Nucleic Acids Research, 2025, **53**, gkaf164
https://doi.org/10.1093/nar/gkaf164
Chemical Biology and Nucleic Acid Chemistry

Imperfect G-quadruplex as an emerging candidate for transcriptional regulation

Sunipa Sarkar[1], Hisae Tateishi-Karimata[1,2,*], Tatsuya Ohyama[1], Naoki Sugimoto[1,*]

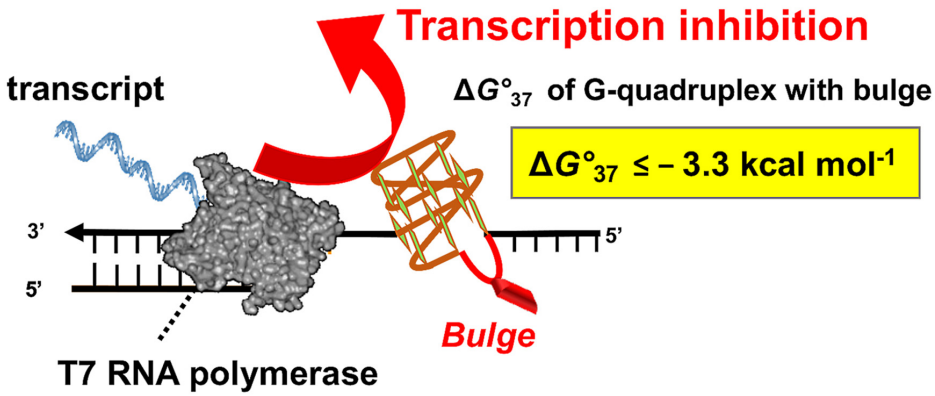
[1]Frontier Institute for Biomolecular Engineering Research (FIBER), Konan University, 7-1-20 Minatojima-minamimachi, Kobe 650-0047, Japan
[2]Graduate School of Frontiers of Innovative Research in Science and Technology (FIRST), Konan University, 7-1-20 Minatojima-minamimachi, Kobe 650-0047, Japan

[*]To whom correspondence should be addressed. Email: sugimoto@konan-u.ac.jp
Correspondence may also be addressed to Hisae Tateishi-Karimata. Email: tateishi@konan-u.ac.jp

Abstract

G-quadruplexes (G4s) with continuous G-tracts are well-established regulators of gene expression and important therapeutic targets for various diseases. However, bioinformatics analyses have identified G4-like sequences containing interrupted G-tracts, incorporating non-G nucleotides as bulges (buG4s). Our findings show that the stability of buG4s is significantly influenced by the bulge position and size within the G-tract, with bulges at the 5′ end exhibiting the highest stability. Moreover, a molecular crowding condition inducing by poly (ethylene glycol), providing a suitable intracellular environment, stabilizes buG4s, especially those with longer bulges, making their formation more pronounced. A transcription assay performed under crowding conditions revealed that the transcription arrested efficiency by buG4s is affected not only by stability but also by the position and size of the bulge. Based on these findings, we propose a model for the preliminary screening of buG4 sequences according to their stability, distinguishing functional sequences capable of transcriptional arrest ($\Delta G^\circ_{37} \leq -3.3$ kcal·mol^{-1}) from nonfunctional sequences ($\Delta G^\circ_{37} > -3.3$ kcal·mol^{-1}). This provides valuable insight into estimating the efficiency of target buG4 sequences in either arresting or facilitating transcription, presenting a novel approach and emphasizing buG4s as emerging therapeutic targets.

Graphical abstract



Introduction

Beyond the DNA double helix, nucleic acids can also adopt noncanonical structures, such as triplexes and tetraplexes including G-quadruplexes (G4s) and i-motifs [1–3]. Notably, G4s are found within the promoter and coding regions of genes associated with various diseases and are considered to play crucial roles in biological processes such as replication, transcription, and translation [4–8]. G4s are four stranded noncanonical structures of nucleic acids, formed by Hoogsteen hydrogen bonding between planar G bases which are further stabilized by K^+ or Na^+ [9]. The G4 forming sequences from the genome follows the sequence pattern $G_{3+}L1_{N=1-7}G_{3+}L2_{N=1-7}G_{3+}L3_{N=1-7}G_{3+}$, where the suffix indicates three or more continuous G bases, referred to as the G-tract which are separated by loops (L1, L2, or L3) with the number of loop nucleotides, N, varying between 1 and 7 nt [7]. Later, evidence was found for the existence of sequences with loop lengths >7 nt [10]. Bioinformatics analyses using more flexible algorithms have identified significant numbers of G4-like sequences that do not follow a perfect G4-forming sequence pattern [11–13]. These deviated structures from the perfect G4 are categorized as "imperfect G4"

Received: October 23, 2024. **Revised:** February 5, 2025. **Editorial Decision:** February 11, 2025. **Accepted:** February 19, 2025

[11, 14–17]. These novel G4 sequences include intramolecular G4s with vacancies in the G-tract (G-vacancy-bearing G4) or those with bulges arising from the incorporation of non-guanine bases in the G-tracts [14, 18]. Potential G4-forming sequences typically consist of four segments of consecutive guanines, whereas the buG4 sequence features G-tracts interrupted by bulges, thus resulting in reduced thermodynamic and mechanical stability when compared to G4s [14, 19]. Interestingly, imperfect G4 sequence patterns are more frequently localized in promoter and coding regions than perfect G4 sequence patterns [13]. However, very few imperfect G4 structures have been investigated thus far because of the complexity of the sequence and the low possibility of stable G4 formation. For example, a (3 + 1) hybrid structure of an adenine (A)-bulge-containing sequence was identified in the *PARP1* promoter, and its formation was confirmed by nuclear magnetic resonance (NMR) spectroscopy [20]. Additionally, a structure has been reported in the human *KRAS* gene nuclease-hypersensitive element (NHE-III1) region, featuring a thymine (T) bulge, that blocks transcriptional regulators [21]. A recent study revealed that the proportion of genes containing bulged sequences is about two times more than the genes containing perfect G4 sequences in transcription regulatory regions, such as intron–exon junctions, near the transcription start site (TSS), and around the transcription termination site (TTS) [11]. This finding indicates that these buG4 sequences may have significant transcription efficiency, similar to that of perfect G4. In previous studies, our group demonstrated that the transcription inhibition efficiency of G4 is dependent on the thermodynamic stability of the G4 structure and independent of topology [5, 22]. Subsequent research has revealed that the inhibition of replication by G4s is not solely dependent on stability but also on the topology of the structure [4]. The presence of bulges at various positions within the G-tract complicates the formation of stable G4 structures; consequently, their ability to regulate transcriptional processes can be significantly altered. Moreover, the inability of current search algorithms to accurately predict stable imperfect G4s with larger bulge sizes makes the genome-wide targeting of these sequences infeasible for therapeutic purposes. Therefore, understanding the effect of bulges on G4 stability is essential for accurately predicting stable buG4 sequences and assessing their functional role in transcriptional regulation.

Altered G4 stability has been reported depending on the cellular conditions associated with certain types of cancer, neurological disorders and other diseases [5, 23]. Moreover, the stability and function of nucleic acids are influenced significantly by the crowded surrounding environment within a living cell [24]. Although the intracellular environment is significantly different from that of homogeneous solutions used for experiments *in vitro*, it is possible to mimic its physical and chemical properties by using certain cosolutes [25]. Molecular crowding conditions have been demonstrated to stabilize G4 structures, while destabilizing canonical duplexes [24, 26]. Our group demonstrated recently that 40 wt% poly (ethylene glycol) with average molecular weight of 200 (PEG 200) can mimic the molecular environment of cellular organelles and that the efficiency of a guide RNA sequence during R-loop formation can be predicted under these crowding condition [24, 26]. It has also been observed that the stability of many different noncanonical structures such as G4 and i-motif are affected greatly by the crowding environment created with PEG 200 [27–30]. Thus, the thermodynamic stability and function of buG4s must be analyzed under molecular crowding condition in order to obtain quantitative values under intracellular condition. Thermodynamic analysis is a fundamental method for predicting the functions of nucleic acid structures [24]. Therefore, quantitative data obtained from thermodynamic studies are required to predict accurately the gene regulatory functions of these novel G4s based on their stability with bulge type, position, and size.

In this study, we quantified systematically the thermodynamic stability and transcription regulatory function of parallel-stranded buG4 sequences with various bulge types, positions and sizes within the G-tracts. Under molecular crowding conditions with PEG 200, both perfect G4 (pG4) and buG4s were stabilized compared to noncrowding conditions without PEG 200, thereby providing a suitable intracellular environment to study stable buG4s. Our detailed thermodynamic analysis and *in vitro* transcription assay demonstrated that the stability and transcription regulatory function of buG4s vary depending on the position and size of the bulge. With this understanding, we developed a predictive model to estimate the efficacy of buG4s in transcriptional regulation. This model can successfully determine whether a sequence is capable of arresting transcriptional regulation based on its thermodynamic stability. This study offers a comprehensive understanding of the destabilization caused due the bulges in a G4 structure, providing valuable insights for bioinformatics searches aiming for more accurate predictions of imperfect G4s, especially those with larger bulge sizes. Predicting the efficacy of buG4s in transcriptional regulation is promising for future therapeutics by specifically targeting these imperfect G4 structures.

Materials and methods

Oligonucleotides and materials

High-performance liquid chromatography (HPLC)-grade oligonucleotides were purchased from Japan Bio Service and used without further purification. To measure the single strand concentration of the sequences, the samples were heated up to 95°C, and the absorbance of 260 nm was recorded using Shimadzu 1900 spectrophotometer (Shimadzu, Kyoto, Japan) connected to a thermoprogrammer. All the solutions were prepared using 10 mM Tris–HCl buffer of pH 7.4 at 37°C containing KCl (concentration of KCl is specified in the respective experimental sections). Polyethylene glycol, with an average molecular weight of 200 (PEG 200) was purchased from Wako Pure Chemical Industries (Osaka, Japan) and used without further purification. T7 RNA polymerase, Ribonucleoside Triphosphate (rNTP), Dithiothreitol (DTT), and RNase inhibitor were purchased from Takara Bio (Shiga, Japan).

Circular dichroism

Circular dichroism (CD) spectra were measured using JASCO 1500 instrument coupled with temperature controller. DNA samples were prepared in 10 mM Tris–HCl buffer of pH 7.4 at 37°C containing KCl in the absence and presence of PEG 200. Before measurement 5 μM DNA solutions were heated up to 95°C and slowly cooled to 4°C at a rate of 1°C·min^{-1} and kept it for overnight at 4°C. The CD spectra were subtracted from the baseline, and each spectrum was the average of three consecutive measurements.

Native gel electrophoresis

Native gel electrophoresis was used to confirm the presence of the monomeric form of G4. DNA oligonucleotides (5 μM) were dissolved in a buffer containing 10 mM Tris–HCl (pH 7.4 at 37°C), 10 mM KCl and 10 wt% PEG 200 were heated up to 95°C and cooled to 4°C at a rate of 1.0°C·min^{-1}. The samples were loaded onto 10% native polyacrylamide gels in Tris–borate–EDTA (TBE) buffer and run for 70 min at 10 V·cm^{-1} at 37°C. The gels were stained with SYBR Gold (Life Technologies, now part of Thermo Fisher Scientific) and visualized using Fluoroimager (Typhoon FLA 9500).

Thermodynamic analysis

To understand the thermodynamic stability of buG4s, we performed a temperature dependent UV-melting study using a Shimadzu 1900 spectrophotometer (Shimadzu, Kyoto, Japan) connected to a thermoprogrammer. The solutions used were prepared in a buffer containing 10 mM Tris–HCl (pH 7.4 at 37°C) containing KCl and PEG 200 (exact concentrations are mentioned in the Results and Discussion section). Before measurement, the solutions were heated up to 95°C and slowly cooled down to 1°C at a rate of 1°C·min^{-1}. For melting, the sample solutions were incubated at 1°C for 10 min and heated up to 95°C at a rate of 1°C·min^{-1}. UV-melting curves were obtained by monitoring the absorbance at 295 nm as a function of temperature ranging 1–95°C. Melting temperatures (T_ms) were determined from the middle point of the melting curves [31]. The free energy change (ΔG^o), enthalpy change (ΔH^o), and entropy change (ΔS^o) for the formation of buG4s were obtained by fitting the melting curves using van't Hoff analysis [31]. The upturn in UV melting curves at high temperatures is attributed to the increased absorbance caused by the exposure of unfolded nucleobases to the solvent. Notably, this effect is more pronounced for longer bulges. To ensure accurate thermodynamic parameter calculations, the significant upturn observed for longer bulge sequences (e.g. between 80 and 95°C) was excluded during the smooth fitting of the melting curves. The method of analysis relies on the assumption that there are two states (folded and unfolded) for the buG4 unfolding process because this simplifies the analysis by allowing the system to be treated as a two-state equilibrium.

