

Review

Sequence Determinants of G-Quadruplex Thermostability: Aligning Evidence from High-Precision Biophysics and High-Throughput Genomics

Ke Xiao ¹ , Jiye Fu ¹, Rongxin Zhang ² and Jing Tu ^{1,*} 

¹ State Key Laboratory of Digital Medical Engineering, School of Biological Science and Medical Engineering, Southeast University, Nanjing 211189, China; kexiao@seu.edu.cn (K.X.)

² Robert Lurie Comprehensive Cancer Center, Department of Obstetrics and Gynecology, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, USA

* Correspondence: jtu@seu.edu.cn

Abstract

G-quadruplexes (G4s) are non-canonical nucleic acid structures that function as key regulatory elements in crucial cellular processes. Their biological functions are intrinsically linked to thermostability, which is governed by specific sequence features. This review systematically synthesizes evidence from high-precision biophysical studies and high-throughput genomic assays to delineate the sequence determinants of G4 thermostability. Analyses align the trends derived from both methodological paradigms and establish that stability emerges from a complex interplay among three structural elements: the G-tract core, whose length and integrity generally govern stability despite notable exceptions such as the anomalous stability of short G-tracts with 1-nt loops and the stabilization induced by large, structured bulges; the loops, which exhibit a consistent inverse relationship between length and stability across methods, though with context-dependent compositional effects and methodological disparities; and the flanking sequences, whose composition modulates stability and can bias topological outcomes. By integrating findings across scales, this work provides a unified conceptual framework connecting biophysical measurements with genomic observations—a critical step toward computationally predicting G4 stability, topology, and function directly from sequence, thereby advancing the understanding of their roles in health and disease.

Keywords: G-quadruplex; thermostability; sequence features; biophysical assay; high-throughput sequencing



Academic Editor: Tigran Chalikian

Received: 23 October 2025

Revised: 18 November 2025

Accepted: 19 November 2025

Published: 20 November 2025

Citation: Xiao, K.; Fu, J.; Zhang, R.;

Tu, J. Sequence Determinants of

G-Quadruplex Thermostability:

Aligning Evidence from

High-Precision Biophysics and

High-Throughput Genomics.

Biomolecules **2025**, *15*, 1632. [https://](https://doi.org/10.3390/biom15111632)

doi.org/10.3390/biom15111632

Copyright: © 2025 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article

distributed under the terms and

conditions of the Creative Commons

Attribution (CC BY) license

([https://creativecommons.org/](https://creativecommons.org/licenses/by/4.0/)

[licenses/by/4.0/](https://creativecommons.org/licenses/by/4.0/)).

1. Introduction: The Challenge of Quantifying G-Quadruplex Stability Across Scales

1.1. The Biological Significance of G4 Stability: From Biophysical Property to Functional Determinant

G-quadruplexes (G4s) are not merely structural curiosities but are widely recognized as functional regulatory elements whose biological roles are intrinsically linked to their stability [1–4]. These four-stranded nucleic acid structures, formed by the folding of guanine (G)-rich sequences, are broadly involved in a range of critical cellular processes, including the regulation of gene expression, DNA replication, telomere maintenance, and the preservation of genomic integrity [1,5–7]. Genomic analyses have revealed that potential G-quadruplex sequences (PQS) are not randomly distributed but are significantly

enriched in key functional regions such as gene promoters, 5' untranslated regions (5' UTRs), telomeres, and replication origins [1,8–10]. In these contexts, G4 structures can function as physical barriers to enzymatic machinery like DNA or RNA polymerases or serve as recognition sites for specific regulatory proteins, thereby facilitating precise control over genetic activity [6,11].

The thermostability of a G4 structure—quantified by its resistance to thermal denaturation—directly governs its persistence and abundance within the cellular environment, which in turn determines the strength and duration of its biological function [3,12]. A highly stable G4 can act as a persistent “molecular switch” [12–15], whereas a less stable structure may form transiently to participate in dynamic regulatory events [15–17]. This correlation between stability and function has positioned G4s as attractive therapeutic targets, particularly in oncology [18–20]. For instance, stabilizing G4 structures in oncogene promoters or telomeric regions using small-molecule ligands can effectively suppress cancer cell proliferation and invasion [6,21].

The critical link between G4 stability and function is further underscored by evolutionary biology. A genome-wide study by Guiblet et al. revealed a pivotal pattern: G4 sequences located within functional genomic elements (e.g., promoters, UTRs, and replication origins) are not only overrepresented but also exhibit significantly higher thermostability and are under purifying selection. In contrast, unstable G4 motifs in regions like the non-transcribed strands of exons evolve neutrally, implying they are either non-functional or potentially deleterious [12]. Parallel evidence from the earlier work of Puig Lombardi et al. demonstrated negative selection against extremely stable G4 structures, which mitigates their adverse effects on genome integrity while preserving beneficial G4-related functions [22]. Together, these findings from both groups converge to suggest that G4 thermostability is not merely a random physicochemical trait, but a parameter tuned by evolution to fulfill specific biological roles. Consequently, the accurate quantification and prediction of G4 stability have progressed from a biophysical characterization to a central question in functional genomics.

1.2. The Methodological Divide in Stability Assessment

The current assessment of G4 stability relies on two distinct methodological paradigms, which differ fundamentally in their precision, throughput, and the nature of the information they yield, creating a significant methodological divide.

The first paradigm encompasses high-precision, low-throughput biophysical techniques. These methods focus on the precise characterization of individual, well-defined oligonucleotide sequences *in vitro*. Representative techniques include UV melting [23,24], circular dichroism (CD) [25], differential scanning calorimetry (DSC) [26], fluorescence resonance energy transfer (FRET) [27,28] and molecular dynamics (MD) simulations [7]. These techniques yield direct, quantitative thermodynamic parameters. The most common metric is the melting temperature (T_m), at which 50% of the G4 structures denature into single strands. By analyzing the melting curve, one can also derive a suite of precise thermodynamic parameters, including the Gibbs free energy change (ΔG), enthalpy change (ΔH), and entropy change (ΔS). The principal advantage of these methods is the directness and precision of the measurements, establishing them as the “gold standard” for G4 stability assessment. However, their critical limitation is extremely low throughput, as each experiment typically analyzes only a single or a few sequences, precluding their application on a genome-wide scale.

