformation of the nucleotides of adjacent guanines in the G-quartet, forming wide, narrow, or medium grooves (Figure 3B) (*24*).

Thermal stability of G4s depends on multiple factors, including the number of stacked G-quartets, loop length and composition, their topology and the presence and nature of metal-cations. *In vitro*, G4s constantly compete with duplex DNA formation with the latter prevailing under near-physiological conditions. *In vivo* however, specific factors within the nucleus can shift this equilibrium to favor quadruplex formation *(27, 28)*. Initial studies suggested that G4s could protect chromosome ends and prevent DNA repair enzymes from mistakenly recognizing them as damaged *(28, 29)*. The first reported crystal structure of G4s in the telomeric region of human genome in 2002 further supported the idea that G4s could also form *in vivo* and may also be of biological importance *(30, 31)*. Besides telomeres, potential G4 forming sequences are also highly enriched in key human regulatory regions of DNA, such as active promoters, enhancers and 5'-untranslated regions. This implies their involvement in transcription regulation, as the sequence of the promoter region combined with its spatial (3D) structure determines

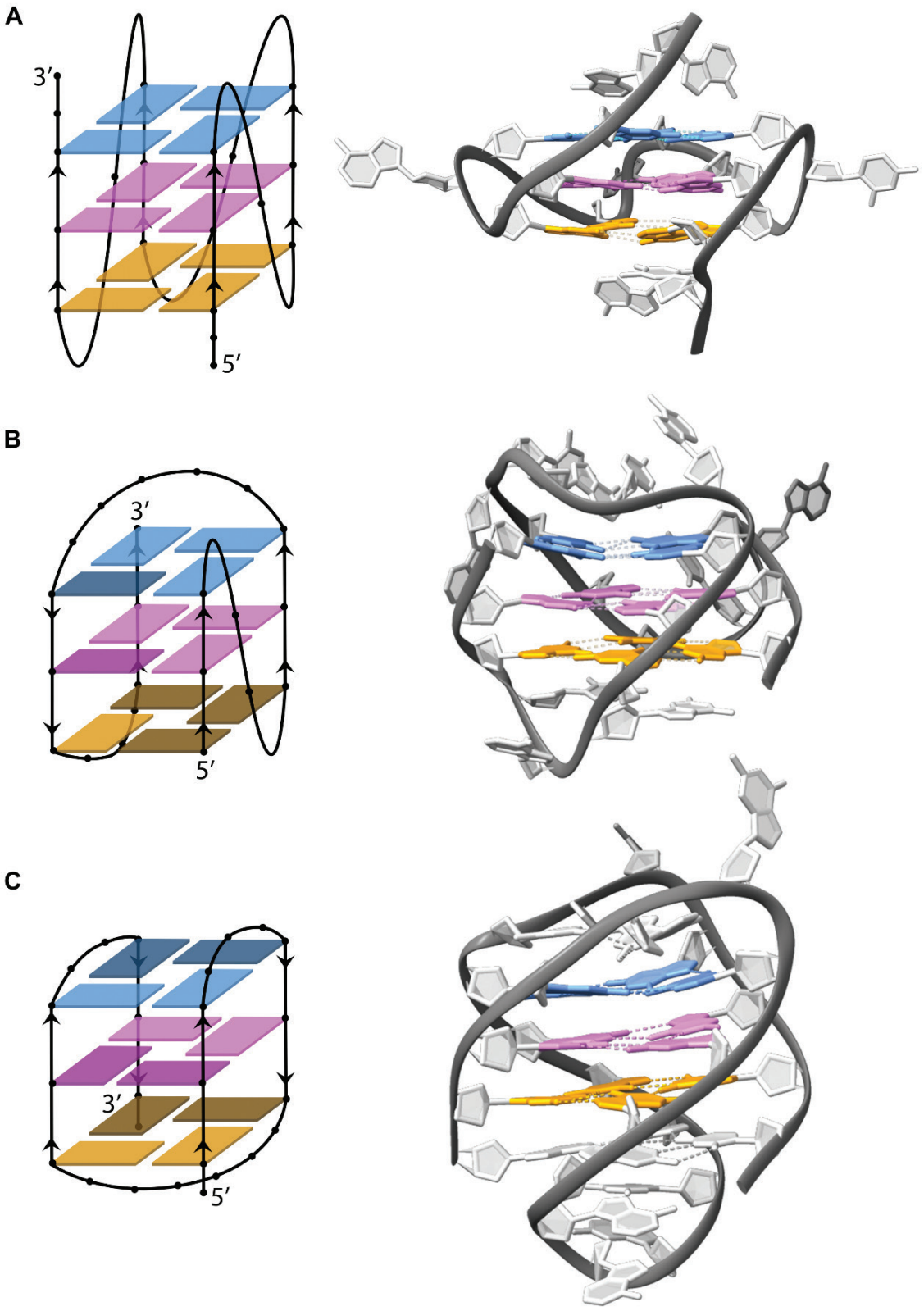

Figure 4. *Schematic representation and structure of three distinct G4s and their inherent polymorphism. The backbone of the structural models is presented as a dark gray cartoon and the three G-quartets forming the G4 core are shown in blue, violet and orange. The schematic representations indicate chain trajectory starting at the 5'-end, providing the basis for topological classification. (**A**) G4 with parallel (4+0) topology (PDB: 2LEE; Ref. 25). (**B**) G4 with hybrid (3+1) topology (PDB: 2LOD; Ref. 26). Darker colorings of planar guanine residues indicate syn conformation of the corresponding glycosidic bond. (**C**) G4 with antiparallel (2+2) topology in basket conformation (PDB: 6ZX7; Ref. 17).*

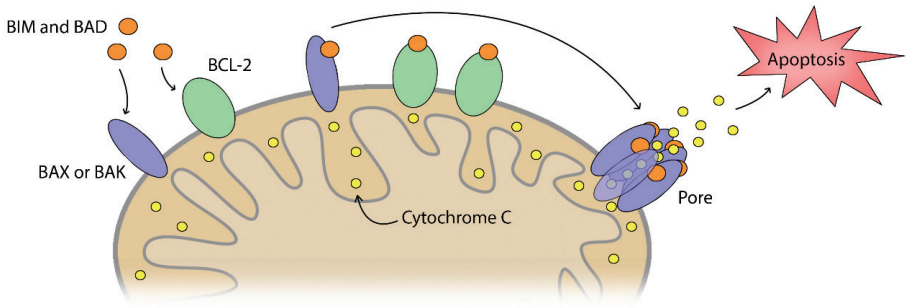

Figure 5. *Schematic representation of intrinsic apoptosis involving BCL-2. High concentrations of proteins BIM and BAD can saturate the anti-apoptotic BCL-2, especially when BCL-2 expression is mitigated. Consequently, the pore-forming ability of BAX and BAK is restored, hence causing the release of apoptosis-inducing cytochrome c into the cytosol.*

the transcription factor binding specificity and thus controls expression levels *(19, 32)*. Initially, G4s were proposed to act as transcription repressors, but this view proved superficial, as newer studies suggest that most G4s may actually function as transcriptional enhancers (*31*). A well-studied example of G4 impact on protein expression is the BCL-2 promoter region.

APOPTOSIS AND BCL-2

Mitochondrial or intrinsic apoptosis begins with the permeabilization of the outer mitochondrial membrane, resulting in the release of apoptogenic factors such as cytochrome c into the cytosol. These, in turn, activate the caspase cascade, causing the disassembly of cellular components, thus killing the cell and preparing it for phagocytosis *(33)*. The initial permeabilization can occur spontaneously as a result of excessive oxidative stress damaging and rupturing the membrane or as a response to cell signalization, for example when extensive DNA damage is detected *(34)*.