Transcription assay

The transcription reactions were carried out using the DNA template sequences. A total 20 μl for each solution was maintained for the transcription reaction. In the solution, 1.5 μM DNA template was mixed with rNTP, DTT, and RNase inhibitor. The solution buffer of pH 7.2 at 37°C contained 100 mM KCl, 40 mM Tris–HCl, and 8 mM $MgCl_2$. T7 RNA polymerase of 0.3 μM concentration was added into the solution at 37°C. The solutions were incubated for 120 min at 37°C and finally the reaction was quenched by adding DNase I. After incubation for 20 min, the reaction was stopped by adding 100 μl of a solution containing 80% (v/v) formamide and 10 mM Na_2EDTA. The reaction product was loaded onto a denaturing gel containing 10% polyacrylamide and 7 M urea and run at 70°C for 45 min at 250 V in TBE buffer. After electrophoresis, the gels were stained with SYBR Gold (Life Technologies, now part of Thermo Fisher Scientific), and the gel images were captured using a Fluoroimager (Typhoon FLA 9500). The intensity of the gel bands was analyzed using Imagequant TL software. The transcription arrested efficiency was calculated as the ratio of the intensity of arrested transcript band to the intensity of the total transcript bands.

Results and discussion

Sequence design and experimental solution conditions

The buG4 sequences of various bulge type, bulge position, and bulge size were designed by incorporating bulge nucleotides in between two Guanine (G) bases within a G-tract (indicated by dots), following the pG4 sequence, 5′-TTGGGTGGGTGGGTGGGT-3′ (Fig. 1A) [14]. Notably, the (GGGT)4 core of the pG4 sequence was derived from the HIV integrase aptamer T30695, which has been widely used in studies by various groups [32–34]. There are eight possible positions in which a bulge could be incorporated, as indicated by the dots in the parallel form of the pG4 structure (Fig. 1B). For example, a parallel buG4 structure with a bulge at position 1 is shown in Fig. 1C (structure determined in a previous study) [14]. We incorporated adenine (A), cytosine (C), and thymine (T) as bulges to understand the effect of bulge type on the G4 stability. G bulges were excluded to prevent the extension of the G-tract with consecutive G bases, which could lead to the formation of unexpected structures. Specifically, only T bulges were selected to investigate the effect of bulge size because sequences with A bulge were found to be significantly more unstable than those with T or C bulges [14]. We avoided continuous C bulges because they may form other noncanonical structures with G-C base pairing [15]. The naming convention for the buG4 sequences used in this study were as follows: bulge type (A, C, or T), bulge size (n), and bulge position (p). For example, bT_n-1 (TTGT$_n$GGTGGGTGGGTGGGT) denotes a bulge type T of size n in the suffix (1–7 nt) placed at position 1.

Previous studies showed that 100–150 mM KCl with 20–40 wt% PEG 200 mimics physiological conditions [24–26, 35–38]; however, these conditions hindered thermodynamic analysis due to the high thermal stability of pG4 with short loops (Supplementary Fig. S1A). To resolve this, melting curves were measured using 1 mM KCl and varying PEG 200 concentrations (10–40 wt%), with 10 wt% PEG 200 yielding clear curves for pG4 and single bulge buG4s under noncrowding and crowding conditions (Supplementary Fig. S1B and C). Low KCl concentrations are known to cause hysteresis in G4 structures with longer loops, making them unsuitable for accurate measurements for thermodynamic parameters [39]. However, we have confirmed experimentally that the model sequences used in this study do not show significant hysteresis in the heating and cooling curves. Therefore, we chose buffer solutions containing 10 mM KCl without PEG 200 (noncrowding) and 10 mM KCl with 10 wt% PEG 200 (crowding) to measure the thermodynamic parameters of buG4s with longer bulges (≥2 nt bulge size), which allowed the measurement of all sequences under identical solution condition, enabling an effective comparison of buG4 stability across different bulge sizes. We used 10 mM KCl without and with 10 wt% PEG 200 to study buG4s with ≥2 nt bulges, ensuring consistent conditions for all sequences. A 10 mM Tris–HCl buffer (pH 7.4 at 37°C) was used to ensure the enzyme worked efficiently during transcription [5, 22], while keeping buG4 stability and thermodynamic parameters

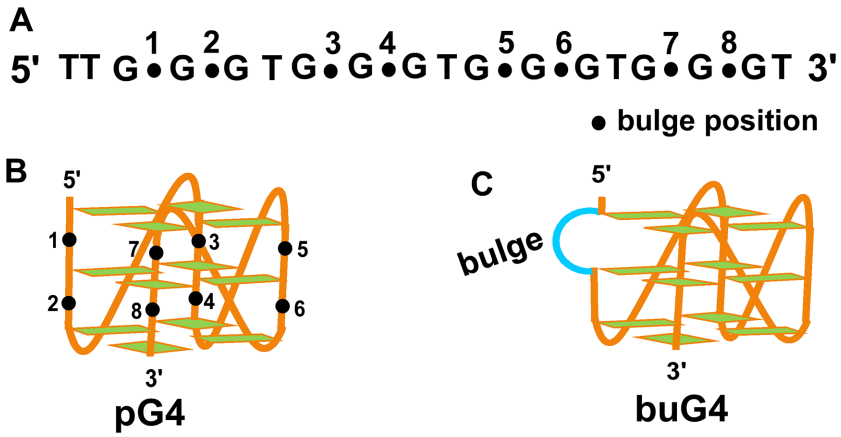


Figure 1. (**A**) Schematic representation of the G-tracts in a G4 (pG4) sequence with possible bulge positions. The bulge positions are marked with dots and are numbered from 1 to 8. (**B**) The model structure of a pG4 where the bulge positions are assigned as dots. (**C**) Example of a buG4 structure with a bulge at position 1. The bulge region is represented as a curved line.

unaffected by temperature-related pH changes of the buffer (Supplementary Fig. S1D).

Quantitative analysis of the effect of single bulges at the different bulge positions on G4 stability

The buG4s are structurally less stable compared to regular G4s [14]. In this study, we quantified the impact of a single bulge on the stability of the G4 structure by incorporating a single nucleotide at eight bulge positions in pG4 (Fig. 1A and Table 1). Initially, the structural characteristics of single-bulge buG4 sequences were analyzed. The CD spectra revealed a distinct positive peak at 265 nm and a negative peak at 245 nm under noncrowding condition, containing 10 mM Tris–HCl buffer (pH 7.4 at 37°C) with 1 mM KCl, indicative of a parallel topology regardless of bulge type or position (Supplementary Fig. S2A). In crowding condition containing 10 mM Tris–HCl buffer (pH 7.4 at 37°C) with 1 mM KCl and 10 wt% PEG 200, the CD spectra retained the same peak positions (265 nm positive and 245 nm negative peaks) (Supplementary Fig. S2B), suggesting that the crowding condition did not alter the topology of these sequences. Previous research has shown that G-rich sequences with short loops, comprising one or two nucleotides, can adopt a parallel topology and often form intermolecular structures [40]. For most of the sequences presented in this study, a concentration-independent melting temperature, determined by UV-melting (strand concentration range 3–30 μM), confirmed that the folding process was intramolecular (Supplementary Fig. S3A and B). Additionally, native gel electrophoresis revealed a single band for each sequence, supporting the formation of intramolecular structures in the single-bulge buG4s under noncrowding and crowding conditions (Supplementary Fig. S4A–F). The presence of an intramolecular parallel topology in buG4 structures is consistent with a previous study [14].

Because all buG4 sequences formed similar intramolecular parallel structures, we attempted to quantify the effect of a single bulge on the stability of pG4 using thermodynamic analysis to avoid any topology effect on stability. The UV-melting curves of buG4s were monitored at 295 nm which is the characteristic peak position for G4 formation [31]. The single transition of the UV melting curves obtained under both noncrowding and crowding conditions indicated that the melting of buG4s proceeded from a folded state to an unfolded state in a single step (Fig. 2A and B) [31]. First, we investigated the effect of different bulge type (A, C, and T) at position 1 by comparing the melting temperatures (T_ms) and free energy changes at 37°C (ΔG°_{37}) for buG4 and pG4 formation (Fig. 2A and B, and Table 1). The T_m value for pG4 was 72.5°C under noncrowding condition, while the T_m values for bA_1-1, bC_1-1, and bT_1-1 decreased to 53.6°C, 58.3°C, and 59.1°C, respectively (Table 1). The ΔG°_{37} value for buG4 formation was also increased when compare with ΔG°_{37} value for pG4 formation, indicating destabilization of the structure. For example, ΔG°_{37} value for pG4 was −6.1 kcal mol^{-1} which was increased to −2.5, −3.8, and −3.8 kcal mol^{-1} for bA_1-1, bC_1-1, and bT_1-1, respectively, under noncrowding condition (Table 1). This result indicated that the pG4 structure was destabilized by the incorporation of a bulge, regardless of the bulge type, following the order of destabilization: A > C ≥ T. This order of destabilization in a G4 structure is opposite to the destabilization observed in duplex structures, which follows the order: T > C > A [41].

To better understand the underlying mechanism, we analyzed the enthalpy change (ΔH°) and entropy change (ΔS°) associated with the free energy change at 37°C (ΔG°_{37}) of the buG4 formation (Table 1). We found that both ΔH° and $T\Delta S^\circ$ values were increased for the formation of buG4 compared to pG4 (Table 1). For instance, under the noncrowding condition, the ΔH° and $T\Delta S^\circ$ for pG4 were −57.3 kcal·mol^{-1} and −51.2 kcal·mol^{-1}, respectively; however, it increased to −52.5 kcal·mol^{-1} and −48.7 kcal·mol^{-1}, respectively, for bT_1-1. Notably, the increase in ΔH° is more pronounced compared to the increase in $T\Delta S^\circ$, suggesting that the destabilization of buG4s is mainly governed by the unfavorable enthalpic contribution during buG4 formation. Similar destabilization of pG4 has been observed in the presence of an abasic site, which is also attributed to unfavorable enthalpic contribution and is dependent on the position of the abasic site [28]. The rationale behind this result can be attributed to the weakening of the base stacking interactions and hydrogen bonding of the G-quartets owing to the incorporation of bulge. However, increase in ΔG°_{37} was most pronounced when the bulge nucleotide was A, while C and T caused similar values of ΔG°_{37} (Table 1). This discrepancy in ΔG°_{37} for different bulge types is likely due to the larger size of the A bulge

Table 1. The melting temperature and thermodynamic parameters for the formation of pG4 and single bulge buG4s under noncrowding and crowding conditions

		Noncrowding condition[a] (without PEG 200)				Crowding condition[b] (with 10 wt% PEG 200)			
Name	Sequences (5′- 3′)	T_m (°C)[c]	$\Delta H°$ (kcal·mol^{-1})	$T\Delta S°$ (kcal·mol^{-1})	$\Delta G°_{37}$ (kcal·mol^{-1})	T_m (°C)[c]	$\Delta H°$ (kcal·mol^{-1})	$T\Delta S°$ (kcal·mol^{-1})	$\Delta G°_{37}$ (kcal·mol^{-1})
pG4	TTGGGTGGGTGGGTGGGT	72.5 ± 0.1	−57.3 ± 0.2	−51.2 ± 0.2	−6.1 ± 0.3	81.2 ± 0.2	−61.1 ± 0.1	−53.1 ± 0.2	−8.0 ± 0.5
bA$_1$-1	TTG**A**GGTGGGTGGGTGGGT	53.6 ± 0.1	−45.4 ± 0.1	−42.9 ± 0.1	−2.5 ± 0.1	63.4 ± 0.1	−52.9 ± 0.1	−48.6 ± 0.1	−4.3 ± 0.2
bC$_1$-1	TTG**C**GGTGGGTGGGTGGGT	58.3 ± 0.5	−52.5 ± 0.1	−48.7 ± 0.1	−3.8 ± 0.1	66.1 ± 0.2	−56.1 ± 0.1	−51.1 ± 0.1	−5.0 ± 0.2
bT$_1$-1	TTG**T**GGTGGGTGGGTGGGT	59.1 ± 1.2	−52.5 ± 0.1	−48.7 ± 0.2	−3.8 ± 0.1	67.0 ± 0.2	−55.0 ± 0.1	−49.9 ± 0.3	−5.1 ± 0.3
bT$_1$-2	TTGG**T**GTGGGTGGGTGGGT	47.8 ± 0.6	−44.8 ± 0.1	−43.1 ± 0.3	−1.7 ± 0.1	56.7 ± 0.3	−49.7 ± 0.1	−46.6 ± 0.2	−3.1 ± 0.3
bT$_1$-3	TTGGGTG**T**GGTGGGTGGGT	50.9 ± 1.1	−52.2 ± 0.1	−49.9 ± 0.2	−2.3 ± 0.2	60.1 ± 0.2	−51.1 ± 0.1	−47.3 ± 0.1	−3.8 ± 0.1
bT$_1$-4	TTGGGTGG**T**GTGGGTGGGT	49.3 ± 0.1	−55.5 ± 0.2	−53.3 ± 0.2	−2.2 ± 0.1	56.7 ± 0.8	−61.3 ± 0.1	−57.5 ± 0.2	−3.8 ± 0.2
bT$_1$-5	TTGGGTGGGTG**T**GGTGGGT	50.5 ± 0.3	−46.4 ± 0.1	−44.5 ± 0.4	−1.9 ± 0.3	59.9 ± 0.6	−55.3 ± 0.1	−51.2 ± 0.4	−4.0 ± 0.1
bT$_1$-6	TTGGGTGGGTGG**T**GTGGGT	48.4 ± 0.1	−49.3 ± 0.2	−47.5 ± 0.1	−1.8 ± 0.2	56.4 ± 0.5	−57.5 ± 0.1	−54.1 ± 0.2	−3.4 ± 0.2
bT$_1$-7	TTGGGTGGGTGGGTG**T**GGT	48.8 ± 1.1	−54.2 ± 0.1	−52.2 ± 0.2	−2.0 ± 0.5	58.4 ± 0.2	−55.5 ± 0.1	−51.7 ± 0.2	−3.8 ± 0.3
bT$_1$-8	TTGGGTGGGTGGGTGG**T**GT	48.3 ± 1.0	−53.9 ± 0.1	−51.9 ± 0.2	−2.0 ± 0.4	56.9 ± 0.8	−53.6 ± 0.2	−50.2 ± 0.3	−3.4 ± 0.2