The second paradigm consists of high-throughput, indirect genomic methods. These techniques leverage the obstructive effect of G4 structures on DNA polymerase, coupled with high-throughput sequencing. The most prominent example is G4-Seq [8,29]. In a stan-

standard G4-Seq experiment, DNA polymerase progression is challenged under G4-stabilizing conditions (e.g., in the presence of K^+ or specific ligands). A stable G4 on the template strand acts as a physical barrier, inducing polymerase pausing or dissociation [8,11]. This stalling event manifests in sequencing data as a sharp decline in read quality and a significant increase in the base mismatch rate (Mismatch Percentage, MM%) immediately downstream (3') of the G4 initiation site [8]. Furthermore, Tu et al. (2021) developed the G4-miner algorithm, which infers G4 formation by detecting subtle fluctuations in sequencing quality scores from standard whole-genome resequencing data [30]. Thus, stability is inferred indirectly: a higher MM% or a more pronounced quality drop is interpreted as evidence of a more stable G4 that more effectively impedes the polymerase [8,30]. The unparalleled advantage of these methods is their throughput, enabling the mapping of hundreds of thousands of potential G4 structures across an entire genome in a single experiment. Their limitation, however, is that the sequencing signal is an indirect proxy for stability, not a direct thermodynamic measurement, and can be confounded by other factors.

1.3. Focusing on Sequence Determinants

An “alignment problem” exists between these two methodological classes. They measure fundamentally different physical quantities: biophysical methods like UV melting assess a thermodynamic equilibrium property (T_m), whereas high-throughput sequencing assays measure the kinetics of an enzymatic process (polymerase stalling efficiency). While a thermodynamically more stable G4 is likely to present a stronger kinetic barrier, this relationship is not guaranteed. Factors such as G4 unfolding kinetics, specific interactions with the polymerase, and local structural topology can independently influence pausing efficiency, decoupling it from the overall T_m .

Researchers have provided initial clues to help bridge this divide. For instance, Wang et al. demonstrated that enzyme-G4 interactions, such as Pif1-mediated unwinding, are sensitive to G4 stability and are thus influenced by structural features like loop length [31]. Similarly, Sahakyan et al. confirmed that the MM% values from G4-Seq correlate with the biophysically assessed stability (via CD, 1H NMR, and UV melting) of selected G4 sequences [32]. However, systematically validating the relationship between thermodynamic stability and high-throughput sequencing signals (e.g., MM% from G4-Seq or quality scores from G4-miner) across an entire genome, with its millions of G4 motifs, remains a formidable challenge.

A pragmatic approach to this problem is to compare the readouts from both experimental paradigms, observing whether they exhibit consistent relative changes in response to identical sequence features. If consistent correlative patterns emerge, it would suggest that both methods capture a core, shared attribute of G4 stability. Establishing such a correspondence would greatly facilitate the extrapolation of G4 stability understanding and its application across the entire genome. Therefore, this review aims to collate and integrate evidence from both fields to provide a theoretical foundation for connecting them.

Before delving into sequence features, it is crucial to acknowledge that G4 stability is profoundly influenced by the intracellular environment. For instance, K^+ generally stabilizes G4s more effectively than Na^+ . This preference is driven by the greater energetic cost of dehydrating Na^+ ; although Na^+ binding to a quadruplex is thermodynamically more favorable, the dehydration penalty for Na^+ far exceeds that for K^+ [4]. Furthermore, the crowded intracellular environment (macromolecular crowding) preferentially stabilizes the compact, folded state of G4s through the excluded volume effect, thereby promoting their formation and stability [33–35]. Detailed discussions on these topics can be found in reviews by Jana and Weisz [36] and Nishio et al. [37].

While these environmental factors provide the essential context for *in vivo* G4 formation, this review will deliberately focus on elucidating the intrinsic sequence features that determine G4 thermostability under constant, well-defined experimental conditions (typically in K⁺-containing buffer). By systematically isolating and analyzing the influence of core sequence elements—the G-tracts, loops, and flanking regions—this study aims to synthesize the empirical rules that link sequence features to stability data, establish a set of generalizable principles by examining both consistencies and discrepancies between the two approaches (Figure 1), and bridge the findings from low-throughput biophysical methods and high-throughput genomic assays.