Intrinsic apoptosis regulation includes numerous proteins. Pro-apoptotic factors such as BIM and BID bind to and activate BAX or BAK membrane proteins to form a pore in the outer mitochondrial membrane which facilitates the cytochrome c escape from the intermembrane region and induces cell death. On the other hand, anti-apoptotic factors balance and regulate the effects of their pro-apoptotic counterparts. The most prominent anti-apoptotic factor in human cells is B-cell lymphoma 2 (BCL-2), also a membrane protein that competitively binds BIM and BID proteins, hence inhibiting BAX and BAK's pore-forming ability (Figure 5) *(33)*. To preserve homeostasis, both pro- and anti-apoptotic proteins must be precisely regulated. Upregulation of BCL-2 is often associated with a downturn in intrinsic apoptosis levels, impairing the cell's capability to respond to serious malfunctions. Moreover, overexpression of BCL-2 is frequently observed in cancer cells, contributing to the importance of understanding BCL-2 regulation as a potent oncogenic agent *(35)*.

THE G-RICH BCL-2 PROMOTER

The BCL-2 gene has 3 promoters, with P1, the one furthest upstream from the coding sequence, being the focal point of most studies due to its G-rich regions. In the current study we examined the P1 extended promoter sequence spanning from –499 nt upstream to +100 nt downstream relative to the transcription start site (TSS), corresponding to –1302 to –1901 nt relative to the start of the coding sequence, acquired from the Eukaryotic Promoter Databank (EPD). According to the JASPAR database, the P1 promoter contains potential binding sites for numerous transcription factors, several of which are strongly associated with apoptosis regulation and are therefore shown in Figure 6 *(36)*. Highlighted in bold text are c-AMP response element (CRE) and binding sites specific for Wt1 and Sp1 transcription proteins, which are among the more recognized contributors influencing BCL-2 expression *(37–39)*.

The P1 promoter consists of several G-rich regions, where the local guanine content reaches or even exceeds 50 % (Figure 6). Generally, such regions have a high propensity to form secondary structures, especially G4s. Previous studies on the BCL-2 promoter identified some G-quadruplex-forming regions, primarily P1G4, Pu39 (also known as bcl2G4-1), P32 and bcl2G4-2 (Figure 6). Within these regions, a variety of G4s have been characterized (Table 1). Furthermore, numerous studies have reported that G4 formation in these regions strongly contributes to the regulation of BCL-2 transcription, not only by disrupting the double helix but also through interactions with certain G4-specific binding proteins.

The P1G4 region is known to form two distinct G4s (form A and form B), though they exhibit a high-degree of similarity as they share the same parallel topology and include a 12- and a 11-nucleotide long hairpin loop, respectively. The most evident distinction between the two G4s is that the form B exhibits a broken second strand that is completed with a guanine residue from the central loop. This not only modifies the position of the loop but

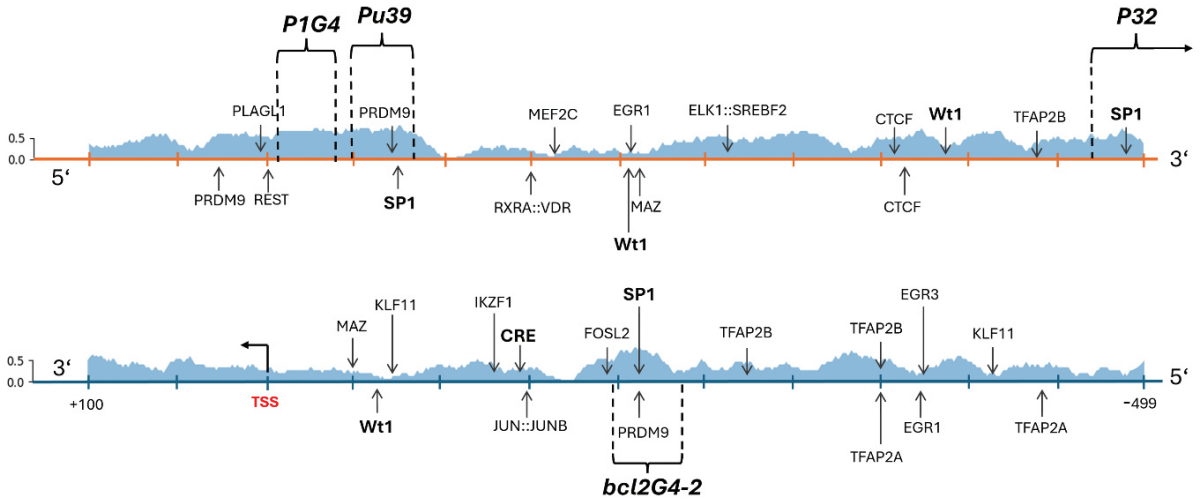

Figure 6. *Human BCL-2 P1 promoter from +100 to –499 relative to TSS. The sense strand is represented with an orange line, while the antisense strand carrying the BCL-2 gene is represented with a blue line. The promoter is annotated with several potential transcription factor binding sites that have a score of 400 or higher according to the JASPAR database. Although the CRE binding site does not score above the imposed limit, we included it for its documented importance on the BCL-2 regulation (38, 39). Additionally, the blue graph along both strands indicates the promoter's local guanine content calculated using a 15-nucleotide sized window, screening each of the two strands. The guanine content values are shown according to the scale on the y-axis, located on the left side of the graph. The black brackets indicate reported G-rich regions known to form G4s.*

also shortens its length highlighting the potential regulatory function of the two entities *(40)*.

Pu39 includes seven G-tracts of different lengths, enabling the formation of multiple G4s, the most prevalent being MidG4 and Pu30. MidG4 adopts an antiparallel topology, while Pu30 is a parallel G4 and includes a 13-nucleotide middle loop *(40)*. Pu30 is thermodynamically more stable than MidG4, however it is kinetically less favored due to its longer loops. The structural differences between both G4s are suggestive of the distinct interactions they could form with various effector proteins, possibly resulting in altered BCL-2 expression *(40)*.

Both P1G4 and Pu39 are located immediately upstream of the TSS and, despite their close proximity, appear to fold independently, hinting at their importance for regulation of the *BCL-2* gene. Moreover, the G4 formed in the Pu39 region serves as a transcription activator, while P1G4 exhibits a potent repressive function, which is further intensified with ligand stabilization of the G4 architecture, thus hindering BCL-2 expression. Additionally, the long loops formed by both G4s can serve as specific recognition sites for small regulatory molecules, enabling further transcription modulation *(40)*.

The P32 region is located the furthest from the TSS among the aforementioned G-rich regions. Nonetheless, it is believed to influence BCL-2 expression. With three G_4-tracts and one G_5-tract the sequence can fold into a single highly stable four-quartet G4 with hybrid topology *(41)*. Its well-defined structural characteristics imply their significance for facilitating specific quadruplex-protein

Table 1. *G-rich regions of the BCL-2 P1 promoter and the sequences corresponding to the known G4 structures.*

G-rich region	G4 sequence (5' to 3')*
P1G4	CGGGCGGGAGCGCGGCGGGCGGGCGGGCA
form A	CGGGCGGGAGCGCGGC**GGG**CGGGCGGGCA
form B	CGGGCGGG**G**AGCGCGGC**GG**GCGGGCGGGCA
Pu39 / bcl2G4-1	AGGGGCGGGCGCGGGAGGAAGGGGGCGGGAGCGGGGCTG
MidG4	GGGCGCGGGA**GG**AA**GG**GGGCGGG
Pu30	A**G**GGGCGGGCGC**GGG**A**GG**AA**GG**GGGCGGGA
P32	TGGGGTCCGCGACGGGGTGGGGGCTCCCCGGGG
bcl2G4-2	CCGGGCCAGGGAGCGGGGCGGAGGGGGCGGTCGGGT

*The G-tracts shown in red are or potentially could be a part of the G4 core. The G-tracts shown in black do not form G4 core.

interactions. In fact, the presence of certain ligands slightly modifies the G4 structure and results in lower BCL-2 expression levels, thus asserting P32 as transcription activating regulatory element *(41)*.