[a]Experiments were performed in 10 mM Tris–HCl buffer containing 1 mM KCl without PEG 200 (pH 7.4 at 37°C).
[b]Experiments were performed in 10 mM Tris–HCl buffer containing 1 mM KCl with 10 wt% PEG 200 (pH 7.4 at 37°C).
[c]DNA concentration used for the measurement was 5 μM.

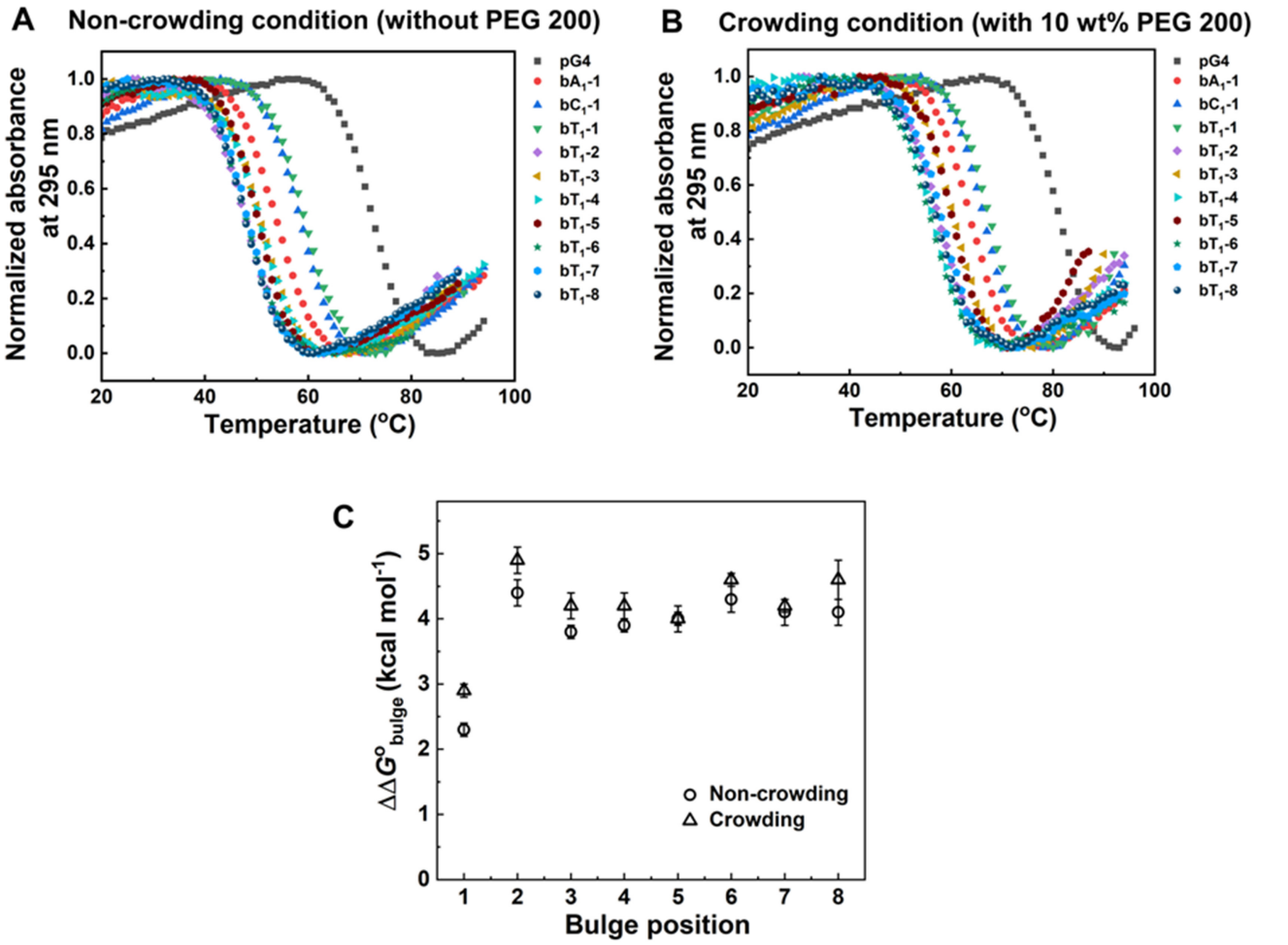


Figure 2. UV-melting curves of pG4 (5 μM) and single bulge containing buG4s (5 μM) in (**A**) noncrowding conditions containing 1 mM KCl, 10 mM Tris–HCl (pH 7.4 at 37°C) without PEG 200 and (**B**) crowding conditions containing 1 mM KCl, 10 mM Tris–HCl (pH 7.4 at 37°C) with 10 wt% PEG 200, monitored at 295 nm. (**C**) The plot of $\Delta\Delta G°_{bulge}$ with different bulge positions under noncrowding and crowding conditions. Here, $\Delta\Delta G°_{bulge} = \Delta G°_{37}$ (buG4) − $\Delta G°_{37}$ (pG4), where $\Delta G°_{37}$ (buG4) and $\Delta G°_{37}$ (pG4) are the free energy change of the buG4s and pG4 at 37°C, respectively.

compared to the C or T bulges, which has the larger steric effect on G4 folding. A similar trend was observed under crowding conditions, indicating that 10 wt% PEG 200 stabilizes the otherwise unstable buG4s, facilitating their study without altering the overall trend significantly (Table 1). It highlights a similar destabilizing effect observed when purine bases were incorporated into G4 loops compared to pyrimidine bases [42, 43]. With this in mind, to further quantify the destabilization caused by the bulge, we selected a single T bulge at different positions within the pG4 sequence in order to better understand how bulge position affects stability under cell mimetic environment.

Incorporating a single T bulge at various bulge positions led to a decrease in the T_m values and an increase in the ΔG°_{37} values associated with buG4 formation under both noncrowding and crowding conditions, as shown in Table 1. Although our study did not use physiological conditions, we found that the ΔG°_{37}value of buG4 decreases linearly increasing with both KCl and PEG 200 concentrations (Supplementary Figs S5–7). This result suggests that comparing ΔG°_{37} values of buG4s and pG4, even at lower KCl or PEG 200 concentrations, reflects the trends observed under cell-mimicking crowding conditions. Thus, we quantified the effect of a single bulge on the thermodynamic stability of the pG4 structure by comparing the ΔG°_{37} of buG4s and pG4 [$\Delta\Delta G^\circ_{bulge} = \Delta G^\circ_{37}$ (buG4) − ΔG°_{37} (pG4), where ΔG°_{37} (buG4) and ΔG°_{37} (pG4) represent the free energy change of buG4 and pG4 formations at 37°C, respectively] (ΔG°_{37} values listed in Table 1). Positive values for $\Delta\Delta G^\circ_{bulge}$ indicate that, regardless of the bulge position, the presence of a bulge leads to destabilization of the buG4 structure when compared to pG4. However, the plot showed $\Delta\Delta G^\circ_{bulge}$ is slightly lower (2.3 kcal·mol^{-1}) for bulge position 1, compared to the other bulge positions showing average of 4.0 kcal·mol^{-1} in the noncrowding condition (Fig. 2C). Similarly, in the crowding condition $\Delta\Delta G^\circ_{bulge}$ is 2.9 kcal·mol^{-1} for bulge position 1, compared to the other bulge positions showing average of 4.4 kcal·mol^{-1}. This result suggests that incorporating a single bulge at position 1 has a less destabilizing effect on the pG4 structure compared to other bulge positions, which is consistent under both noncrowding and crowding conditions. One of the major factors that stabilizes the G4 is the base stacking interactions, in addition to hydrogen bonds, electrostatic interactions and the hydration shell, similar to the forces that stabilize duplex DNA [44]. The presence of a bulge in buG4 structure increases the distance between the quartets, weakening the quartets compared to a similar parallel G4 structure without a bulge (Supplementary Fig. S8A and B). In a typical G4 structure, the O6 atoms form a square in each quartet, creating a bipyramidal antiprism in the quadruplex with an interquartet distance of 3.4 Å which increased to 4 Å (Supplementary Fig. S8A). This suggests that bulges at other positions likely further increase the interquartet distance, weakening the stacking interactions and causing variations in buG4 stability based on the bulge position. Having quantified the bulge position dependency for a single bulge, it is now important to understand how changing the bulge size at different positions influences the stability of buG4, which is discussed in the next section.

Quantitative analysis of the effect of bulge size at different bulge positions on the stability of buG4 structure

To explore how varying bulge sizes affect the stability of intramolecular buG4 structures, their thermodynamic properties were analyzed using UV-melting curves for all the buG4 sequences as they all can form intramolecular parallel structure (Supplementary Fig. S4A–F and Supplementary Fig. S9A–L). Moreover, unlike G4 sequences with longer loops (>6 nt), which often exhibit hysteresis in their melting curves [45], the buG4 sequences showed no significant hysteresis, even with bulge sizes up to 7 nt (Supplementary Fig. S10A and B). This observation further supports the formation of intramolecular buG4 structures, enabling reliable thermodynamic analysis of their melting curves. The typical example UV-melting curves of bT_1-1 to bT_7-1 sequences (Fig. 3A and B) revealed a decrease in melting temperature with an increase in bulge size under both noncrowding and crowding conditions, consistent with a previous report [14]. The ΔG°_{37} values for buG4s with longer bulge increases drastically compared to the single bulge at every bulge position with bulge size varying from 2 to 7 nt, indicating that the longer bulges affect the thermodynamic stability of buG4 structures more effectively compared to single bulge. For example, for bulge position 1, ΔG°_{37} value for bT_1-1 was −6.6 kcal·mol^{-1} which increased to −0.2 kcal·mol^{-1} for bT_5-1 under the noncrowding condition (Table 2). Moreover, the ΔG°_{37} values for bT_6-1 and bT_7-1, were showing positive values of 0.2 and 0.5 kcal mol^{-1}, respectively, under noncrowding condition, indicating the structures are thermodynamically not favorable at 37°C. The associated ΔH° and $T\Delta S^\circ$ for bT_1-1 were −61.2 and −54.6 kcal·mol^{-1} which increased to −36.8 and −37.2 kcal·mol^{-1} for bT_7-1, respectively, suggesting an unfavorable enthalpy contribution which leads to destabilization of structure due to increase in bulge size (Supplementary Table S1). In the presence of 10 wt% PEG 200, the ΔG°_{37} values for bT_6-1 and bT_7-1 were showing negative values of −1.5 and −1.6 kcal·mol^{-1}, respectively, indicating that PEG 200 provides a favorable condition for buG4 formation with long bulges. The PEG 200 mimics effectively crowding effects and stabilizes buG4s; however, it cannot replicate fully the complexity of the cellular environment. The cellular milieu includes macromolecules, amino acids, and divalent cations that influence nucleic acid folding and thermodynamics through electrostatic interactions, hydrogen bonding, and stacking interactions, factors absent in PEG 200.