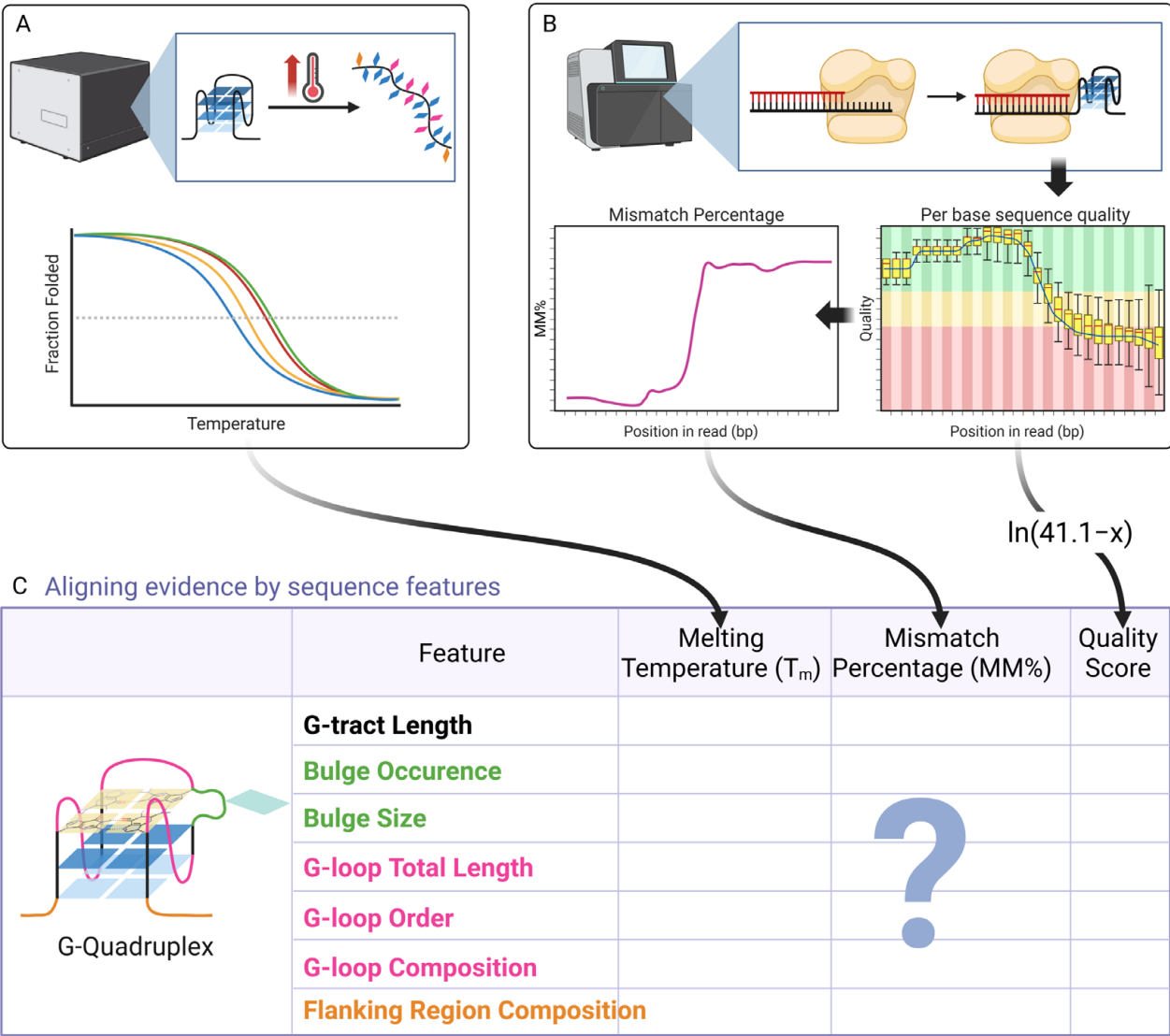


Figure 1. Conceptual framework for aligning evidence from high-precision biophysical and high-throughput sequencing methods. **(A)** High-precision biophysical methods, exemplified by UV melting and CD melting assays, primarily yield thermodynamic parameters such as melting temperature (T_m). **(B)** High-throughput sequencing methods, such as G4-Seq and G4-miner, provide genome-wide metrics including base mismatch percentage (MM%) and sequencing quality scores. The original quality scores were transformed to ensure that higher values correspond to greater G4 stability, consistent with the interpretation of MM%. **(C)** Readouts from both methodological paradigms were systematically compared and aligned to determine whether they exhibit consistent responses to identical sequence features, as illustrated by six representative sequence features analyzed in this study.

For each sequence feature, studies based on high-precision biophysical methods were reviewed, typically focusing on T_m trends derived from assays such as UV melting and CD melting (Figure 1A). Accompanying explanations for these trends—including topological alterations, conformational details, and associated energy changes—were also compiled. These T_m -based trends were subsequently compared and aligned with corresponding data from high-throughput sequencing methods, specifically Mismatch Percentage (MM%) from G4-Seq and sequencing quality scores from G4-miner (Figure 1B, Appendix A). To ensure consistent interpretation across metrics, the original G4-miner quality scores (x) were transformed using the function $\ln(41.1 - x)$. This transformation was reasonable as the native Phred + 33 quality scores (range: 0–41) are inversely correlated with sequencing error rates. The operation $41.1 - x$ reverses this relationship to generate positive values, while the subsequent natural logarithm enhances discrimination near the maximum score, i.e., 41, thereby aligning the metric with MM% where higher values indicate greater G4 stability.

2. The G-Tract Core: The Impact of Size and Imperfections on Stability

The G-tracts that constitute the core of the G-quadruplex are fundamental to its structural integrity [38,39]. This section examines how the length of these G-tracts and their sequence integrity—specifically, the presence of non-guanine base insertions that create bulges—influence the overall stability of the G4 structure.

2.1. G-Tract Length: Incomplete Monotonic Relationship with Stability

Intuitively, longer G-tracts would be expected to confer greater stability due to the stacking of additional G-tetrad layers. However, experimental evidence reveals that the relationship between G-tract length and G4 stability is not a simple monotonic progression. Instead, it is strongly modulated by the length of the connecting loops, exhibiting complex, incomplete monotonic characteristics [40].

The research by Rachwal et al. provides a demonstration of this phenomenon. They compared the stability of two sequence series: one with G-tracts connected by a single thymine (T). For these $d(G_nT)_4$ sequences, the stability order was unexpectedly non-monotonic: $n = 3 > n = 7 > n = 6 > n = 5 > n = 4$. The sequence with the shortest G-tracts, $d(G_3T)_4$, exhibited “anomalous stability,” possessing a T_m significantly higher than counterparts with longer G-tracts (G_4 , G_5 , G_6). In contrast, when the loop length was increased to two thymines (T2) in the case of $d(G_nT_2)_4$, the stability trend returns to the intuitive pattern, increasing monotonically with G-tract length: $n = 7 > n = 6 > n = 5 > n = 4 > n = 3$.