Lastly, bcl2G4-2 consists of seven G-tracts with the first, second, third and fifth runs being the main contributors for the G4 formation *(42)*. However, the overall stability of the G4 has empirically been shown to vary, implicating intricate forming patterns and conformational diversity. Still, limited information exists about the bcl2G4-2 G4 and its potential role in BCL-2 regulation with further research required for thorough understanding *(43)*. Furthermore, reports on this region suggest that bcl2G4-2 oligonucleotide is unable to form G4s when present as a double stranded DNA. G4 formation could be induced only after stabilizing the complementary C-rich strand *(42)*. However, it should be taken into consideration that *in vitro* conditions do not always reflect the *in vivo* environment accurately. *In vivo*, many proteins twist, unwind and otherwise modulate the DNA structure, hence promoting G4 formation without exogenous interference *(44)*.

POTENTIAL LONG-LOOPED G4S IN THE BCL-2 PROMOTER

As discussed in the previous chapter, the BCL-2 promoter is well documented for its G4-forming properties, providing insight into the complexity and diversity of G4 structures. Broken strands, bulges, mismatches, and long

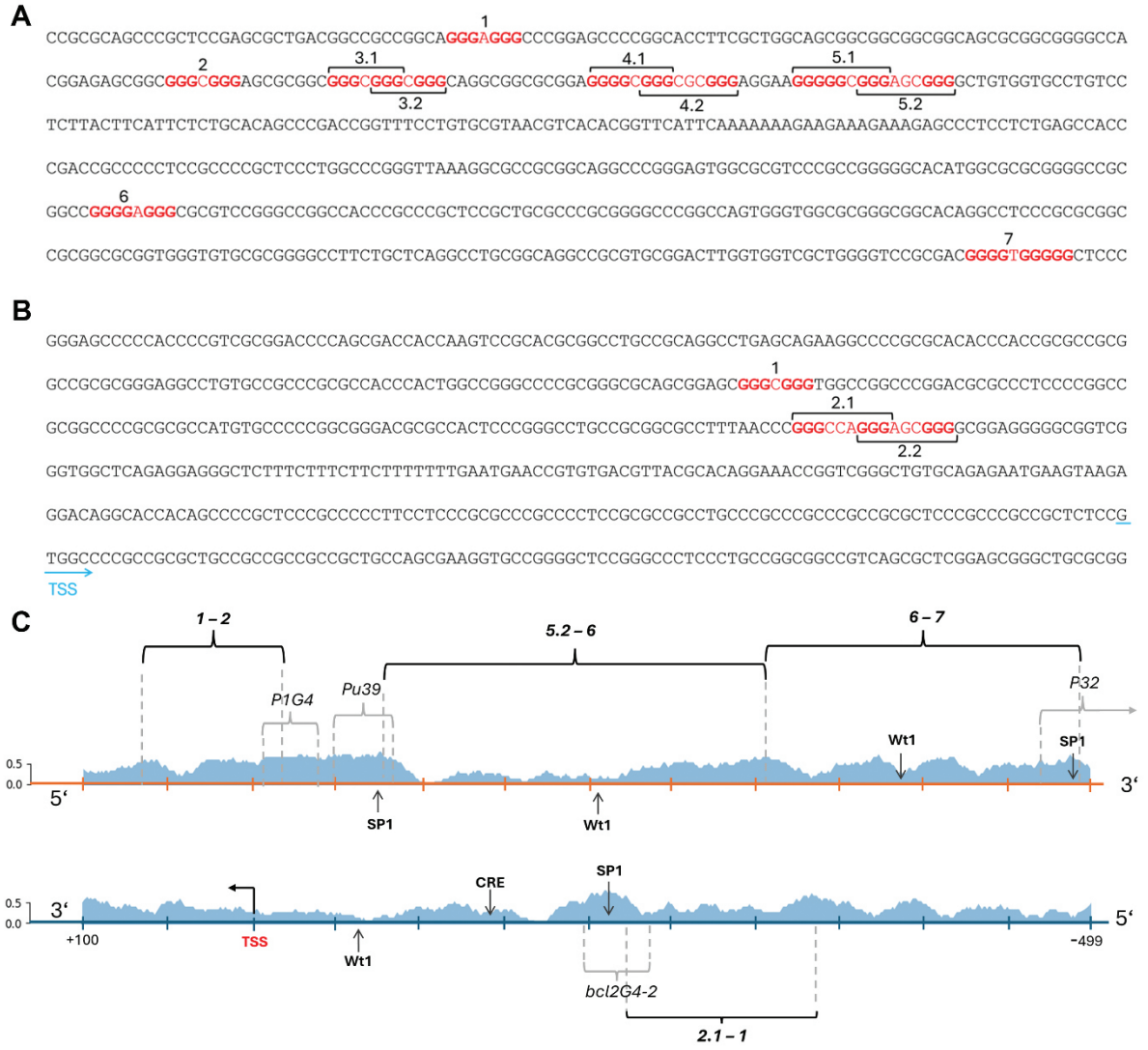

Figure 7. *(**A**) BCL-2 promoter sense strand with identified double-G-tract motifs shown in red. These motifs are labeled with numbers increasing from the 5'-end. If two double-G-tracts are overlapping, they are numbered with the same leading number. In total, the sense strand contains ten double-G-tracts with three overlaps. (**B**) Antisense strand, also starting with the 5'-end. The sequence contains three double-G-tracts with one overlap. (**C**) The position of suggested long-looped G4s in context of the BCL-2 promoter annotated with dark curvy brackets. The numbers on the brackets indicate the indexes of double-G-tracts that join to form the G4 core.*

hairpin loops are among the characteristics of G4s that defy the commonly accepted consensus sequence (*18*). For future reference, such G4s will be referred to as non-standard. Three long-looped G4s were identified in the P1G4 and Pu39 regions alone, with the longest loops extending 11, 12, and 13 nucleotides (*40, 45*). Here, we hypothesize that G4s with even longer loops, possibly reaching triple-digit lengths, could exist. These structures could facilitate promoter-enhancer interactions by bringing two sequentially distant elements into spatial proximity, adding significant complexity to the already intricate mechanism of gene regulation.

Deviations from the accepted G4 consensus sequence have long hindered research on long-looped G4s, as uncertainty about their thermodynamic stability and formation kinetics has raised doubts about their existence and relevance (*15*). Consequently, most methods for identifying G4 sequences either penalize long loops with a low overall score or do not detect them at all. As a result, standard software such as G4Hunter, QGRSMapper, and pqsfinder was not suitable for our bioinformatic analysis (*46*). For this reason, we wrote our own Python script to find all $G_3N_{0\text{-}3}G_3$ motifs, referred to as double-G-tracts, within a given sequence. Two such motifs should be able to form a three-quartet G4 core, with the intervening sequence serving as the connecting middle loop of the G4. Moreover, the idea that a long central loop does not compromise the stability of the G4 has already been discussed and affirmed in several studies (*47*). Results of the search for double-G-tracts in both strands of the BCL-2 promoter are shown in Figures 7A and 7B.