To understand the effect of the bulge size on the thermodynamic stability of buG4s at every bulge position, we plotted $\Delta\Delta G^\circ_{bulge\ size}$ [= ΔG°_{37} (longer bulge) − ΔG°_{37} (single bulge), where ΔG°_{37} (longer bulge) and ΔG°_{37} (single bulge) are the free energy change of the buG4s with longer bulge (2–7 nt) and buG4s with single bulge (1 nt) at 37°C, respectively] against bulge size in the noncrowding and crowding conditions (Fig. 3C and D). The plot revealed that the $\Delta\Delta G^\circ_{bulge\ size}$ values for buG4 stability with increasing bulge size do not follow a consistent trend across the eight bulge positions. At position 1 on the 5′ end, $\Delta\Delta G^\circ_{bulge\ size}$ gradually increased with bulge size and saturated after 5 nt under the noncrowding condition, suggesting that increases in bulge size beyond 5 nt did not significantly affect buG4 stability (Fig. 3C). This saturation effect is more pronounced under the crowding condition, highlighting the stronger stabilizing influence of molecular crowding on longer bulges (Fig. 3D). At bulge positions 3, 5, and 7, a similar trend of a gradual increase in $\Delta\Delta G^\circ_{bulge\ size}$ with increasing bulge size was observed. However, in the noncrowding conditions, longer bulge sizes could not be fully assessed due to the inability of the sequences to fold into buG4 structure (Fig. 3C). In contrast, bulge size variations at positions 2, 4, 6, and 8 showed minimal destabilization up to 3 nt, with a more pronounced increase in $\Delta\Delta G^\circ_{bulge\ size}$ occurring only beyond 3 nt. This suggests that bulge size has a less significant effect on stability at these positions compared to positions 1, 3, 5, and 7. The variability in $\Delta\Delta G^\circ_{bulge\ size}$ with longer bulges (≥2 nt) across different positions indicates a strong bulge position dependency on buG4 stability. Previ-

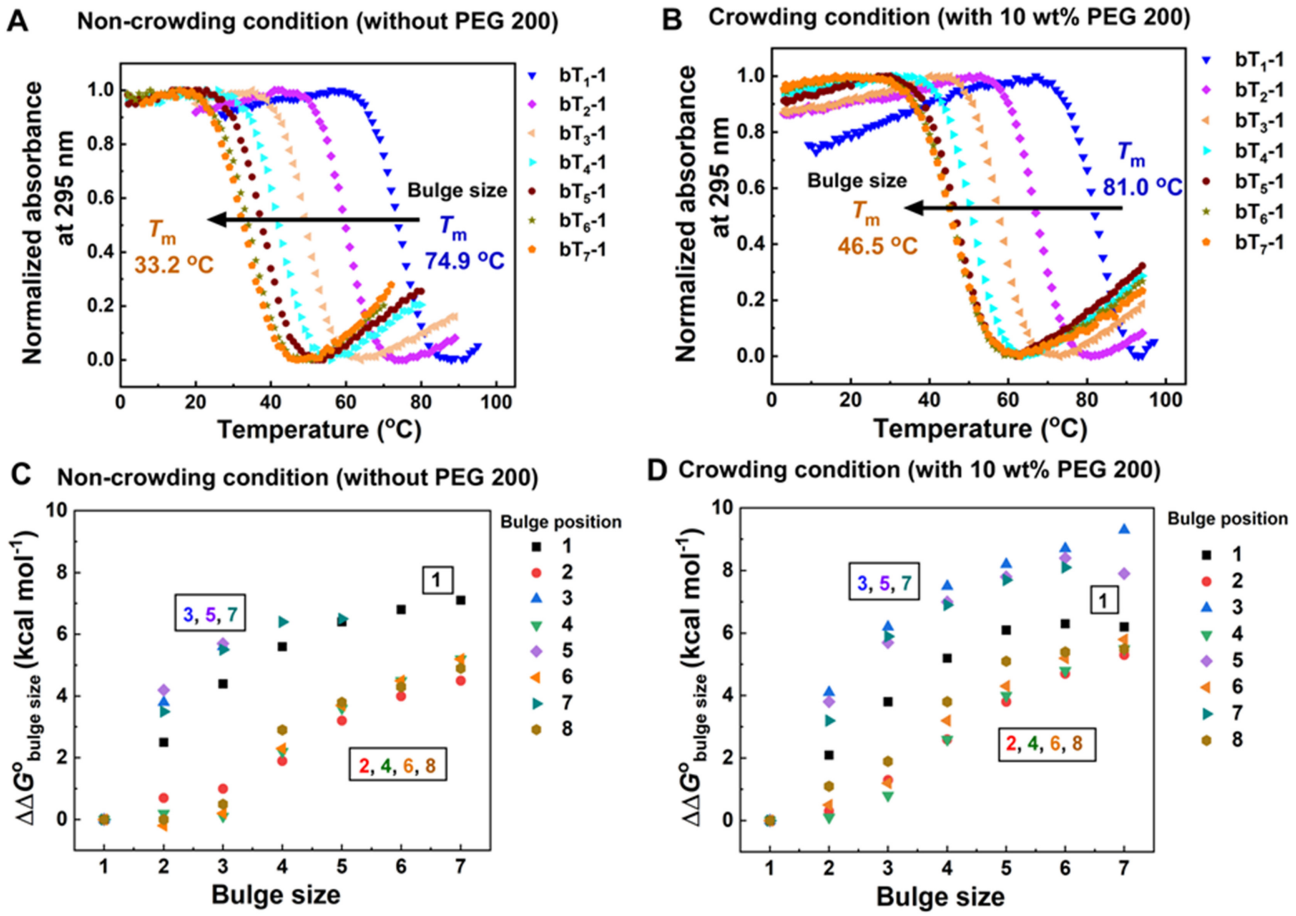


Figure 3. UV-melting curves of buG4s (5 μM) with varied bulge size at position 1 in (**A**) noncrowding condition containing 10 mM KCl, 10 mM Tris–HCl buffer (pH 7.4 at 37°C) without PEG 200 and (**B**) crowding condition containing 10 mM KCl, 10 mM Tris–HCl (pH 7.4 at 37°C) with 10 wt% PEG 200, monitored at 295 nm. The plot of $\Delta\Delta G^\circ_{bulge\ size}$ with increasing bulge size at various bulge positions under (**C**) noncrowding and (**D**) crowding conditions (with 10 wt% PEG200). Here, $\Delta\Delta G^\circ_{bulge\ size} = \Delta G^\circ_{37}$ (longer bulge) − ΔG°_{37} (single bulge), ΔG°_{37} (longer bulge), and ΔG°_{37} (single bulge) are the free energy change of the buG4s with longer bulges (2–7 nt) and buG4s with single bulge at 37°C, respectively.

ous studies have shown that an increase in loop size similarly affects the T_m values of G4 structures at both the 5′ and 3′ ends [45, 46]. However, our findings demonstrated that bulge size influences stability of G4 structure differently from loop size, with a notable destabilization on the bulge positions (1, 3, 5, and 7) compared to the bulge positions (2, 4, 6, and 8). Thus, we hypothesized that the influence of bulge position and size on both their structure and thermodynamic stability could, in turn, affect their gene regulatory function, such as transcription, as discussed in the following section.

Impact of bulge position on transcriptional regulation induced by buG4 with single bulge

To investigate the effect of bulges on the transcriptional regulatory functions of buG4, we performed an *in vitro* transcription assay under both noncrowding and crowding conditions (Fig. 4 and Supplementary Fig. S11). The template DNAs were designed by placing pG4 sequence (forming perfect G4 structures) and buG4 sequences (containing a bulge), at the downstream of the TSS (Fig. 4A and sequences are listed in Supplementary Table S2). The template design is illustrated in Fig. 4A, which highlights the bulge positions as dots. At first, the template strand was transcribed using a linear sequence (incapable of forming G4 structures) as an internal control (Supplementary Table S2) to normalize T7 RNA polymerase activity for different measurements. For linear sequence T7 RNA polymerase runs off until the end of the template DNA, producing a 55 nt full-transcript (Fig. 4B). A schematic illustration for the full transcript production is shown in Supplementary Fig. S12A. On the other hand, if a quadruplex-forming sequence was present, it could block the T7 RNA polymerase movement, resulting in shorter 15-nt RNA transcript (Fig. 4B). Previous studies demonstrated that transcription of linear sequences by T7 RNA polymerase reaches saturation within 120 min and 37°C [5, 22]. Thus, we transcribed the DNA template sequences using T7 RNA polymerase for 120 min in a buffer containing 100 mM KCl, 40 mM Tris–HCl (pH 7.2), and 8 mM $MgCl_2$, both without and with 10 wt% PEG 200. The denaturing gel electrophoresis results showed that the 78-nt template DNA with a linear sequence was transcribed into a 55-nt full-length transcript (Fig. 4B and Supplementary Fig. S11). On the other hand, transcription of the pG4-containing template strand (L-pG4) was completely stopped, producing only a short 15 nt RNA product under both noncrowding and crowding conditions (Fig. 4B and Supplementary Fig. S11). This suggests that T7 RNA polymerase encounters a fully folded G4 structure involving four G-tracts, halting transcription immediately before the G-rich sequence (Fig 4A and Supplementary Fig. S12B). We con-

Table 2. The thermodynamic parameters for buG4 formation with different bulge size under noncrowding and crowding conditions