The “anomalous stability” of $d(G_3T)_4$ underscores a delicate balance between two opposing forces: the enthalpic gain from increased G-tetrad stacking and the entropic/steric cost imposed on the connecting loops. While each additional G-tetrad layer contributes favorable enthalpy, in an intramolecular G4 connected by the very short (1-nt) V-shaped loops, longer G-tracts force these loops to span a greater spatial distance. This imposes significant conformational strain on the phosphodiester backbone, offsetting the energetic benefits of extra tetrads. This interpretation will be supported by later discussions on loop length (Section 3.1), wherein short loops favor parallel topologies due to steric constraints, while longer loops afford the flexibility for alternative antiparallel or hybrid folds [41–43].

2.2. Bulges in G-Tracts: The Impact of Imperfections

Although the canonical G4 model requires continuous G-tracts, accumulating evidence demonstrates that G4s with structural imperfections in their core are not only feasible but

also biologically significant. These deviations include bulges (non-guanine nucleotides inserted within a G-tract) [44,45], vacancies (missing guanines in a G-tetrad) [46], and dynamic fluctuations like strand register shifts [47]. Rather than being mere structural defects, these features expand the structural diversity and functional repertoire of G4s.

Early foundational work [48] established that bulges generally reduce G4 stability, and they systematically investigated the effects of residue type, size, position, and number on G-quadruplex structure and stability. Their UV-melting experiments demonstrated that the number and size of bulges substantially diminish stability. This finding was later corroborated by Sarkar et al., who further attributed the destabilization of bulged G4s (buG4s) to an unfavorable enthalpic contribution during formation. They proposed that this arises from weakened base stacking interactions and compromised hydrogen bonding within the G-quartets due to the incorporated bulge [45].

The chemical identity of the bulge residue significantly influences its impact on stability. Mukundan and Phan observed a pronounced decrease in T_m values when the bulge was an adenine (A), suggesting that the lower stability might result from an increased solvation entropy associated with the exposure of its larger aromatic surface area [48]. Sarkar et al. confirmed this trend, reporting that the destabilization follows the order: $A > C \geq T$. They suggested that this discrepancy is likely due to the larger steric hindrance imposed by an adenine bulge, which more severely perturbs G4 folding [45].

The influence of bulge position appears to be context-dependent. In the UV-melting assays conducted by Mukundan and Phan, the positional effect of a single bulge was minimal, with T_m variations of only about 1 °C across different locations [48]. In contrast, Sarkar et al. found that the destabilizing effect is indeed modulated by position; for instance, bulges at position 1 in the first G-tract were less destabilizing than at other positions. They suggested that bulges at other sites may increase the inter-quartet distance, thereby weakening stacking interactions and leading to position-dependent stability variations [45]. Furthermore, Zhang et al. indicated that bulge position not only influences G4 topology but also the kinetics and pathway of folding. In guanine-vacancy-bearing G-quadruplexes, which are key folding intermediates, the location of the vacancy can strongly dictate the stability of the structure [16].

Despite their inherent destabilizing effect, several compensatory structural mechanisms can mitigate, or in some cases reverse, the negative impact of bulges. For example, a bulge can form a structured hairpin stem that stacks onto an outer G-tetrad. In such Quadruplex–Duplex hybrids, a larger, structured duplex bulge can confer greater thermostability than an unstructured loop of the same sequence [49]. Environmental factors can also play a compensatory role. Sarkar et al. showed that under molecular crowding conditions, the stability of both canonical and bulged G4s increases, thereby reducing the relative destabilization caused by the bulge [45]. This suggests that the cellular environments may effectively “rescue” certain bulged structures that would be marginally stable under standard dilute in vitro conditions.

2.3. Aligning the Evidence: G-Tract and Bulge Features in High-Throughput Sequencing Analysis

The influence of G-tract length on stability, as established by Rachwal et al. [40], is largely recapitulated by high-throughput sequencing data, both MM% from G4-Seq and quality score from G4-miner. In general, the inferred stability score increases with G-tract length, with a notable exception for three-layer G4s connected exclusively by 1-nt loops (Figure 2). The G4-miner algorithm successfully captured all documented trends, whereas G4-Seq accurately reflected the behavior of sequences with 1-nt loops but showed discrepancies in the case of 5-layer G4s when the loops exceeded 1 nt in length.