Since the vast array of all possible G4s was not feasible to analyze in detail, we limited further analysis to G4s formed by two neighboring double-G-tracts so that the linker length between them was more than 15 nt. Doing so, we not only reduced the number of potential G4s to four – three on the sense strand and one on the antisense strand (Figure 7C), but also ensured the long loop includes no additional double-G-tracts that could compete for G4 core formation. To label the said G4s, we settled on a systematic principle that specifies the location of the G4 (sense or antisense strand) and lists its composing double-G-tracts in order 5' to 3'. For example, the G4 connecting double-G-tracts 1 and 2 on the sense strand is named GS[1-2].

To evaluate the suggested G4s, we utilized RNAfold set to DNA computing parameters to predict the secondary structure and provide estimates of the thermodynamic stability of the loop connecting the two double-G-tracts *(48)*. Should the loop exhibit high stability, it would be a strong indicator that the corresponding G4 could form *(49)*.

There are two main factors to consider when assessing RNAfold predictions. The first is the similarity between the minimal free energy (MFE) structure and the centroid structure. A centroid structure exhibits minimal base-pair distance to all structures in the thermodynamic ensemble (48), where the base-pair distance between two structures is defined as the number of unique base pairs present in one structure but absent in the other (50). As such, the centroid structure is a representative of the ensemble *(51)*. In turn, if the MFE structure closely resembles the centroid, it indicates that the most thermodynamically stable structure is also representative of the ensemble, thus increasing prediction reliability (*48*). The second key factor is ensemble diversity, defined as the average base-pair distance between all structures in the ensemble. Lower diversity reflects greater uniformity in the set, thereby improving reliability (*48*).

Table 2. *RNAfold predictions for the long middle loops of the selected G4s.*

G4	Central loop sequence (5' to 3')	Loop length	ΔG [kcal/mol]*	Ensemble diversity
GS[1-2]	CCCGGAGCCCCGGCACCTTCGCTGGCAGCGGCGGCGGC-GGCAGCGCGGCGGGGCCACGGAGAGCGGC	67 nt	–33.67	2.60
GS[5.2-6]	GCTGTGGTGCCTGTCCTCTTACTTCATTCTCTGCA-CAGCCCGACCGGTTTCCTGTGCGTAACGTCACACGGTTC ATTCAAAAAAAGAAGAAAGAAAGAGCCCTCCTCTGAGC-CACCCGACCGCCCCCTCCGCCCCGCTCCCTGGCCCGGG-TTAAAGGCGCCGCGGCAGGCCCGGGAGTGGCGCGTCCC-GCCGGGGGCACATGGCGCGCGGGGCCGCGGCC	220 nt	–43.69	76.45
GS[6-7]	CGCGTCCGGGCCGGCCACCCGCCCGCTCCGCTGCGCCC-GCGGGGCCCGGCCAGTGGGTGGCGCGGGCGGCACAG-GCCTCCCGCGCGGCCGCGGCGCGGTGGGTGTGCGC-GGGGCCTTCTGCTCAGGCCTGCGGCAGGCCGCGTGCG-GACTTGGTGGTCGCTGGGGTCCGCGAC	173 nt	–63.78	47.00
GA[1-2.1]	TGGCCGGCCCGGACGCGCCCTCCCCGGCCGCGGCCCC-GCGCGCCATGTGCCCCCGGCGGGACGCGCCACTCCC-GGGCCTGCCGCGGCGCCTTTAACCC	98 nt	–27.49	31.98

*The value corresponds to the minimal free energy (MFE) of the thermodynamic ensemble.

The results of the RNAfold prediction for the proposed loops of G4s in the BCL-2 promoter are summarized in Table 2. The predictions for GS[5.2-6], GS[6-7], and GA[1-2.1] are not relevant, as all show excessively high ensemble diversity, indicating there is no single well-defined free energy minimum achievable with the corresponding sequence. Additionally, the MFE structures of the GS[5.2-6] and GA[1-2.1] loops differ significantly from their centroid counterparts, suggesting it is highly unlikely that these G4s would form. In contrast, GS[1-2] displays exceptionally low ensemble diversity, indicative of the confidence of the predicted hairpin-like structure. Moreover, the MFE model closely resembles the centroid structure, with a base-pair distance of just 1, further validating the prediction. The reason for such high certainty may originate from the 5'-GGCGGCGGCGG-3' sequence located in the middle of the loop, which RNAfold recognized to be able to form a two-quartet G4 (Figure 8A).

Additionally, when RNAfold was instructed not to incorporate G4s into its calculation and prediction, the resulting ensemble diversity was 22.88, with significant mismatching between the MFE and centroid structures, suggesting that the loop is stable only when the G4 is present (Figure 8B). If this is the case, it could represent a potential GS[1-2] folding mechanism, as the energy required to form the shorter G4 in the middle loop would be significantly lower than that required for GS[1-2] itself. Its formation could in turn facilitate further loop stabilization and ultimately lead to the formation of the long-looped GS[1-2] *(49)*. Furthermore, the structural integrity of the GS[1-2] loop would result in well-defined structural characteristics which could be targeted by specific proteins and provide GS[1-2] with a potential biological role. Being located directly at the TSS (Figure 7C), its formation could be associated with the transcription bubble, possibly promoting transcription by stabilizing the complementary strand *(52)*. In addition, the 3' double-G-tract of GS[1-2] is also a key structural component of P1G4, a potent repressor of transcription. Therefore, the GS[1-2] formation would also facilitate transcription indirectly via the disruption of P1G4 inhibitor architecture.

It is important to note that although RNAfold is a useful tool for assessing stability and predicting secondary structure, its results should be interpreted with caution. RNAfold relies on an energy model that approximates the unknown complete model, so it may not perform well in some cases and can be error-prone (*51*).The limitations of the existing energy model are also evident when accounting for G4 formation, as the program searches only for the G4s which follow the consensus sequence and penal-