		Noncrowding[a] (without PEG 200)		Crowding[b] (10 wt% PEG 200)	
Name	Sequences (5′- 3′)	T_m (°C)[c]	$\Delta G°_{37}$ (kcal·mol^{-1})	T_m (°C)[c]	$\Delta G°_{37}$ (kcal·mol^{-1})
bT$_1$-1	TTGTGGTGGGTGGGTGGGT	74.9 ± 0.5	−6.6 ± 0.6	81.0 ± 0.9	−7.8 ± 0.3
bT$_2$-1	TTGTTGGTGGGTGGGTGGGT	59.8 ± 0.9	−4.1 ± 0.3	68.4 ± 0.2	−5.7 ± 0.4
bT$_3$-1	TTGTTTGGTGGGTGGGTGGGT	49.4 ± 0.6	−2.2 ± 0.1	58.9 ± 0.5	−4.0 ± 0.2
bT$_4$-1	TTGTTTTGGTGGGTGGGTGGGT	42.0 ± 1.0	−1.0 ± 0.4	51.6 ± 0.2	−2.6 ± 0.2
bT$_5$-1	TTGTTTTTGGTGGGTGGGTGGGT	37.7 ± 0.3	−0.2 ± 0.3	48.0 ± 0.4	−1.7 ± 0.1
bT$_6$-1	TTGTTTTTTGGTGGGTGGGTGGGT	34.2 ± 0.2	0.2 ± 0.2	46.4 ± 0.8	−1.5 ± 0.4
bT$_7$-1	TTGTTTTTTTGGTGGGTGGGTGGGT	33.2 ± 0.2	0.5 ± 0.1	46.5 ± 0.4	−1.6 ± 0.3
bT$_1$-2	TTGGTGTGGGTGGGTGGGT	64.2 ± 0.2	−4.3 ± 0.1	71.9 ± 0.5	−6.1 ± 0.2
bT$_2$-2	TTGGTTGTGGGTGGGTGGGT	64.2 ± 0.2	−3.6 ± 0.5	69.2 ± 0.1	−5.8 ± 0.4
bT$_3$-2	TTGGTTTGTGGGTGGGTGGGT	58.4 ± 0.1	−3.3 ± 0.2	63.3 ± 1.2	−4.8 ± 0.3
bT$_4$-2	TTGGTTTTGTGGGTGGGTGGGT	50.2 ± 0.2	−2.4 ± 0.1	55.8 ± 0.1	−3.5 ± 0.3
bT$_5$-2	TTGGTTTTTGTGGGTGGGTGGGT	44.4 ± 2.0	−1.1 ± 0.1	49.4 ± 3.0	−2.3 ± 0.5
bT$_6$-2	TTGGTTTTTTGTGGGTGGGTGGGT	39.4 ± 0.6	−0.3 ± 0.2	44.4 ± 0.3	−1.4 ± 0.2
bT$_7$-2	TTGGTTTTTTTGTGGGTGGGTGGGT	35.6 ± 0.6	0.2 ± 0.1	41.1 ± 1.1	−0.8 ± 0.4
bT$_1$-3	TTGGGTGTGGTGGGTGGGT	65.6 ± 0.9	−5.3 ± 0.3	75.8 ± 1.1	−7.2 ± 0.3
bT$_2$-3	TTGGGTGTTGGTGGGTGGGT	46.1 ± 0.1	−1.5 ± 0.1	57.6 ± 0.5	−3.6 ± 0.3
bT$_3$-3	TTGGGTGTTTGGTGGGTGGGT	34.6 ± 0.6	0.3 ± 0.3	44.9 ± 0.5	−1.5 ± 0.5
bT$_4$-3	TTGGGTGTTTTGGTGGGTGGGT	24.9 ± 2.1	–[d]	37.4 ± 1.0	−0.2 ± 0.3
bT$_5$-3	TTGGGTGTTTTTGGTGGGTGGGT	19.6 ± 1.4	–[d]	32.4 ± 0.8	0.5 ± 0.2
bT$_6$-3	TTGGGTGTTTTTTGGTGGGTGGGT	14.7 ± 1.5	–[d]	28.7 ± 0.9	1.0 ± 0.2
bT$_7$-3	TTGGGTGTTTTTTTGGTGGGTGGGT	–[d]	–[d]	27.2 ± 0.5	1.6 ± 0.1
bT$_1$-4	TTGGGTGGTGTGGGTGGGT	65.3 ± 0.7	−5.3 ± 0.3	72.0 ± 0.6	−6.6 ± 0.1
bT$_2$-4	TTGGGTGGTTGTGGGTGGGT	66.5 ± 0.6	−5.1 ± 0.2	70.5 ± 0.3	−6.5 ± 0.2
bT$_3$-4	TTGGGTGGTTTGTGGGTGGGT	63.7 ± 0.8	−5.2 ± 0.2	66.8 ± 0.5	−5.8 ± 0.2
bT$_4$-4	TTGGGTGGTTTTGTGGGTGGGT	53.2 ± 0.6	−3.1 ± 0.2	56.8 ± 0.5	−4.0 ± 0.3
bT$_5$-4	TTGGGTGGTTTTTGTGGGTGGGT	45.5 ± 0.5	−1.7 ± 0.1	50.1 ± 0.5	−2.6 ± 0.1
bT$_6$-4	TTGGGTGGTTTTTTGTGGGTGGGT	40.1 ± 0.5	−0.8 ± 0.2	45.5 ± 0.3	−1.8 ± 0.2
bT$_7$-4	TTGGGTGGTTTTTTTGTGGGTGGGT	37.2 ± 0.4	−0.1 ± 0.1	42.5 ± 0.9	−1.1 ± 0.1
bT$_1$-5	TTGGGTGGGTGTGGTGGGT	64.7 ± 1.0	−5.3 ± 0.3	74.8 ± 0.5	−7.4 ± 0.3
bT$_2$-5	TTGGGTGGGTGTTGGTGGGT	45.5 ± 1.2	−1.1 ± 0.4	57.2 ± 0.1	−3.6 ± 0.1
bT$_3$-5	TTGGGTGGGTGTTTGGTGGGT	34.2 ± 1.5	0.4 ± 0.2	44.9 ± 0.2	−1.7 ± 0.5
bT$_4$-5	TTGGGTGGGTGTTTTGGTGGGT	26.2 ± 1.0	–[d]	35.6 ± 0.3	−0.4 ± 0.1
bT$_5$-5	TTGGGTGGGTGTTTTTGGTGGGT	22.0 ± 1.2	–[d]	31.9 ± 1.5	0.4 ± 0.1
bT$_6$-5	TTGGGTGGGTGTTTTTTGGTGGGT	18.6 ± 1.4	–[d]	29.3 ± 1.3	1.0 ± 0.1
bT$_7$-5	TTGGGTGGGTGTTTTTTTGGTGGGT	16.3 ± 0.4	–[d]	26.5 ± 1.5	0.5 ± 0.2
bT$_1$-6	TTGGGTGGGTGGTGTGGGT	63.2 ± 0.5	−4.7 ± 0.2	71.8 ± 0.8	−6.4 ± 0.2
bT$_2$-6	TTGGGTGGGTGGTTGTGGGT	63.3 ± 0.8	−4.9 ± 0.1	67.4 ± 0.6	−5.9 ± 0.2
bT$_3$-6	TTGGGTGGGTGGTTTGTGGGT	60.5 ± 0.6	−4.5 ± 0.1	64.0 ± 0.6	−5.2 ± 0.1
bT$_4$-6	TTGGGTGGGTGGTTTTGTGGGT	49.6 ± 0.6	−2.4 ± 0.2	54.3 ± 0.5	−3.2 ± 0.3
bT$_5$-6	TTGGGTGGGTGGTTTTTGTGGGT	41.6 ± 0.8	−1.0 ± 0.1	48.1 ± 0.6	−2.2 ± 0.1
bT$_6$-6	TTGGGTGGGTGGTTTTTTGTGGGT	36.7 ± 0.5	−0.2 ± 0.1	42.8 ± 0.3	−1.2 ± 0.1
bT$_7$-6	TTGGGTGGGTGGTTTTTTTGTGGGT	33.2 ± 0.9	0.5 ± 0.1	39.9 ± 1.1	−0.6 ± 0.1
bT$_1$-7	TTGGGTGGGTGGGTGTGGT	63.4 ± 0.5	−4.8 ± 0.5	73.8 ± 0.2	−7.0 ± 0.4
bT$_2$-7	TTGGGTGGGTGGGTGTTGGT	47.0 ± 0.1	−1.3 ± 0.5	57.2 ± 0.1	−3.8 ± 0.4
bT$_3$-7	TTGGGTGGGTGGGTGTTTGGT	31.9 ± 1.1	0.7 ± 0.2	43.6 ± 0.9	−1.1 ± 0.2
bT$_4$-7	TTGGGTGGGTGGGTGTTTTGGT	25.2 ± 1.2	1.6 ± 0.3	35.1 ± 0.6	−0.1 ± 0.2
bT$_5$-7	TTGGGTGGGTGGGTGTTTTTGGT	20.2 ± 0.3	1.7 ± 0.4	29.9 ± 0.8	0.7 ± 0.2
bT$_6$-7	TTGGGTGGGTGGGTGTTTTTTGGT	15.2 ± 1.2	–[d]	26.2 ± 0.6	1.1 ± 0.2
bT$_7$-7	TTGGGTGGGTGGGTGTTTTTTTGGT	–[d]	–[d]	24.9 ± 0.9	–[d]
bT$_1$-8	TTGGGTGGGTGGGTGGTGT	64.3 ± 0.5	−4.0 ± 0.2	73.6 ± 0.4	−6.6 ± 0.4
bT$_2$-8	TTGGGTGGGTGGGTGGTTGT	58.5 ± 1.5	−4.0 ± 0.2	65.6 ± 0.3	−5.5 ± 0.5
bT$_3$-8	TTGGGTGGGTGGGTGGTTTGT	56.3 ± 0.7	−3.5 ± 0.5	60.5 ± 0.3	−4.7 ± 0.4
bT$_4$-8	TTGGGTGGGTGGGTGGTTTTGT	47.8 ± 1.5	−1.1 ± 0.6	52.1 ± 1.2	−2.8 ± 0.4
bT$_5$-8	TTGGGTGGGTGGGTGGTTTTTGT	38.8 ± 1.2	−0.2 ± 0.2	46.2 ± 0.1	−1.5 ± 0.3
bT$_6$-8	TTGGGTGGGTGGGTGGTTTTTTGT	34.9 ± 0.9	0.3 ± 0.1	43.8 ± 0.6	−1.2 ± 0.5
bT$_7$-8	TTGGGTGGGTGGGTGGTTTTTTTGT	33.3 ± 0.3	0.9 ± 0.4	42.7 ± 0.7	−1.1 ± 0.5

[a]Experiments were performed in 10 mM Tris–HCl buffer containing 10 mM KCl (pH 7.4 at 37°C).
[b]Experiments were performed in 10 mM Tris–HCl buffer containing 10 mM KCl and 10 wt% PEG 200 (pH 7.4 at 37°C).
[c]DNA concentration used for the measurement was 5 μM. [d](-) represents the values cannot be calculated as the buG4 structures were unstable at 37°C.

ducted a systematic transcription assay using templates containing buG4 sequences with a single bulge at various positions (Fig. 4A) under both noncrowding and crowding conditions. For templates with buG4 sequence containing bulge at position 1 (L-bT$_1$-1), T7 RNA polymerase was arrested, producing a 15-nt arrested transcript-1, similar to pG4. This suggests that, despite the presence of a bulge, a fully folded buG4 structure was still formed, creating a barrier for T7 RNA polymerase movement. A similar observation was made for bulges at positions 2, 3, 4, 5, 6, and 7 (Fig. 4B and Supplementary Fig. S12C–F), suggesting that transcription arrest was due to the formation of buG4. Notably, a second transcript, transcript-2, was observed for L-bT$_1$-7 with single bulge, indicating that the T7 RNA polymerase was likely stalled at a second position within the buG4 sequence (Supplementary Fig. S12F), possibly due to the formation of a partially folded structure. An equilibrium between fully folded and partially folded buG4s was previously discussed based on bulge position which supports our result [19]. Interestingly, when the bulge was located at the 3′ end (L-bT$_1$-8), full-length transcript production increased compared to other buG4-containing templates. This suggests that a buG4 with a bulge at position 8 is less effec-

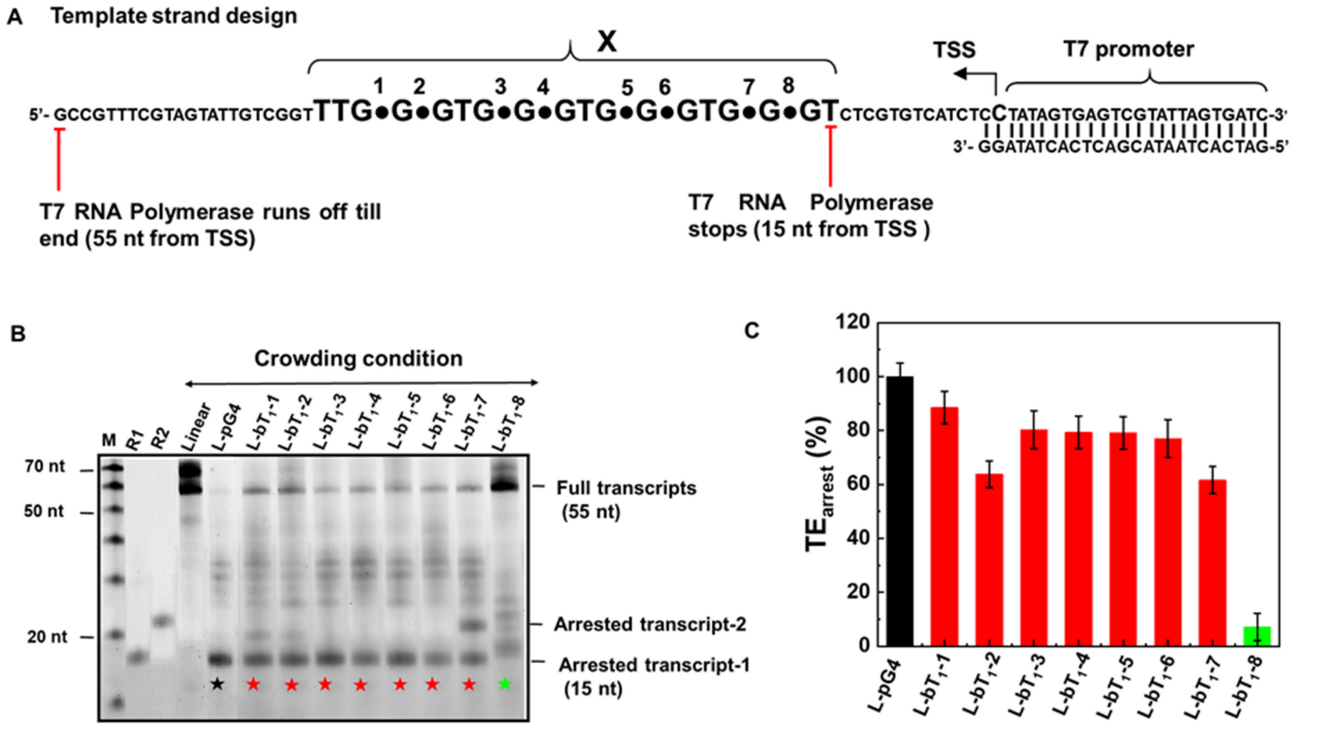


Figure 4. (**A**) Template DNA sequence where the X region contains the pG4 or buG4 sequences. The T7 promoter region where the T7 RNA polymerase will bind and the TSS where the transcription will start, are highlighted in the template sequence. The bulge positions 1, 2, 3, 4, 5, 6, 7, and 8 are represented as dots. When T7 RNA polymerase stops in front of the G-rich region a 15 nt short RNA will be produced, whereas, when it runs off till the end of the sequence a 55 nt RNA will be produced. Template sequences are shown in Supplementary Table S2. (**B**) Transcript production of the linear, pG4 and buG4 containing template sequences in the presence of 100 mM KCl and 10 wt% PEG200 as crowding condition. Denaturing gel electrophoresis was used to identify the products of transcription reactions carried out for 120 min at 37°C. The marked lanes represent the transcript formation from the template DNA containing Linear, pG4, bT_1-1, bT_1-2, bT_1-3, bT_1-4, bT_1-5, bT_1-6, bT_1-7, and bT_1-8. *M* is the size marker. R1 and R2 are the 15 nt and 20 nt RNA sequences, respectively as markers. The 15 nt arrested bands from pG4 and buG4 containing templates are marked with stars. (**C**) Transcription arrested efficiency (TE_{arrest}) for the template DNA containing pG4 and single bulge buG4s, calculated as the ratio of arrested transcript-1 intensity to total transcript intensity. TE_{arrest} values below 10% are considered as background, hence considered as no arrest.

tive at arresting transcription than buG4s with bulges at other positions. No arrested transcript-1 was produced; instead, a slightly longer arrested transcript was observed of length more than 15 nt (Supplementary Fig. S12G), indicating the presence of only a partially folded buG4 structure. Although both fully folded and partially folded intramolecular structures of buG4 have been reported for bT_1-8 sequence without the template strand [19], the bulge at position 8 in the template strand likely prevents full buG4 folding due to interactions with T7 RNA polymerase during transcription. Additionally, the intensity of the longer transcript band was weaker than that of transcript-1 observed for buG4s with bulges at other positions, indicating that the partially folded structure is less effective at halting transcription, highlighting the importance of fully folded buG4 formation for efficient transcriptional arrest under both noncrowding and crowding conditions.