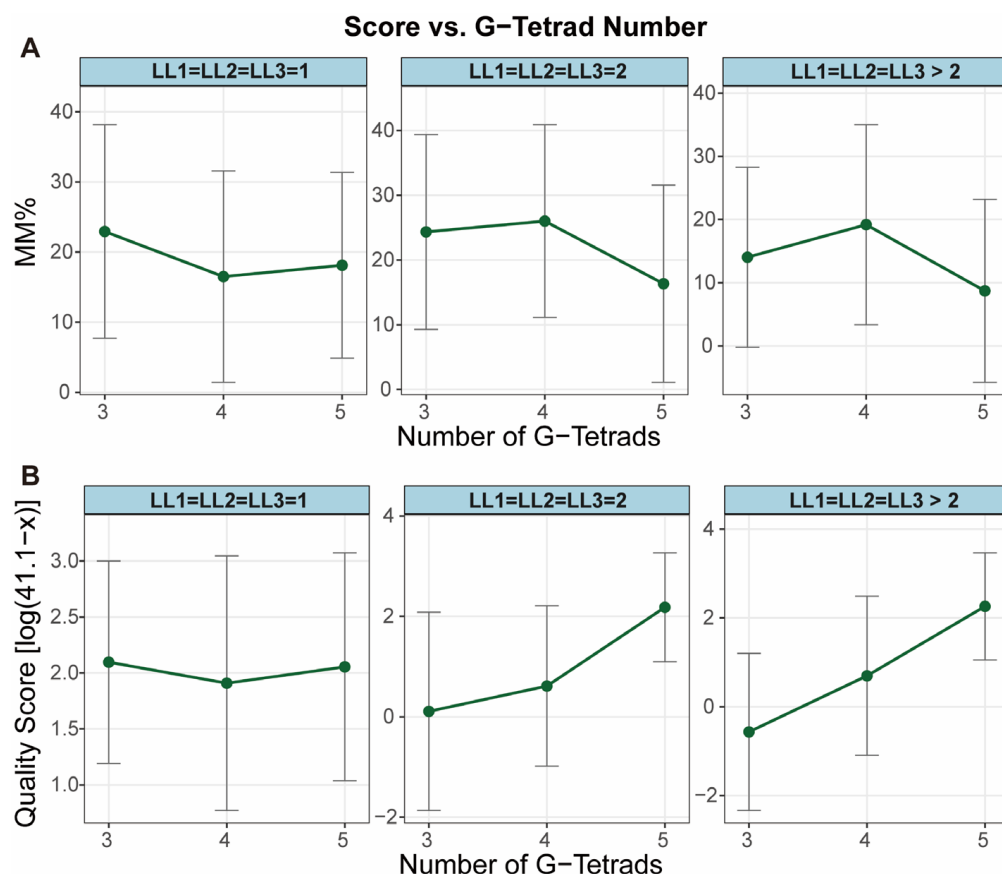


Figure 2. Dependence of G4 stability scores on the number of G-tetrad layers. Mean stability scores are plotted against the number of G-tetrad layers, with data stratified by loop length (1 nt, 2 nt, and >3 nt). **(A)** Mean MM% from G4-Seq. **(B)** Mean transformed quality score from G4-miner. The original G4-miner quality score (x) was transformed as $\ln(41.1 - x)$. Error bars represent standard deviations across samples. The terms LL1, LL2, LL3 represent lengths of loop 1, loop 2 and loop 3, respectively.

A possible explanation for the anomalous behavior of 5-layer G4s in G4-Seq may lie in the experimental design, which compares sequencing results under K^+ versus Na^+ stabilization conditions. Since 5-layer G4s exhibit inherently high stability in both ionic environments, the additional stabilizing effect conferred by K^+ is relatively minor. Consequently, the MM% values—reflecting the differential stability between ionic conditions—are smaller for these structures compared to their counterparts with fewer G-tetrad layers.

Both high-throughput methods are also capable of detecting the destabilizing influence of bulges on G4 stability. When compared to G4s with a bulge, G4s without bulges consistently yield higher stability scores in both G4-Seq and G4-miner analyses (Figure 3A,B). Furthermore, the impact of bulge size was correctly captured by both G4-Seq and G4-miner: stability levels decrease as the bulge size increases up to 5 nucleotides, confirming that bulges act as a general destabilizing factor within this size range. Intriguingly, for bulges larger than 5 nt, a positive correlation between stability metric and bulge size emerges. This reversal in trend likely reflects stabilization conferred by structured motifs, such as the formation of an internal duplex within the elongated bulge, which can compensate for the initial destabilization.

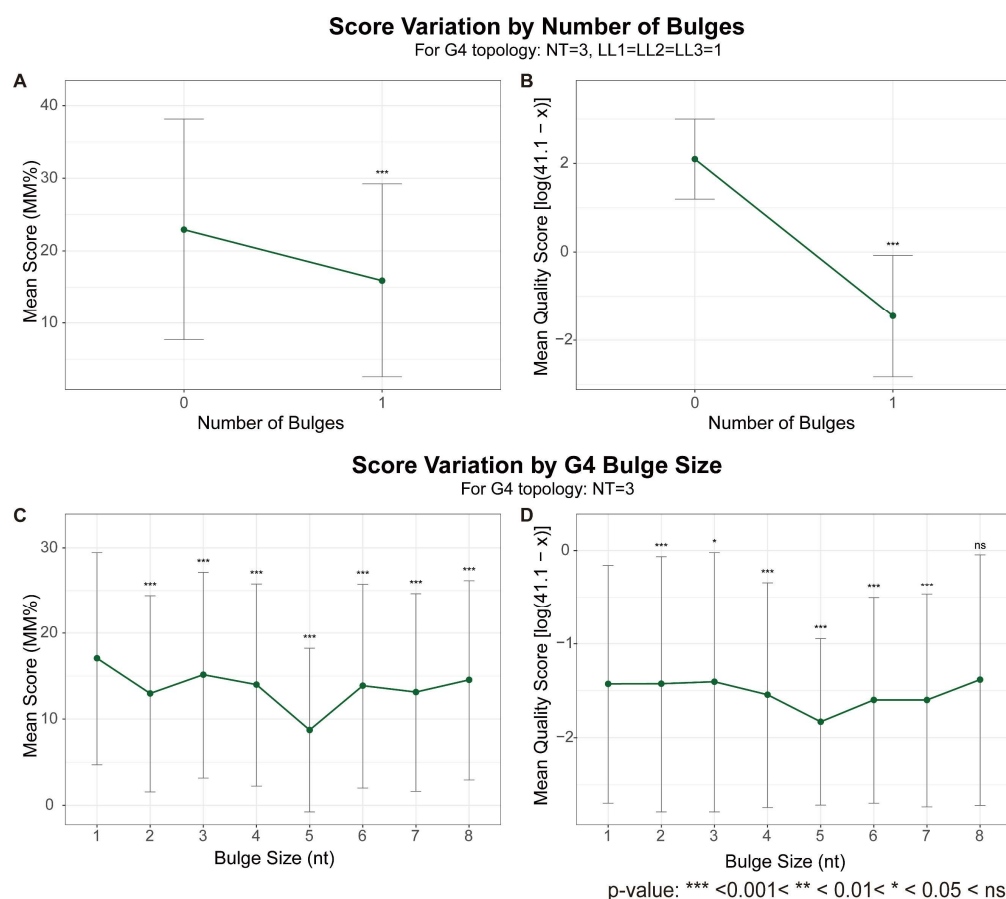


Figure 3. Influence of bulge presence and size on G4 stability scores. (A,B) Effect of bulge occurrence. Mean MM% from G4-Seq (A) and mean transformed quality score from G4-miner (B) are shown for three-layer G4s—with or without bulges—where all loops are 1 nt in length. (C,D) Effect of bulge size. Mean MM% (C) and transformed quality score (D) are shown for three-layer G4s with only 1-nt loops, grouped by bulge size (1 to 8 nt). The original G4-miner quality score (x) was transformed as $\ln(41.1 - x)$ in panels B and D. Error bars represent standard deviations. Significance was assessed using the Wilcoxon rank-sum test: in (A,B), between the two groups; in (C,D), between each bulge size and the first group (indicated by symbols above points). The terms LL1, LL2, LL3 represent lengths of loop 1, loop 2 and loop 3, respectively, and NT represents number of G-tetrads.