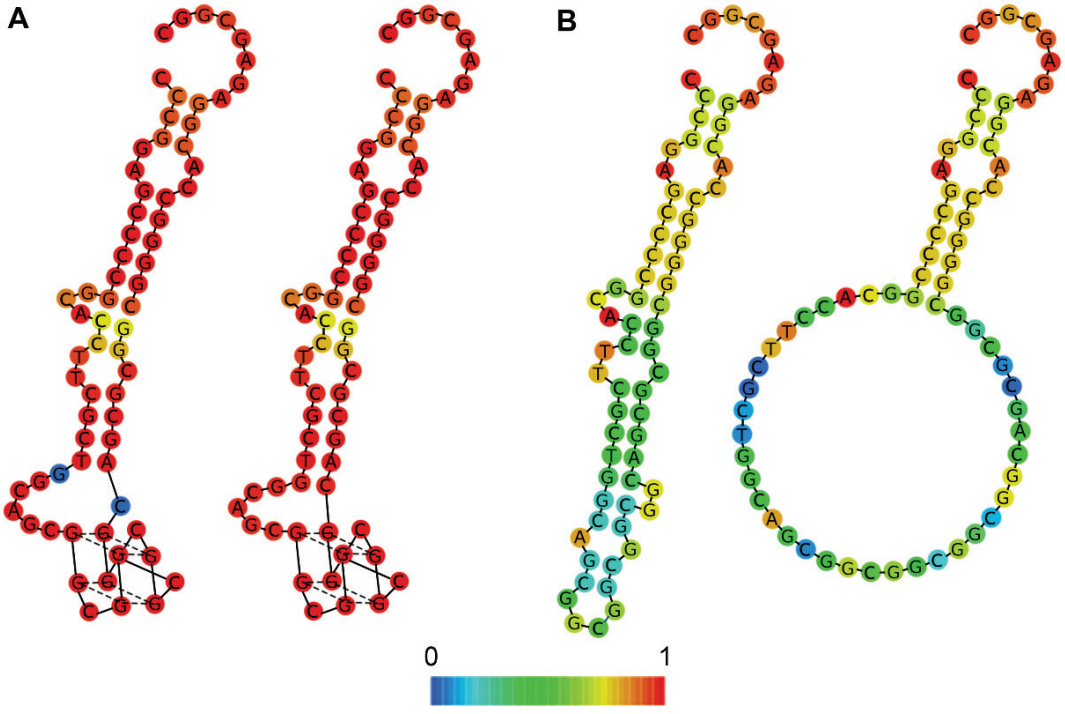

Figure 8. *RNAfold predictions for the longest loop of GS[1-2], color coded according to the base-pairing probability. Although RNAfold is unable to predict the topology of the additional G4 within GS[1-2], we manually constructed and illustrated a representative topology to provide a more intuitive visualization. (**A**) The MFE structure is shown on the left, while the centroid structure is on the right. Both loop structures include the G4 formed at the tip of the central loop of GS[1-2] with the dashed lines representing Hoogsteen base pairing within the two G-quartets. (**B**) The predicted loop structure was generated without allowing the formation of an additional G4 within the loop. The resulting minimum free energy (MFE) structure on the left differs significantly from the centroid structure on the right.*

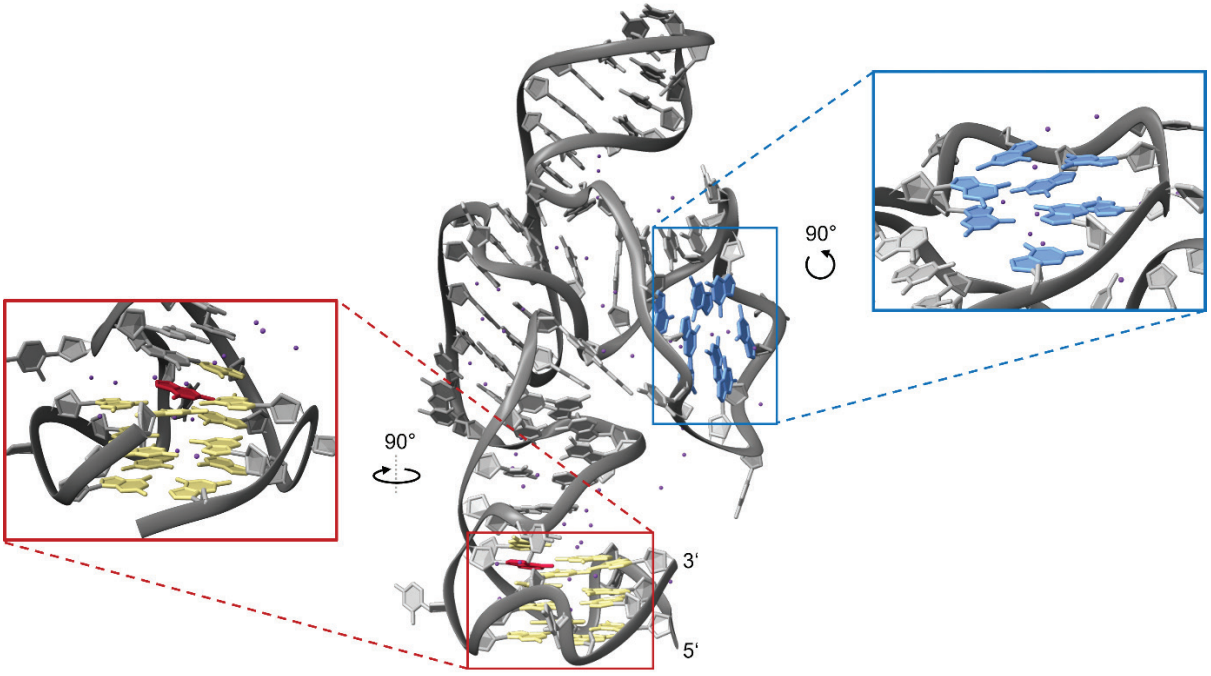

Figure 9. *An AlphaFold3 prediction of the 3D structure of GS[1-2] and its long middle loop shows that the loop adopts a hairpin-like structure with a two-quartet parallel G4, shown in blue. The main GS[1-2] G4 is also of parallel topology and consists of three G-quartets, indicated by the red frame. The yellow guanine residues are those from the originating double-G-tracts, which were hypothesized to form the G4. In this prediction, however, an additional guanine residue from the 3' end, shown in crimson, displaces one from the 5' double-G-tract and participates in forming of the top G-quartet. The purple spheres represent potassium ions.*

izes loop length in a logarithmic manner *(53)*. Consequently, RNAfold was unable to provide the prediction for the complete GS[1-2] due to the harsh energy restraints hindering G4 formation entirely. Moreover, RNAfold specializes in RNA secondary structure prediction, which may not always translate well to DNA sequences, even when DNA computing parameters are selected *(54)*. For these reasons, all results from this chapter remain strictly hypothetical and present only a general idea that should be further investigated experimentally.

With intention of gaining further insight into the three-dimensional architecture of the suggested G4s, we tried implementing AlphaFold3, which generally provides decent results for a variety of DNA and RNA sequences, including their potential G4 forming characteristics *(55)*. However, the predictions in our cases showed extremely low pLDDT confidence values, likely due to the extensive lengths of the nucleotide chains investigated, which allow for significant flexibility and structural dynamics. Consequently, the resulting predictions also vary significantly, failing to provide a clear answer regarding the structural characteristics of the suggested G4s. However, AlphaFold3 consistently predicts the formation of both GS[1-2] and the G4 in its central loop when the entry includes 50 potassium ion entities, indicating its folding potential (Figure 9). Naturally, this does not offer definitive proof regarding GS[1-2] formation, yet it provides a vague but important contribution to future research not only on BCL-2 promoter regulation, but also on long-looped G4s in general.

TARGETING OF QUADRUPLEX-DUPLEX STRUCTURES

Accumulating evidence of *in vivo* G4 formations, their involvement in gene expression and the connection with cancer and neurodegenerative diseases sparked significant interest in G4 targeting and influencing their formation. Various ligands capable of binding and stabilizing or destabilizing G4s have already been studied and present a potential for novel therapeutic and diagnostic strategies *(32, 56)*.

Ligands can interact with different regions of the G4 structure, specifically with the terminal G quartets of the core – an interaction known as end stacking – or with the G4 backbone, referred to as groove or loop binding (32, 56). Intercalation between the G-quartets is also possible, though this interaction is highly invasive and causes major structural changes to the G4 (57). Intercalation and end stacking are characteristic of large planar ligands that rely on stacking interactions with the aromatic guanine residues and electrostatic interactions with the corresponding heteroatoms, thus exhibiting low selectivity

(*32*). In contrast, the grooves and loops surrounding the core can provide some degree of discrimination, as the varying groove dimensions and loop composition offer distinct environments for ligand binding (*58*). Still, most ligand-G4 interactions are relatively indiscriminate, recognizing a greater number of G4s with similar topologies, which makes them effective multi-targeting agents (56). A promising solution for the lack of specificity is the use of binders specialized in targeting non-standard G4s, which can contribute unique structural features such as bulges and stem-loops that may play a key role in selective ligand design. However, the similarity among most G4 cores makes it difficult to achieve selectivity by targeting the stacked G quartets alone, consequently limiting their therapeutic potential (*18, 32*).