Furthermore, we estimated the efficiency of buG4 sequences in arresting the transcription process by calculating the transcription-arrested efficiency (TE_{arrest}) of the buG4 sequences from the amount of arrested transcript-1 product obtained from the denaturing gel bands (Fig. 4C). The value of TE_{arrest} was determined by dividing the band intensity of arrested transcript-1 (signifying the formation of a fully folded buG4) by the sum of the intensities of the total transcript production. We considered only the arrested transcript-1 to calculate the efficiency, as it is produced due to the formation of a fully folded buG4 structure. pG4 completely arrested transcription, yielding a TE_{arrest} value of 100%. For the buG4 sequence with a bulge at position 1, the TE_{arrest} value decreased to 88.5% (Fig. 4C). When bulges were introduced at middle positions (2-6, and 7), the TE_{arrest} decreased further to ~60–70%. These findings align with the thermodynamic stability data, demonstrating that the buG4 with the bulge at position 1, being the most stable, achieves the highest TE_{arrest}. This supports previous studies suggesting that G4 structure stability plays a key role in determining the efficiency of a sequence to effectively arrest the transcription process [5, 22]. However, an exception was observed for the L-bT_1-8 buG4 sequence, which contains a bulge at position 8. Despite the comparable stability of bT_1-8 with other buG4 structures, its TE_{arrest} dropped significantly to 7% which signifies no arrest (Fig. 4C). This discrepancy likely results from the inability of the L-bT_1-8 sequence to form a fully folded buG4 during transcription. This finding highlights the importance of structural integrity in buG4-mediated transcriptional arrest for single bulges, prompting us to investigate the role of bulge size in transcriptional regulation.

Impact of bulge size on transcriptional regulation induced by buG4

To quantify the effect of the bulge size on the transcription efficiency of buG4, we selected positions 1, 3, and 5 from the upper and 4, 6, and 8 from the lower bulge po-

sitions in a G4 structure (Fig. 5A). We focused primarily on crowding conditions because buG4s with longer bulges were not sufficiently stable under noncrowding conditions to block transcription at physiological temperatures (37°C), resulting in unclear arrested bands that made the analysis inaccurate (Supplementary Fig. S13A–D). The denaturing gel electrophoresis result showed that templates with longer bulges at bulge position 1 (L-bT_2-1, L-bT_4-1, and L-bT_6-1) produced an arrested transcript-1 of 15 nt, similar to the L-pG4 sequence, thus indicating the formation of a fully folded buG4 structure during transcription (Fig. 5B). This result indicates that buG4 with a longer bulge can efficiently block transcription, whereas its blocking ability decreases with increasing bulge size. Additionally, a second transcript (transcript-2) was observed in the bulges at position 1 and the intensity of that band increased slightly with increasing bulge size (Fig. 5B). This suggested the formation of a partially folded structure during transcription. Based on the arrested positions of T7 RNA polymerase in the template strand (Supplementary Fig. S12C), we hypothesized the presence of an equilibrium between the fully and partially folded structures for buG4s with a bulge at position 1 (Fig. 6A). For bulge size changes in the upper part of buG4 (positions 3 and 5), arrested transcript-1 was produced (Fig. 5C), indicating the formation of fully folded buG4, which blocked the transcription process. Interestingly, along with transcript-1, longer arrested transcripts were also observed at position 5 (Fig. 5C). As the bulge size changed from 2 to 6 nt at position 5, the transcript length also increased, indicating that T7 RNA polymerase movement was hindered at the bulge position, leading to the production of transcripts of various lengths (Supplementary Fig. S12D). This result suggests that the possible formation of a fully folded buG4 and a misfolded structure during transcription caused the production of arrested transcript-1 and longer arrested transcripts, respectively (Fig. 6B). Unlike the bulge position 1, longer full transcripts were produced at both cases (positions 3 and 5). This suggests that the slippage of T7 RNA polymerase at the bulge position facilitates the production of elongated transcripts. Our previous study showed that T7 RNA polymerase can slip, producing transcripts that are 8–10 nt longer or short [22]. In contrast, the presence of a bulge at lower bulge positions (positions 4 and 6), blocked transcription with the formation of arrested transcript-1 without forming a slippage product (Fig. 5D). Therefore, the presence of a longer bulge region in the upper bulge positions may have induced more slippage products than bulges at the lower positions. Moreover, forming a fully folded structure, even for longer bulges, enhanced the arresting ability of buG4s with bulges in the lower positions, as represented for position 6 (Fig. 6C). On the contrary, increasing the bulge size at 3′ end, position 8 (bT_2-8, bT_4-8, and bT_6-8) did not result in the production of arrested transcript-1 of 15 nt (Fig. 5E). Instead, slightly longer, arrested bands were observed. The lengths of the arrested bands increased as the bulge size increased from 2 to 6 nt. As bulge position 8 was closer to the TSS, slippage along this bulge region appeared to restrict the formation of a fully folded buG4 structure (Supplementary Fig. S12G). Instead, it is likely that a partially folded structure and G-triplex were formed in equilibrium (Fig. 6D). Therefore, the structural diversity of buG4s, based on their bulge position, functions as a barrier for T7 RNA polymerase movement during transcription.

We further performed quantitative analysis to evaluate the efficiency of transcription arrest based on bulge size. The TE_{arrest} values calculated for various bulge sizes at different positions significantly decreased as the bulge size increased (Fig. 5F). For example, at position 1, the TE_{arrest} values for bT_1-1, bT_2-1, bT_4-1, and bT_6-1 are 88.5%, 52.4%, 31.9%, and 33.4%, respectively. A gradual decrease in TE_{arrest} with an increasing bulge size also observed for sequences with bulges at positions 3, 4, 5, and 6. However, the decrease was more pronounced at upper bulge positions 1, 3, and 5 than lower bulge positions 4 and 6. For example, increasing the bulge size from one to two nucleotides at position 3 (bT_1-3 to bT_2-3) resulted in a sharp decrease in TE_{arrest} from 80.3% to 38.2%. In contrast, increasing the bulge size from one to two nucleotides at position 4 (bT_1-4 to bT_2-4) decreased the TE_{arrest} from 79.3% to 62.9%. This result aligns well with our thermodynamic analysis, which demonstrated that the upper bulge positions 1, 3, 5, and 7 exhibited more destabilizing effects owing to the increased bulge size when compared to the lower bulge positions 2, 4, 6, and 8. The bulge position and size-dependent reduction in TE_{arrest} values indicate a strong dependence of transcriptional inhibition on the stability of the fully folded buG4 structure. However, the contrasting behavior of buG4 containing a bulge at position 8, which had the lowest TE_{arrest} value (<10%), revealed that TE_{arrest} in buG4 sequences was influenced not only by stability but also by the precise positioning of the bulge. Therefore, understanding the correlation between the bulge position-dependent stability of buG4 and transcription arrest efficiency due to fully folded buG4 formation is essential for accurately predicting the arrest efficiency of these sequences in genes.

We found a good correlation between the stability of buG4 structures, defined by the free energy change at 37°C ($-\Delta G^{\circ}_{37}$) and their transcriptional function, measured as TE_{arrest} (Supplementary Table S3 and Supplementary Fig. S14A). The data showed that as buG4 stability increases (more negative ΔG°_{37} values), TE_{arrest} also increases, though not indefinitely. The plot indicated that TE_{arrest} increased rapidly with buG4 stability, until it reached saturation, thus showing an exponential fit (equation 1).

$$TE_{arrest}\,(\%) = -200 \cdot e^{\frac{\Delta G^{o}_{37}}{3.0}} + 77 \qquad (1)$$

As shown in equation (1), when ΔG°_{37} becomes more negative, the exponential term decreases, resulting in smaller changes in TE_{arrest}, which approaches the value to 77%. Conversely, as ΔG°_{37} becomes less negative (or positive), the exponential term increases, significantly reducing TE_{arrest}. This suggests that TE_{arrest} is highly sensitive to the thermodynamic stability of the buG4 structures. We defined sequences with TE_{arrest} below 10% as unable to arrest transcription as below this value was similar to the background of the gel and this threshold corresponds to a ΔG°_{37} value of −3.3 kcal·mol^{-1} (equation 1). Therefore, buG4 sequences with a ΔG°_{37} value $\leq$ −3.3 kcal·mol^{-1} can efficiently arrest transcription. This equation provides a strategy to understand the relationship between the energetic properties of buG4s and their functional roles in transcriptional regulation (Supplementary Fig. S14B). This insight is crucial for predicting whether buG4 sequences in cancer-related genes can effectively inhibit transcription.