3. The Loops: The Impact of Length, Order and Nucleotide Composition on Stability

The loop regions connecting the G-tracts are critical structural components beyond the core G-tetrad. Their length, nucleotide composition, and sequential arrangement collectively form a fundamental “structural code” that dictates both the topology and thermostability of a G-quadruplex.

3.1. Loop Length: A General Inverse Correlation with Stability

Extensive biophysical evidence supports a fundamental principle: G4 thermostability exhibits a strong inverse correlation with the length of its connecting loops, considering both the total loop length and the length of any specific individual loop. As early as 2004, Phan et al. attributed the differential stability of two G4 structures from the MYC promoter to the size of their central loops, finding the variant with a two-residue loop to be more stable than the one with a six-residue loop [50]. Subsequently, Risitano and Fox demonstrated that the structure and stability of intramolecular quadruplexes are profoundly influenced by loop length and composition, using substitutions with non-nucleosidic linkers [51]. Moreover,

an example by Fernando et al. about the c-kit21 G4 showed that the addition of a single nucleotide in the third loop can reduce the T_m by a remarkable 10 °C [52].

This principle was quantitatively defined by Guédin et al., who established an empirical rule by a systematic study on G4 loop length: each additional nucleotide in the total loop length reduces the T_m by approximately 2 °C in a K^+ environment, corresponding to a loss in free energy (ΔG) of about 0.3 kcal/mol [53]. A broadly similar trend is observed in a Na^+ environment, although the linear relationship between T_m and loop length is less pronounced.

From a thermodynamic perspective, the destabilizing effect of long loops is primarily entropy-driven. A longer, more flexible loop possesses greater conformational entropy in the unfolded, single-stranded state. Upon folding into a G4 structure, this flexibility is constrained, leading to a significant entropic penalty. Hao et al. demonstrated that the longer the central loop, the greater this entropic cost, thereby exerting a more substantial destabilizing effect on the folded structure [54]. Although this general trend is robust, its impact is context-dependent. For example, a G4 structure containing two very short loops (e.g., 1-nt or 2-nt) can often tolerate a single, considerably longer third loop without a catastrophic loss of stability [53]. This observation suggests a synergistic effect among loops, whereby the high stability conferred by two short loops can partially compensate for the large entropic penalty imposed by one long loop.

Furthermore, loop length is a key determinant of G4 topology. Short loops, due to steric constraints, typically favor a parallel G-tract alignment [55]. In contrast, longer loops provide the flexibility necessary for the adoption of more complex antiparallel or hybrid topologies [33,41–43]. Additionally, the loop length may influence the formation of multimeric G4s [55]. A more detailed review about the influence of loops on G4 topology was provided by ref [36].

It is also noteworthy that under specific conditions, long loops can form stable secondary structures, such as hairpins. These can lead to the formation of Quadruplex–Duplex hybrids, which may mitigate, or in some cases even reverse, the intrinsic destabilizing effect typically associated with long loops [56–60].

3.2. Loop Permutation: The Critical Role of Sequential Order

The complexity of G4 sequence design is further exemplified by the phenomenon of loop permutation. Even when the G-tracts, total loop length, and nucleotide composition are held constant, merely altering the sequential order of the three loops can profoundly impact both the topology and thermostability of the resulting G4 structure [61–63].

The magnitude of this effect is substantial. A systematic investigation by Cheng et al. demonstrated that different loop permutations can alter the T_m by up to 17 °C—an effect comparable to that of modifying loop length or key base compositions by several nucleotides [62]. Agrawal et al. observed similar trends in their study of G4s within the VEGF promoter [64]. A central principle emerging from these studies, as proposed by Cheng et al., is the dominant role of the central (second) loop length in determining the final architecture. Sequences with a longer central loop exhibit a strong propensity to adopt stable non-parallel (hybrid or antiparallel) topologies, whereas those with a shorter central loop typically form parallel topologies, which are often less thermostable [62]. This finding is corroborated by Beniaminov et al., who showed that replacing TTA loops with 1-nt loops can shift the conformation of human telomeric G4s from a 3 + 1 hybrid type to a parallel form [65].

The significant influence of loop permutation underscores that loops are not merely passive linkers but play an active, guiding role in the G4 folding pathway, effectively functioning as a “topological code.” This effect cannot be fully explained by the intrinsic

thermodynamic properties of the loops alone; it must also involve the geometric constraints and kinetic choices encountered by the single strand during its folding process. Supporting this view, Chen et al. reported a sequence with a 4T central loop that forms a structure exhibiting cation-sensitive G-switch behavior, highlighting how loop arrangement can encode a regulatory potential within the folding landscape [63].

3.3. Loop Nucleotide Composition: Contradictions Depending on Context

Beyond loop length, the specific nucleotide composition of the loops adds a further layer of complexity to G4 stability. Based on a study of quadruplex-forming sequences from the promoter regions of various proto-oncogenes, Kumar and Maiti observed that these naturally occurring quadruplexes exhibit a wide range of thermodynamic stabilities (ΔG from 1.7 to 15.6 kcal/mol) and suggested that both loop length and composition collectively influence structure and thermodynamics, making it challenging to establish generalized correlations based on length alone [66].