To address the challenges of specificity, several methods are being developed. The most prominent is a dual-specific targeting approach based on simultaneous recognition of the G4 architecture and a nearby duplex DNA sequence. This method combines the strong affinity of G4 binding with the selective targeting of duplex-binding ligands, fulfilling both essential requirements. Two such ligands can be conjugated to form a single compound capable of targeting specific quadruplex-duplex sites within the human genome, thereby serving therapeutic purposes *(59, 60)*.

The duplex domains of quadruplex-duplex motifs, however, do not always include just the double stranded DNA surrounding the G4s. Though not often observed,

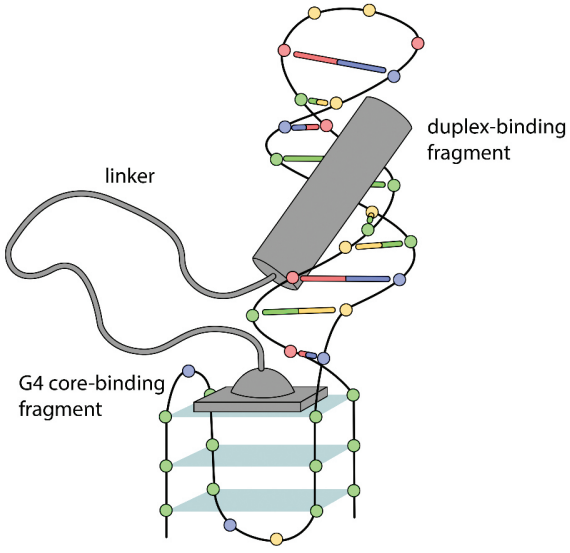

Figure 10. *Schematic representation of a potential selective G4 targeting mechanism. The ligand consists of two fragments joined by a linker. The G4 core-binding fragment specifically binds to the G4 core, while the duplex-binding fragment recognizes the groove and exposed functional groups corresponding to the target sequence in the stem loop of the G4. Together, the two fragments provide both high affinity and specificity for a particular long-looped G4.*

stem loops containing G4s have been characterized in numerous biologically significant regions of the genome (49), such as the myelin transcription factor 1-like (MYT1L) and PIM1 genes, for which high-resolution structures have also been resolved (61, 62). Sufficiently long stem loops exhibit a duplex structure resembling the architecture of the DNA double helix, implying that dual-specific targeting can also be achieved through stem-loop recognition (Figure 10). Studies have already shown that this approach enabled specific targeting of G4s in *hTERT* promoter region and could also be used to influence expression of other genes *(63)*. Furthermore, dual-specific targeting would also be the method of choice for long-looped G4 recognition as their stability is heavily dictated by the stabilizing interactions in the loop regions, predominantly resulting in duplex formation *(49)*. The GS[1-2] quadruplex, suggested in the previous chapter, also includes duplex-type architecture within its predicted loop. Should GS[1-2] in fact form, it too could be specifically targeted in this fashion, yielding potential guidelines for further research.

DISCUSSION

The variety of DNA architecture extends far beyond the commonly recognized double helix. DNA is a highly dynamic biological polymer whose structure depends greatly on the surrounding conditions. G4s are just one example of non-canonical DNA structures, highlighting the complexity of DNA's structural diversity and regulatory features. G-rich tracts are not randomly distributed throughout the genome; instead, they are concentrated mainly in regulatory and telomeric regions of chromosomes, offering insight into their biological functions. Numerous studies have shown that G4 formation significantly influences the expression of key developmental and signal transduction genes, demonstrating their role not only in degenerative and developmental disorders but also in carcinogenesis and related diseases.

Due to its central role in apoptosis regulation and significant G4-forming potential, the BCL-2 promoter has become a well-known model for G4-related research. With four recognized G-rich regions, G4s have been established as the main contributors to transcription regulation, either promoting or silencing BCL-2 expression. Recent studies supporting the relevance of non-standard G4s suggest that the proposed long-looped G4s within the BCL-2 promoter, specifically GS[1-2], play an important role in the BCL-2 regulatory mechanism. However, much information regarding their formation, stability, and potential functions remains unknown, and further comprehensive research is required for a deeper understanding.

G4s have, for their unique architecture and exhibited regulatory functions in relation to relevant human pathological conditions, become promising targets for drug

development. Although established methods exist for targeting G4s with large aromatic ligands, achieving specificity for desired G4s remains the most persistent challenge. Some ligands display a preference for particular G4 topologies; however, the abundance of G4s with similar topologies limits their therapeutic application. To address the urgent need for a practical and effective approach, dual-specific targeting enables simultaneous specific binding to nearby duplex structures without reducing binding affinity. This method could be generalized to all G4 structures by combining a robust, non-specific G4 core-binding fragment with a specific duplex-binding fragment designed to target the desired duplex sequence. Additionally, dual-specific targeting could be used to target long-looped G4s by recognizing their hairpin loops, potentially facilitating research on the important topic of non-standard G4s and their roles in various fundamental biological processes.

Acknowledgements: *The authors acknowledge the financial support from the Slovenian Research and Innovation Agency [ARIS, grants P1-0242, J1-60019 and GC-0005], the European Union programme HORIZON-WIDERA-2023-ACCESS-04, Pathways to Synergies MILESTONE project (grant GA 101159708), the CERIC-ERIC Consortium for the access to experimental facilities and financial support, and the GIGA NMR project co-financed by the Republic of Slovenia, the Ministry of Higher Education, Science and Innovation, and the European Union under the European Regional Development Fund."*