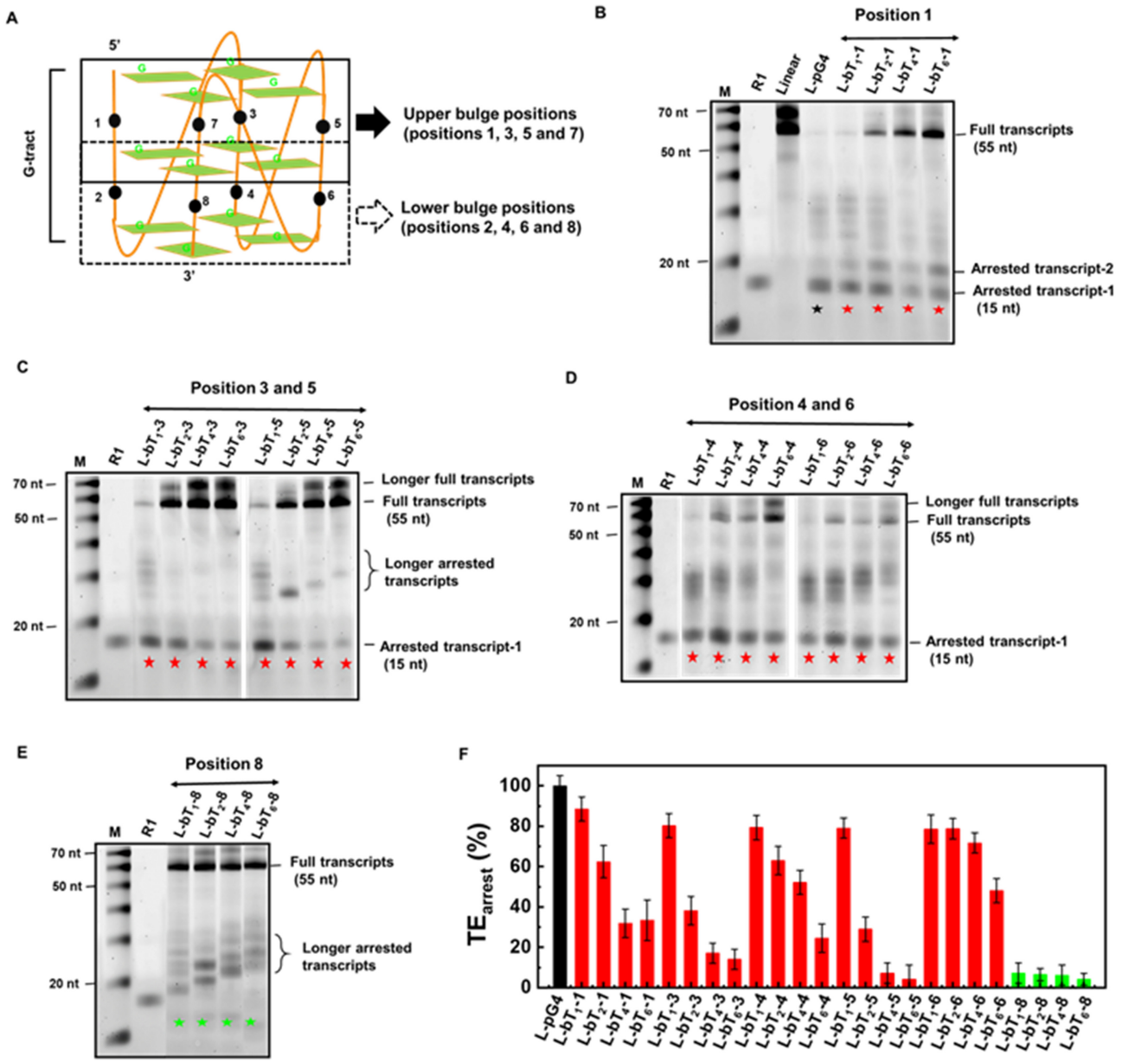


Figure 5. (**A**) The bulge positions in a G4 structure are separated in two categories. The upper region includes the bulge positions 1, 3, 5, and 7 and the lower region includes the bulge positions 2, 4, 6, and 8. Denaturing gel electrophoresis results for the transcript production containing template sequences with bulges at (**B**) position 1, (**C**) positions 3 and 5, (**D**) positions 4 and 6, and (**E**) position 8 under crowding condition. Denaturing gel electrophoresis was used to identify the products of transcription reactions carried out for 120 min at 37°C. The marked lanes represent the templates containing Linear, pG4 and buG4s with bulges at position 1 (bT_1-1, bT_2-1, bT_4-1, and bT_6-1), at positions 3 and 5 (bT_1-3, bT_2-3, bT_4-3, bT_6-3, bT_1-5, bT_2-5, bT_4-5, and bT_6-5), at positions 4 and 6 (bT_1-4, bT_2-4, bT_4-4, bT_6-4, bT_1-6, bT_2-6, bT_4-6 and bT_6-6) and at position 8 (Tb_1-8, Tb_2-8, Tb_4-8, and Tb_6-8), in crowding condition. *M* is the size marker. The 15 nt arrested bands from pG4 and buG4 containing templates are marked with stars. (**F**) The transcription arrested efficiency (TE_{arrest}) from each template DNA under crowding condition. TE_{arrest} values below 10% are considered as no arrest.

Table 3. The $\Delta G°_{37}$ values and measured transcription arrested efficiency (TE_{arrest}) of the wild type buG4 sequences from human genes

Name	Sequences (5′- 3′) [a]	$\Delta G°_{37}$ [b] (kcal mol^{-1})	Predicted Arrest	Measured TE_{arrest} (%)[c]	Reference
kras-bT_1-4	AGGGCGGTGTGGGAATAGGGAA*	−7.1 ± 0.3	Yes	39.8 ± 7.2	[21]
parp1-bA_1-3	TGGGGCCGAGGCGGGGCTTGGG*	–[d]	Yes	33.4 ± 7.5	[20]
myc- bC_1-1	TGCGGCTAGGGGACAGGGGCGGGT	−4.2 ± 0.5	Yes	16.8 ± 6.5	Our study
kit-bT_1-5	TGGGCTGGGCCACTGTGGCTGGGT	−1.5 ± 0.3	No	<10	Our study
myc-bA_1-7	TGGGAGGGCTGGGCCAGAGGT	−3.2 ± 0.2	No	<10	Our study

[a]Wild-type buG4 sequences containing bulges. The predicted bulge positions are indicated in bold and are underlined.
[b]Experiments were performed in 10 mM Tris–HCl buffer containing 100 mM KCl, 8 mM $MgCl_2$ and 10 wt% PEG 200 (pH 7.4 at 37°C). [c]TE_{arrest} < 10% indicated no arrest. [d](–) indicates very high thermal stability of the buG4.

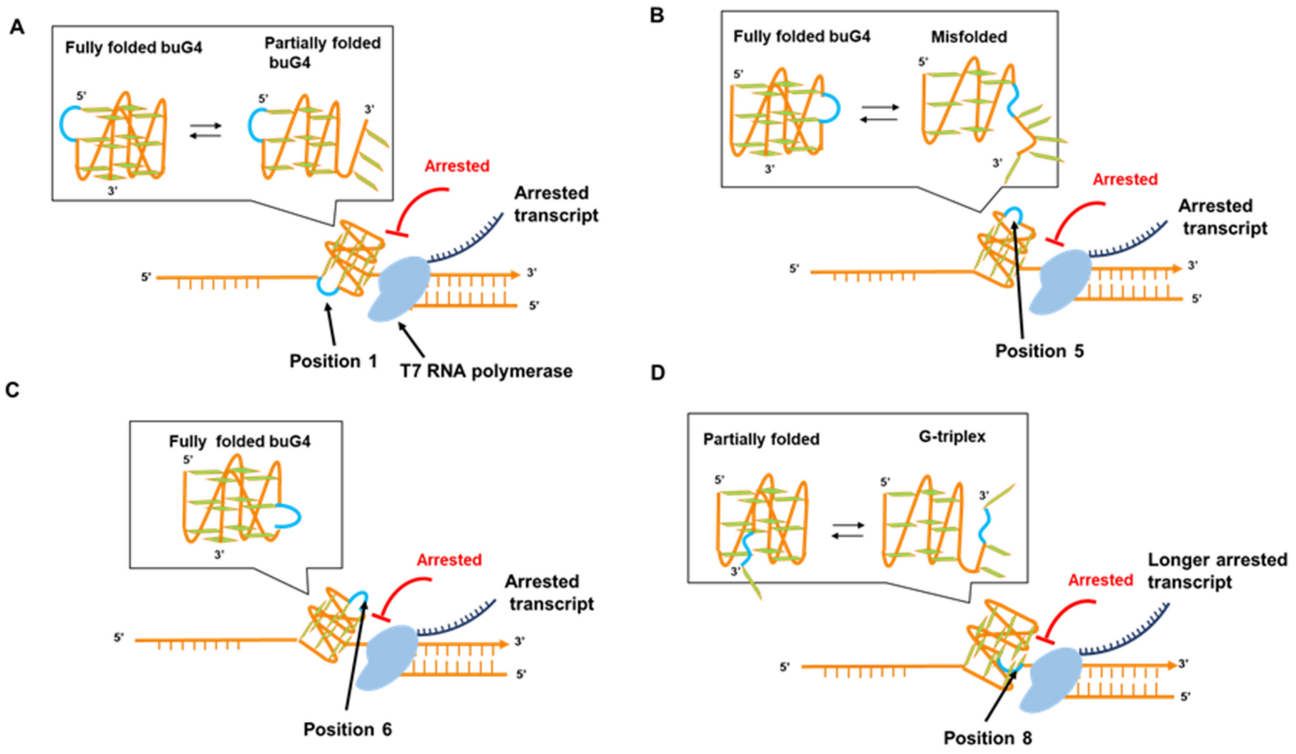


Figure 6. The probable structural equilibrium during transcription for buG4 containing template sequences with bulge at (**A**) position 1, (**B**) position 5, (**C**) position 6, and (**D**) position 8.

Transcription regulatory function of the buG4 sequences from cancer-related genes

The presence of bulges in gene sequences is inherently more complex than in the model sequences used in this study because the sequences contain longer loops with various base compositions. To date, very few bulge-containing sequences of human disease-associated genes have been studied [20, 21]. Here, we chose two known buG4 sequences from *PARP1* and *KRAS* genes, because the formation of bulge-containing intramolecular structures is known (Table 3) [20, 21]. Additionally, we identified potential buG4 forming sequences from both the template and nontemplate strands of the crucial cancer-related genes, *MYC* and *KIT*, guided by G4-iM Grinder predictions [47] (sequences are listed in Supplementary Table S4). Furthermore, formation of a stable buG4 structure was confirmed using by CD spectroscopy and UV-melting curves (Supplementary Fig. S15A–H and Supplementary Fig. S16A–H, detailed explanations are provided in the figure captions).

Because, the selected sequences formed parallel structures under the experimental conditions (Supplementary Fig. S17), to validate our prediction model, *in vitro* transcription assays were performed using the gene sequences to investigate their potential to inhibit transcriptional regulation (template sequences are provided in Supplementary Table S2). The buG4 sequences from these genes were integrated into the X region of the template sequence as shown in Fig. 4A. Transcription was initially performed under 100 mM LiCl conditions for the buG4 sequences from the *MYC* gene (Supplementary Fig. S18). Controls included linear sequences (unable to form secondary structures) and pG4 sequences (capable of forming G4 structures), transcribed under the same conditions. The performed gel analysis showed no arrested transcript-1 (Supplementary Fig. S18), indicating that buG4 sequences did not impede transcription under these conditions. This confirmed that the buG4 structures did not form in the presence of 100 mM LiCl, supporting the hypothesis that efficient transcriptional arrest occurs due to stable buG4 structures. When transcription was performed under crowding conditions, the transcription process was halted, producing a 15 nt shorter RNA transcript-1, marked with stars (Supplementary Fig. S19). Additionally, a longer arrested transcript-2 was detected, suggesting that both the fully folded and partially folded structures coexisted within the buG4 sequences, as seen in the model sequences. The length of transcript-2 varied for each buG4 sequence, likely due to differences in sequence length and loop size across gene sequences (Supplementary Fig. S19). However, by focusing solely on fully folded buG4 formation, we quantified TEarrest and found that sequences predicted by our model to efficiently arrest transcription exhibited TE_{arrest} values >10%. In contrast, sequences predicted not to arrest transcription had TE_{arrest} values <10%. For example, kras-bT_1-4, with a ΔG°_{37} value of −7.1 kcal·mol^{-1}, showed 39.8% arrest, whereas kit-bT_1-5, with a ΔG°_{37} value of −1.5 kcal·mol^{-1}, showed <10% arrest. Therefore, it is apparent that our model can efficiently screen potential buG4 sequences from genes based on their ΔG°_{37} values. It is important to note that this model cannot be applied to buG4 sequences with a bulge at position 8 (3′ end), as these were excluded during the development of the model. Despite the strong predictive power of the model for buG4 function, topology remains a critical factor for improving accuracy. Our model was developed using only parallel G4 conformations, whereas G-rich sequences can adopt diverse topologies depending on the base compositions of the loops [4, 46, 48]. Incorporating other G4 conformations in fu-

ture models could enhance predictive power, offering deeper insights into the complex dynamics of gene regulation.

Conclusion

In conclusion, our study advances the understanding of the role of bulges on G4 stability and transcriptional regulation, particularly in the context of bulge position and size. The detailed thermodynamic stabilities will be useful in broadening the current bioinformatics analysis to accurately predict buG4 sequences, especially under cell-mimetic crowding condition containing PEG 200. Moreover, we propose a relationship between the free energy change of buG4 formation (ΔG°_{37}) and transcription arrest efficiency under intracellular crowding conditions. According to our model, a buG4 sequence with a ΔG°_{37} value ≤ -3.3 kcal·mol^{-1} can arrest the transcription process, whereas a buG4 sequence with a ΔG°_{37} value > -3.3 kcal·mol^{-1} cannot arrest the transcription process. This result offers valuable insight into estimating the efficiency of target buG4 sequences in either arresting or facilitating transcription, presenting a novel approach for exploiting imperfect G4 sequences in cancer-related genes as an emerging target.

Acknowledgements

The authors thank Ms. Kayoko Kishi and Ms. Shoko Sakai for their assistance with native gel experiments and UV measurements.

Author contributions: Sunipa Sarkar (Conceptualization, Methodology, Investigation, Visualization, Writing—original draft), Hisae Tateishi-Karimata (Conceptualization, Methodology, Supervision, Writing—original draft, Writing—review & editing), Tatsuya Ohyama (Methodology), and Naoki Sugimoto (Conceptualization, Supervision, Writing—review & editing)

Supplementary data

Supplementary data is available at NAR online.

Conflict of interest

None declared.

Funding

This work was supported by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) and the Japan Society for the Promotion of Science (JSPS), especially for Grant-in-Aid for Scientific Research (S) (22H04975) and Grant-in-Aid for Transformative Research Areas (B) [21H05107 and JSPS Core-to-Core Program (grant number: JPJSCCA20220005)]. The Konan New Century Strategic Research Project, Okazaki Kazuo Foundation of Konan Gakuen for Advanced Scientific Research and Chubei Itoh Foundation. Funding to pay the Open Access publication charges for this article was provided by Chubei Itoh Foundation.