Nevertheless, certain general principles can be discerned. The identity of the bases within the loops fine-tunes stability through specific interactions, including base stacking, hydrogen bonding, and steric effects [67–69]. Research by Guédin et al. indicated that the presence of adenine is equally unfavorable in either the first or the last loop [67]. A destabilizing effect of adenine was also observed by Nagatoishi et al. in a two-layer G4 system [70]. Furthermore, a subsequent study by Guédin et al. demonstrated that in a three-base loop, an adenine at the first position immediately following a G-tract exerts a significantly more destabilizing effect compared to other bases [71]. This effect has been corroborated in the thermodynamic study of adenine-to-thymine substitutions in the G4 loops [72].

However, a notable counterexample is found in the human telomeric repeat sequence (TTAGGG)_n. The TTA loops are evolutionarily conserved, and research by Tucker et al. has shown that in a Na⁺ environment, the adenines within these loops are critical for maintaining G4 stability. Systematic substitution of these adenines with thymines resulted in fluctuations in both melting temperature and unfolding enthalpy, indicating that adenine stabilizes the structure via stacking interactions with the adjacent G-tetrad and through the formation of intra-loop hydrogen bonds [73]. Further complicating the picture, Agrawal et al.'s study on VEGF promoter G4s revealed that G-to-T mutations of two residues in the central loops did not significantly alter T_m [64].

One possible explanation for these apparent contradictions could be: the contribution of any single base to stability might be dependent on its specific three-dimensional structural context. In the study of Tippana et al., the larger purine ring of adenine may introduce steric hindrance within the loop regions of parallel conformations, thereby disrupting optimal structural packing [43]. In the MD simulations by Zhu et al., the adenine in the TTA loop is precisely positioned to engage in favorable π - π stacking interactions with the terminal G-tetrad, thereby providing additional enthalpic stabilization [74]. Suggested by Lenarčič Živković et al., specific interactions mediated by loop adenines can drive significant structural changes [75], and thus influence the topology of G4s and act as a structural switch that narrows a broad range of putative folds into an antiparallel conformation.

This comparison underscores that any simple rule regarding base composition must be considered in conjunction with the likely topological context. It also implies that predicting the effect of loop composition may first require predicting the most probable topology—a feature that is itself influenced by loop length and arrangement. This interdependence creates a complex, multi-parameter optimization problem in G4 structure prediction [76].

3.4. Aligning the Evidence: Interpreting Loop Features in High-Throughput Data

The negative correlation between thermostability of a G4 structure and the total loop length is also discernible in both MM% from G4-Seq and quality score from G4-miner (Figure 4). Despite inherent data variability, the trends are clear: both methods show a decrease in indicated stability with increasing total loop length, consistent with the principles established by low-throughput biophysical techniques [50–53].

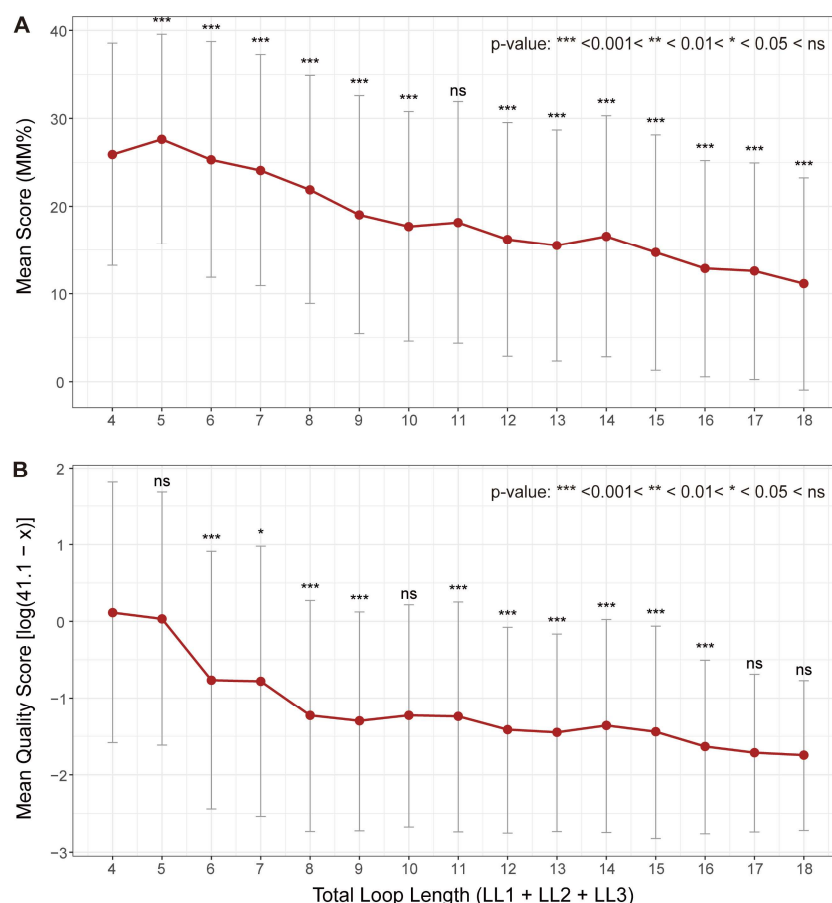


Figure 4. Variation in G4 stability scores with total loop length. Mean stability scores derived from (A) MM% in G4-Seq and (B) transformed quality scores from G4-miner are shown for G4s grouped by total loop length. Each point represents the mean of 1000 resampled sequences. To ensure that higher values correspond to greater stability—consistent with the interpretation of MM%—the original G4-miner quality score (x) was transformed as $\ln(41.1 - x)$. Error bars indicate standard deviations across loop length groups. Significance between adjacent loop length groups was assessed using the Wilcoxon rank-sum test; symbols above data points denote the statistical significance relative to the group immediately to the left. The terms LL1, LL2, LL3 represent lengths of loop 1, loop 2 and loop 3, respectively.