REFERENCES

1. WATSON J D and CRICK F H C 1953 Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. Nature 171(4356): 737–738 https://doi.org/10.1038/171737a0
2. SUGIMOTO N (ed) 2023 Handbook of Chemical Biology of Nucleic Acids, Springer Nature
3. NAKANO S, MIYOSHI D, and SUGIMOTO N 2014 Effects of Molecular Crowding on the Structures, Interactions, and Functions of Nucleic Acids. Chem Rev 114(5): 2733–2758 https://doi.org/10.1021/cr400113m
4. SUZUKI M and YAGI N 1996 An In-the-Groove View of DNA Structures in Complexes with Proteins. J Mol Biol 255(5): 677–687 https://doi.org/10.1006/jmbi.1996.0055
5. BRÄNDÉN C-I and TOOZE J (eds) 1999 Introduction to protein structure, Garland Pub, New York
6. GHOSH A, KAR R K, JANA J, SAHA A, JANA B, KRISHNAMOORTHY J, KUMAR D, GHOSH S, CHATTERJEE S, and BHUNIA A 2014 Indolicidin Targets Duplex DNA: Structural and Mechanistic Insight through a Combination of Spectroscopy and Microscopy. ChemMedChem 9(9): 2052–2058 https://doi.org/10.1002/cmdc.201402215
7. HUD N V and PLAVEC J 2006 The Role of Cations in Determining Quadruplex Structure and Stability in Quadruplex Nucleic Acids. In Quadruplex Nucleic Acids, Neidle, S.; Balasubramanian, S., Eds. The Royal Society of Chemistry: Cambridge, 2006; pp 100–130.
8. ŠKET P and PLAVEC J 2015 Diversity of DNA and RNA G-quadruplex structures. In *Biological Relevance & Therapeutic Applications of DNA- & RNA-Quadruplexes*, David, M., Ed. Future Science Ltd: 2015; pp 22–36.
9. Richter S (ed) 2019 G-quadruplex and Microorganisms, the special issue of Molecules, p 1–169
10. PLAVEC J 2020 Quadruplex targets in neurodegenerative diseases. In *Annual Reports in Medicinal Chemistry*, Neidle, S., Ed. Academic Press: 2020; Vol. 54, pp 441–483.
11. PLAVEC J 2023 NMR Study on Nucleic Acids. In *Handbook of Chemical Biology of Nucleic Acids*, Sugimoto, N., Ed. Springer Nature Singapore: Singapore, 2023; pp 169–212.
12. DAVIS J T 2004 G-Quartets 40 Years Later: From 5′-GMP to Molecular Biology and Supramolecular Chemistry. Angew Chem Int Ed 43(6): 668–698 https://doi.org/10.1002/anie.200300589
13. PLAVEC J 2009 Metal Ion Coordination in G-Quadruplexs. In Metal Complex-DNA Interactions, Hadjiliadis, N.; Sletten, E., Eds. Blackwell Publising Ltd: Cambridge, 2009; pp 55–93.
14. TODD A K, JOHNSTON M, and NEIDLE S 2005 Highly prevalent putative quadruplex sequence motifs in human DNA. Nucleic Acids Res 33(9): 2901–2907 https://doi.org/10.1093/nar/gki553
15. HUPPERT J L and BALASUBRAMANIAN S 2005 Prevalence of quadruplexes in the human genome. Nucleic Acids Res 33(9): 2908–2916 https://doi.org/10.1093/nar/gki609
16. CHAMBERS V S, MARSICO G, BOUTELL J M, DI ANTONIO M, SMITH G P, and BALASUBRAMANIAN S 2015 High-throughput sequencing of DNA G-quadruplex structures in the human genome. Nat Biotechnol 33(8): 877–881 https://doi.org/10.1038/nbt.3295
17. BIELSKUTĖ S, PLAVEC J, and PODBEVŠEK P 2021 Oxidative lesions modulate G-quadruplex stability and structure in the human BCL2 promoter. Nucleic Acids Res 49(4): 2346–2356 https://doi.org/10.1093/nar/gkab057
18. JANA J, MOHR S, VIANNEY Y M, and WEISZ K 2021 Structural motifs and intramolecular interactions in non-canonical G-quadruplexes. RSC Chem Biol 2(2): 338–353 https://doi.org/10.1039/D0CB00211A
19. SATO K and KNIPSCHEER P 2023 G-quadruplex resolution: From molecular mechanisms to physiological relevance. DNA Repair 130: 103552 https://doi.org/10.1016/j.dnarep.2023.103552
20. VARSHNEY D, SPIEGEL J, ZYNER K, TANNAHILL D, and BALASUBRAMANIAN S 2020 The regulation and functions of DNA and RNA G-quadruplexes. Nat Rev Mol Cell Biol 21(8): 459–474 https://doi.org/10.1038/s41580-020-0236-x
21. CADONI E, DE PAEPE L, MANICARDI A, and MADDER A 2021 Beyond small molecules: targeting G-quadruplex structures with oligonucleotides and their analogues. Nucleic Acids Res 49(12): 6638–6659 https://doi.org/10.1093/nar/gkab334
22. CRNUGELJ M, SKET P, and PLAVEC J 2003 Small change in a G-rich sequence, a dramatic change in topology: new dimeric G-quadruplex folding motif with unique loop orientations. J Am Chem Soc 125(26): 7866–7871 https://doi.org/10.1021/ja0348694
23. CHUNG W J, HEDDI B, SCHMITT E, LIM K W, MECHULAM Y, and PHAN A T 2015 Structure of a left-handed DNA G-quadruplex. Proc Natl Acad Sci U S A 112(9): 2729–2733 https://doi.org/10.1073/pnas.1418718112
24. WEBBA DA SILVA M 2007 Geometric Formalism for DNA Quadruplex Folding. Chem Eur J 13(35): 9738–9745 https://doi.org/10.1002/chem.200701255
25. TRAJKOVSKI M, WEBBA DA SILVA M, and PLAVEC J 2012 Unique Structural Features of Interconverting Monomeric and Dimeric G-Quadruplexes Adopted by a Sequence from the Intron

of the N-myc Gene. J Am Chem Soc 134(9): 4132–4141 https://doi.org/10.1021/ja208483v

26. MARUŠIČ M, ŠKET P, BAUER L, VIGLASKY V, and PLAVEC J 2012 Solution-state structure of an intramolecular G-quadruplex with propeller, diagonal and edgewise loops. Nucleic Acids Res 40(14): 6946–6956 https://doi.org/10.1093/nar/gks329

27. KÖNIG S L B, EVANS A C, and HUPPERT J L 2010 Seven essential questions on G-quadruplexes. Biomol Concepts 1(2): 197–213 https://doi.org/10.1515/bmc.2010.011

28. RHODES D and LIPPS H J 2015 G-quadruplexes and their regulatory roles in biology. Nucleic Acids Res 43(18): 8627–8637 https://doi.org/10.1093/nar/gkv862

29. BIFFI G, TANNAHILL D, MCCAFFERTY J, and BALASUBRAMANIAN S 2013 Quantitative visualization of DNA G-quadruplex structures in human cells. Nat Chem 5(3): 182–186 https://doi.org/10.1038/nchem.1548

30. PARKINSON G N, LEE M P H, and NEIDLE S 2002 Crystal structure of parallel quadruplexes from human telomeric DNA. Nature 417(6891): 876–880 https://doi.org/10.1038/nature755

31. ROBINSON J, RAGUSEO F, NUCCIO S P, LIANO D, and DI ANTONIO M 2021 DNA G-quadruplex structures: more than simple roadblocks to transcription? Nucleic Acids Res 49(15): 8419–8431 https://doi.org/10.1093/nar/gkab609

32. WANG K-B, WANG Y, DICKERHOFF J, and YANG D 2024 DNA G-Quadruplexes as Targets for Natural Product Drug Discovery. Engineering 38: 39–51 https://doi.org/10.1016/j.eng.2024.03.015

33. SAROSIEK K A and WOOD K C 2023 Endogenous and imposed determinants of apoptotic vulnerabilities in cancer. Trends Cancer 9(2): 96–110 https://doi.org/10.1016/j.trecan.2022.10.004

34. MARTENS M D, KARCH J, and GORDON J W 2022 The molecular mosaic of regulated cell death in the cardiovascular system. Biochim Biophys Acta - Mol Basis Dis 1868(1): 166297 https://doi.org/10.1016/j.bbadis.2021.166297

35. TESSOULIN B, PAPIN A, GOMEZ-BOUGIE P, BELLANGER C, AMIOT M, PELLAT-DECEUNYNCK C, and CHIRON D 2019 BCL2-Family Dysregulation in B-Cell Malignancies: From Gene Expression Regulation to a Targeted Therapy Biomarker. Front Oncol 8: https://doi.org/10.3389/fonc.2018.00645

36. EPD - BCL2_1 viewer, Available from: https://epd.expasy.org/cgi-bin/epd/get_doc?db=hgEpdNew&format=genome&entry=BCL2_1 (Last accessed 28-08-2025)

37. CHEEMA S K, MISHRA S K, RANGNEKAR V M, TARI A M, KUMAR R, and LOPEZ-BERESTEIN G 2003 Par-4 Transcriptionally Regulates Bcl-2 through a WT1-binding Site on the bcl-2 Promoter*. J Biol Chem 278(22): 19995–20005 https://doi.org/10.1074/jbc.M205865200

38. HECKMAN C A, MEHEW J W, and BOXER L M 2002 NF-κB activates Bcl-2 expression in t(14;18) lymphoma cells. Oncogene 21(24): 3898–3908 https://doi.org/10.1038/sj.onc.1205483

39. DONG L, WANG W, WANG F, STONER M, REED J C, HARIGAI M, SAMUDIO I, KLADDE M P, VYHLIDAL C, and SAFE S 1999 Mechanisms of Transcriptional Activation of bcl-2 Gene Expression by 17β-Estradiol in Breast Cancer Cells*. J Biol Chem 274(45): 32099–32107 https://doi.org/10.1074/jbc.274.45.32099