Data availability

The data underlying this article are available in the article and in its online Supplementary Data.

References

1. Moser HE, Dervan PB. Sequence-specific cleavage of double helical DNA by triple helix formation. *Science* 1987;**238**:645–50. https://doi.org/10.1126/science.3118463
2. Gellert M, Lipsett MN, Davies DR. Helix formation by guanylic acid. *Proc Natl Acad Sci USA* 1962;**48**:2013–8. https://doi.org/10.1073/pnas.48.12.2013
3. Gehring K, Leroy J-L, Guéron M. A tetrameric DNA structure with protonated cytosine–cytosine base pairs. *Nature* 1993;**363**:561–5. https://doi.org/10.1038/363561a0
4. Takahashi S, Kotar A, Tateishi-Karimata H *et al.* Chemical modulation of DNA replication along G-quadruplex based on topology-dependent ligand binding. *J Am Chem Soc* 2021;**143**:16458–69. https://doi.org/10.1021/jacs.1c05468
5. Tateishi-Karimata H, Kawauchi K, Sugimoto N. Destabilization of DNA G-quadruplexes by chemical environment changes during tumor progression facilitates transcription. *J Am Chem Soc* 2018;**140**:642–51. https://doi.org/10.1021/jacs.7b09449
6. Endoh T, Sugimoto N. Conformational dynamics of the RNA G-quadruplex and its effect on translation efficiency. *Molecules* 2019;**24**:1613. https://doi.org/10.3390/molecules24081613
7. Huppert JL, Balasubramanian S. Prevalence of quadruplexes in the human genome. *Nucleic Acids Res* 2005;**33**:2908–16. https://doi.org/10.1093/nar/gki609
8. Huppert JL, Balasubramanian S. G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res* 2007;**35**:406–13. https://doi.org/10.1093/nar/gkl1057
9. Largy E, Marchand A, Amrane S *et al.* Quadruplex turncoats: cation-dependent folding and stability of quadruplex–DNA double switches. *J Am Chem Soc* 2016;**138**:2780–92. https://doi.org/10.1021/jacs.5b13130
10. Rachwal PA, Findlow IS, Werner JM *et al.* Intramolecular DNA quadruplexes with different arrangements of short and long loops. *Nucleic Acids Res* 2007;**35**:4214–22. https://doi.org/10.1093/nar/gkm316
11. Papp C, Mukundan VT, Jenjaroenpun P *et al.* Stable bulged G-quadruplexes in the human genome: identification, experimental validation and functionalization. *Nucleic Acids Res* 2023;**51**:4148–77. https://doi.org/10.1093/nar/gkad252
12. Cagirici HB, Budak H, Sen TZ. Genome-wide discovery of G-quadruplexes in barley. *Sci Rep* 2021;**11**:7876. https://doi.org/10.1038/s41598-021-86838-3
13. Hon J, Martínek T, Zendulka J *et al.* pqsfinder: an exhaustive and imperfection-tolerant search tool for potential quadruplex-forming sequences in R. *Bioinformatics* 2017;**33**:3373–9. https://doi.org/10.1093/bioinformatics/btx413
14. Mukundan VT, Phan AT. Bulges in G-quadruplexes: broadening the definition of G-quadruplex-forming sequences. *J Am Chem Soc* 2013;**135**:5017–28. https://doi.org/10.1021/ja310251r
15. Ngoc Nguyen TQ, Lim KW, Phan AT. Duplex formation in a G-quadruplex bulge. *Nucleic Acids Res* 2020;**48**:10567–75. https://doi.org/10.1093/nar/gkaa738
16. Wang K-B, Dickerhoff J, Wu G *et al.* PDGFR-β promoter forms a vacancy G-quadruplex that can Be filled in by dGMP: solution structure and molecular recognition of guanine metabolites and drugs. *J Am Chem Soc* 2020;**142**:5204–11. https://doi.org/10.1021/jacs.9b12770
17. Wu F, Niu K, Cui Y *et al.* Genome-wide analysis of DNA G-quadruplex motifs across 37 species provides insights into G4 evolution. *Commun Biol* 2021;**4**:98. https://doi.org/10.1038/s42003-020-01643-4
18. He Y-d, Zheng K-w, Wen C-j *et al.* Selective targeting of guanine-vacancy-bearing G-quadruplexes by G-quartet complementation and stabilization with a guanine–peptide conjugate. *J Am Chem Soc* 2020;**142**:11394–403. https://doi.org/10.1021/jacs.0c00774
19. Zhang Y, Cheng Y, Chen J *et al.* Mechanical diversity and folding intermediates of parallel-stranded G-quadruplexes with a bulge.

Nucleic Acids Res 2021;**49**:7179–88. https://doi.org/10.1093/nar/gkab531
20. Sengar A, Vandana JJ, Chambers VS *et al.* Structure of a (3+1) hybrid G-quadruplex in the PARP1 promoter. *Nucleic Acids Res* 2019;**47**:1564–72. https://doi.org/10.1093/nar/gky1179
21. Kerkour A, Marquevielle J, Ivashchenko S *et al.* High-resolution three-dimensional NMR structure of the KRAS proto-oncogene promoter reveals key features of a G-quadruplex involved in transcriptional regulation. *J Biol Chem* 2017;**292**:8082–91. https://doi.org/10.1074/jbc.M117.781906
22. Tateishi-Karimata H, Isono N, Sugimoto N. New insights into transcription fidelity: thermal stability of non-canonical structures in template DNA regulates transcriptional arrest, pause, and slippage. *PLoS One* 2014;**9**:e90580. https://doi.org/10.1371/journal.pone.0090580
23. Tateishi-Karimata H, Sugimoto N. Roles of non-canonical structures of nucleic acids in cancer and neurodegenerative diseases. *Nucleic Acids Res* 2021;**49**:7839–55. https://doi.org/10.1093/nar/gkab580
24. Banerjee D, Tateishi-Karimata H, Toplishek M *et al.* In-cell stability prediction of RNA/DNA hybrid duplexes for designing oligonucleotides aimed at therapeutics. *J Am Chem Soc* 2023;**145**:23503–18. https://doi.org/10.1021/jacs.3c06706
25. Takahashi S, Yamamoto J, Kitamura A *et al.* Characterization of intracellular crowding environments with topology-based DNA quadruplex sensors. *Anal Chem* 2019;**91**:2586–90. https://doi.org/10.1021/acs.analchem.8b04177
26. Ghosh S, Takahashi S, Banerjee D *et al.* Nearest-neighbor parameters for the prediction of RNA duplex stability in diverse *in vitro* and cellular-like crowding conditions. *Nucleic Acids Res* 2023;**51**:4101–11. https://doi.org/10.1093/nar/gkad020
27. Cui J, Waltman P, Le VH *et al.* The effect of molecular crowding on the stability of human c-MYC promoter sequence I-motif at neutral pH. *Molecules* 2013;**18**:12751–67. https://doi.org/10.3390/molecules181012751
28. Fujimoto T, Nakano S, Miyoshi D *et al.* The effects of molecular crowding on the structure and stability of g-quadruplexes with an abasic site. *J Nucleic Acids* 2011;**2011**:857149. https://doi.org/10.4061/2011/857149
29. Nakano S, Miyoshi D, Sugimoto N. Effects of molecular crowding on the structures, interactions, and functions of nucleic acids. *Chem Rev* 2014;**114**:2733–58. https://doi.org/10.1021/cr400113m
30. Cheng M, Chen J, Ju H *et al.* Drivers of i-DNA formation in a variety of environments revealed by four-dimensional UV melting and annealing. *J Am Chem Soc* 2021;**143**:7792–807. https://doi.org/10.1021/jacs.1c02209
31. Mergny J-L, Lacroix L. UV melting of G-quadruplexes. *CP Nucleic Acid Chem* 2009;**37**:17.11.11–5. https://doi.org/10.1002/0471142700.nc1701s37
32. Esposito V, Pirone L, Mayol L *et al.* Exploring the binding of d (GGGT)4 to the HIV-1 integrase: an approach to investigate G-quadruplex aptamer/target protein interactions. *Biochimie* 2016;**127**:19–22. https://doi.org/10.1016/j.biochi.2016.04.013
33. Jing N, Marchand C, Liu J *et al.* Mechanism of inhibition of HIV-1 integrase by G-tetrad-forming oligonucleotides *in vitro* *. *J Biol Chem* 2000;**275**:21460–7. https://doi.org/10.1074/jbc.M001436200
34. Rivieccio E, Tartaglione L, Esposito V *et al.* Structural studies and biological evaluation of T30695 variants modified with single chiral glycerol-T reveal the importance of LEDGF/p75 for the aptamer anti-HIV-integrase activities. *Biochim Biophys Gen Sub* 2019;**1863**:351–61. https://doi.org/10.1016/j.bbagen.2018.11.001
35. Matsumoto S, Tateishi-Karimata H, Takahashi S *et al.* Effect of molecular crowding on the stability of RNA G-quadruplexes with various numbers of quartets and lengths of loops. *Biochemistry* 2020;**59**:2640–9. https://doi.org/10.1021/acs.biochem.0c00346
36. Ghosh S, Takahashi S, Ohyama T *et al.* Nearest-neighbor parameters for predicting DNA duplex stability in diverse molecular crowding conditions. *Proc Natl Acad Sci USA* 2020;**117**:14194–201. https://doi.org/10.1073/pnas.1920886117
37. Buscaglia R, Miller MC, Dean WL *et al.* Polyethylene glycol binding alters human telomere G-quadruplex structure by conformational selection. *Nucleic Acids Res* 2013;**41**:7934–46. https://doi.org/10.1093/nar/gkt440
38. Kilburn D, Roh JH, Guo L *et al.* Molecular crowding stabilizes folded RNA structure by the excluded volume effect. *J Am Chem Soc* 2010;**132**:8690–6. https://doi.org/10.1021/ja101500g
39. Risitano A, Fox KR. Stability of intramolecular DNA quadruplexes: comparison with DNA duplexes. *Biochemistry* 2003;**42**:6507–13. https://doi.org/10.1021/bi026997v
40. Wan C, Fu W, Jing H *et al.* NMR solution structure of an asymmetric intermolecular leaped V-shape G-quadruplex: selective recognition of the d (G2NG3NG4) sequence motif by a short linear G-rich DNA probe. *Nucleic Acids Res* 2019;**47**:1544–56. https://doi.org/10.1093/nar/gky1167
41. Minetti CA, Remeta DP, Dickstein R *et al.* Energetic signatures of single base bulges: thermodynamic consequences and biological implications. *Nucleic Acids Res* 2010;**38**:97–116. https://doi.org/10.1093/nar/gkp1036
42. Guédin A, De Cian A, Gros J *et al.* Sequence effects in single-base loops for quadruplexes. *Biochimie* 2008;**90**:686–96. https://doi.org/10.1016/j.biochi.2008.01.009
43. Piazza A, Adrian M, Samazan F *et al.* Short loop length and high thermal stability determine genomic instability induced by G-quadruplex-forming minisatellites. *EMBO J* 2015;**34**:1718–34. https://doi.org/10.15252/embj.201490702
44. Reshetnikov RV, Kopylov AM, Golovin AV. Classification of g-quadruplex DNA on the basis of the quadruplex twist angle and planarity of g-quartets. *Acta Naturae* 2010;**2**:72–81. https://doi.org/10.32607/20758251-2010-2-4-72-81
45. Guédin A, Gros J, Alberti P *et al.* How long is too long? Effects of loop size on G-quadruplex stability. *Nucleic Acids Res* 2010;**38**:7858–68. https://doi.org/10.1093/nar/gkq639
46. Cheng M, Cheng Y, Hao J *et al.* Loop permutation affects the topology and stability of G-quadruplexes. *Nucleic Acids Res* 2018;**46**:9264–75. https://doi.org/10.1093/nar/gky757
47. Belmonte-Reche E, Morales JC. G4-iM Grinder: when size and frequency matter. G-Quadruplex, i-motif and higher order structure search and analysis tool. *NAR Genom Bioinform* 2019;**2**:lqz005. https://doi.org/10.1093/nargab/lqz005
48. Chen J, Cheng M, Salgado GF *et al.* The beginning and the end: flanking nucleotides induce a parallel G-quadruplex topology. *Nucleic Acids Res* 2021;**49**:9548–59. https://doi.org/10.1093/nar/gkab681

Received: October 23, 2024. **Revised:** February 5, 2025. **Editorial Decision:** February 11, 2025. **Accepted:** February 19, 2025