40. ONEL B, CARVER M, WU G, TIMONINA D, KALARN S, LARRIVA M, and YANG D 2016 A New G-Quadruplex with Hairpin Loop Immediately Upstream of the Human BCL2 P1 Promoter Modulates Transcription. J Am Chem Soc 138(8): 2563–2570 https://doi.org/10.1021/jacs.5b08596

41. SUN H, XIANG J, SHI Y, YANG Q, GUAN A, LI Q, YU L, SHANG Q, ZHANG H, TANG Y, and XU G 2014 A newly identified G-quadruplex as a potential target regulating Bcl-2 expression. Biochim Biophys Acta – Gen Subj 1840(10): 3052–3057 https://doi.org/10.1016/j.bbagen.2014.07.014

42. ONYSHCHENKO M I, GAYNUTDINOV T I, ENGLUND E A, APPELLA D H, NEUMANN R D, and PANYUTIN I G 2009 Stabilization of G-quadruplex in the BCL2 promoter region in double-stranded DNA by invading short PNAs. Nucleic Acids Res 37(22): 7570–7580 https://doi.org/10.1093/nar/gkp840

43. CHETIA P and KUMAR A 2025 From unearthing to an intriguing cancer-fighting target: the human BCL-2 promoter G-quadruplex and i-motif. Biochim Biophys Acta – Rev Cancer 1880(5): 189391 https://doi.org/10.1016/j.bbcan.2025.189391

44. DELL'OCA M C, QUADRI R, BERNINI G M, MENIN L, GRASSO L, RONDELLI D, YAZICI O, SERTIC S, MARINI F, PELLICIOLI A, MUZI-FALCONI M, and LAZZARO F 2024 Spotlight on G-Quadruplexes: From Structure and Modulation to Physiological and Pathological Roles. Int J Mol Sci 25(6): 3162 https://doi.org/10.3390/ijms25063162

45. AGRAWAL P, LIN C, MATHAD R I, CARVER M, and YANG D 2014 The Major G-Quadruplex Formed in the Human BCL-2 Proximal Promoter Adopts a Parallel Structure with a 13-nt Loop in K+ Solution. J Am Chem Soc 136(5): 1750–1753 https://doi.org/10.1021/ja4118945

46. PUIG LOMBARDI E and LONDOÑO-VALLEJO A 2020 A guide to computational methods for G-quadruplex prediction. Nucleic Acids Res 48(1): 1–15 https://doi.org/10.1093/nar/gkz1097

47. GUÉDIN A, GROS J, ALBERTI P, and MERGNY J-L 2010 How long is too long? Effects of loop size on G-quadruplex stability. Nucleic Acids Res 38(21): 7858–7868 https://doi.org/10.1093/nar/gkq639

48. GRUBER A R, LORENZ R, BERNHART S H, NEUBÖCK R, and HOFACKER I L 2008 The Vienna RNA Websuite. Nucleic Acids Res 36(suppl_2): W70–W74 https://doi.org/10.1093/nar/gkn188

49. LIM K W, JENJAROENPUN P, LOW Z J, KHONG Z J, NG Y S, KUZNETSOV V A, and PHAN A T 2015 Duplex stem-loop-containing quadruplex motifs in the human genome: a combined genomic and structural study. Nucleic Acids Res 43(11): 5630–5646 https://doi.org/10.1093/nar/gkv355

50. CHAN C Y and DING Y 2008 Boltzmann ensemble features of RNA secondary structures: a comparative analysis of biological RNA sequences and random shuffles. J Math Biol 56(1): 93–105 https://doi.org/10.1007/s00285-007-0129-z

51. DING Y, CHAN C Y, and LAWRENCE C E 2005 RNA secondary structure prediction by centroids in a Boltzmann weighted ensemble. RNA 11(8): 1157–1166 https://doi.org/10.1261/rna.2500605

52. BUGLIONE E, SALERNO D, MARRANO C A, CASSINA V, VESCO G, NARDO L, DACASTO M, RIGO R, SISSI C, and MANTEGAZZA F 2021 Nanomechanics of G-quadruplexes within the promoter of the KIT oncogene. Nucleic Acids Res 49(8): 4564–4573 https://doi.org/10.1093/nar/gkab079

53. LORENZ R, BERNHART S H, QIN J, SIEDERDISSEN C H zu, TANZER A, AMMAN F, HOFACKER I L, and STADLER P F 2013 2D Meets 4G: G-Quadruplexes in RNA Secondary Structure Prediction. IEEE/ACM Trans Comput Biol Bioinform 10(4): 832–844 https://doi.org/10.1109/TCBB.2013.7

54. BINET T, PADIOLLEAU-LEFÈVRE S, OCTAVE S, AVALLE B, and MAFFUCCI I 2023 Comparative Study of Single-stranded Oligonucleotides Secondary Structure Prediction Tools. BMC Bioinform 24(1): 422 https://doi.org/10.1186/s12859-023-05532-5

55. ABRAMSON J, ADLER J and DUNGER J et al. 2024 Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature 630(8016): 493–500 https://doi.org/10.1038/s41586-024-07487-w

56. FIGUEIREDO J, MERGNY J-L, and CRUZ C 2024 G-quadruplex ligands in cancer therapy: Progress, challenges, and clinical perspectives. Life Sci 340: 122481 https://doi.org/10.1016/j.lfs.2024.122481

57. GHOSH A, TRAJKOVSKI M, TEULADE-FICHOU M-P, GABELICA V, and PLAVEC J 2022 Phen-DC3 Induces Refolding of Human Telomeric DNA into a Chair-Type Antiparallel G-Quadruplex through Ligand Intercalation. Angew Chem Int Ed 61(40): e202207384 https://doi.org/10.1002/anie.202207384

58. ASAMITSU S, BANDO T, and SUGIYAMA H 2019 Ligand Design to Acquire Specificity to Intended G-Quadruplex Structures. Chem Eur J 25(2): 417–430 https://doi.org/10.1002/chem.201802691

59. ASAMITSU S, OBATA S, PHAN A T, HASHIYA K, BANDO T, and SUGIYAMA H 2018 Simultaneous Binding of Hybrid Molecules Constructed with Dual DNA-Binding Components to a G-Quadruplex and Its Proximal Duplex. Chem Eur J 24(17): 4428–4435 https://doi.org/10.1002/chem.201705945

60. NGUYEN T Q N, LIM K W, and PHAN A T 2017 A Dual-Specific Targeting Approach Based on the Simultaneous Recognition of Duplex and Quadruplex Motifs. Sci Rep 7(1): 11969 https://doi.org/10.1038/s41598-017-10583-9

61. TAN D J Y, WINNERDY F R, LIM K W, and PHAN A T 2020 Coexistence of two quadruplex–duplex hybrids in the PIM1 gene. Nucleic Acids Res 48(19): 11162–11171 https://doi.org/10.1093/nar/gkaa752

62. LIU L-Y, WANG K-N, LIU W, ZENG Y-L, HOU M-X, YANG J, and MAO Z-W 2021 Spatial Matching Selectivity and Solution Structure of Organic–Metal Hybrid to Quadruplex–Duplex Hybrid. Angew Chem Int Ed 60(38): 20833–20839 https://doi.org/10.1002/anie.202106256

63. KARNA D, LIANG L, SHARMA G, MANDAL S, ASAMITSU S, KAWAMOTO Y, HASHIYA K, BANDO T, SUGIYAMA H, and MAO H 2024 Modulation of dynamic DNA G-quadruplex structures in the hTERT promoter region by ligands. Nucleic Acids Res 52(18): 10775–10787 https://doi.org/10.1093/nar/gkae754