We further compared the influence of loop permutation on stability scores from high-throughput methods with the findings of Cheng et al. [62]. For G4s with three unequal loop lengths (denoted as $x \neq y \neq z$), similar trends were partially observed (Figure 5A,B). For instance, the s-l-m (short-long-medium) and l-s-m (long-short-medium) permutations, corresponding to groups A2 and A5 in Cheng et al.'s study, generally exhibited high stability scores. Conversely, the m-l-s (medium-long-short) permutation (group A4 in Cheng et al.) tended to yield lower scores, indicating destabilization of the G4 structures. However, discrepancies were noted, such as unexpectedly high scores for the l-m-s group in G4-Seq and the s-m-l group in G4-miner. For G4s with loop length combinations $x = y > z$ and $x = y < z$, the scores from G4-Seq were inconsistent with Cheng et al.'s experimental

results [62] (Figure 5C–F). In contrast, G4-miner scores better captured the expected trends; for example, the x - y - x permutation (group A2 in Cheng et al.) correctly showed the lowest or highest stability levels for the $x = y > z$ and $x = y < z$ combinations, respectively.

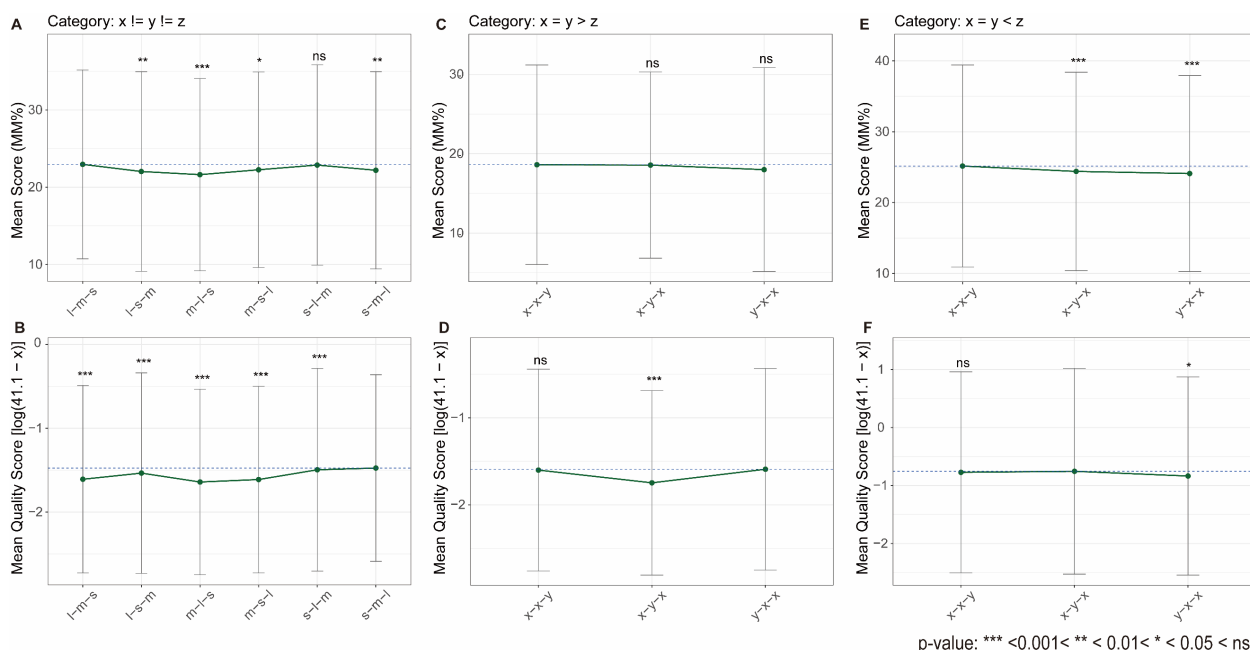


Figure 5. G4 stability score variation with loop permutation. Stability scores are shown for G4s grouped by different loop length permutations: (A,B) three unequal loop lengths ($x \neq y \neq z$), (C,D) two long loops of equal length and one shorter loop ($x = y > z$), and (E,F) two short loops of equal length and one longer loop ($x = y < z$). The upper panels (A,C,E) show the mean MM% from G4-Seq, while the lower panels (B,D,F) show the mean transformed quality score from G4-miner. The original G4-miner quality score (x) was transformed as $\ln(41.1 - x)$. Error bars represent standard deviations. Statistical significance was assessed using the Wilcoxon rank-sum test, comparing each group with the highest mean stability score within each panel.

It should be noted that the original biophysical observations are highly context-dependent, as trends in T_m across loop permutations are not universally consistent (Table 1 in [62]), which may partly explain the variability in high-throughput data. Moreover, the magnitude of stability score differences attributable to loop permutation in sequencing-based methods is considerably attenuated compared to the $\sim 17^\circ\text{C}$ T_m variations reported in biophysical studies, indicating that genomic assays are generally less sensitive to such structural nuances. This attenuated sensitivity aligns with earlier analyses based on high-throughput data and molecular dynamics simulations [77], which concluded that while total loop length exerts a clear negative effect on G4 stability, the specific distribution of loop lengths has minimal influence. This inherent limitation of sequencing-based approaches may explain not only the discrepancies with biophysical measurements but also the divergent outputs between G4-Seq and G4-miner, as the topological and stabilizing effects of K^+ [36] employed in G4-Seq likely dominate the sequencing signal, thereby overshadowing the weaker modulation induced by loop permutation.

Regarding loop composition, we first tested the ability of these methods to detect the effect of A-to-T mutations in the evolutionarily conserved TTA loops of telomeric G4s (Figure 6A,B). Both G4-Seq and G4-miner correctly reflected the destabilizing influence of these mutations, as reported by Tucker et al. [73], with A-to-T mutant groups showing consistently lower stability scores than the wild-type. This successful detection is likely facilitated by the well-defined highly characteristic sequence and structure of telomeric G4s.

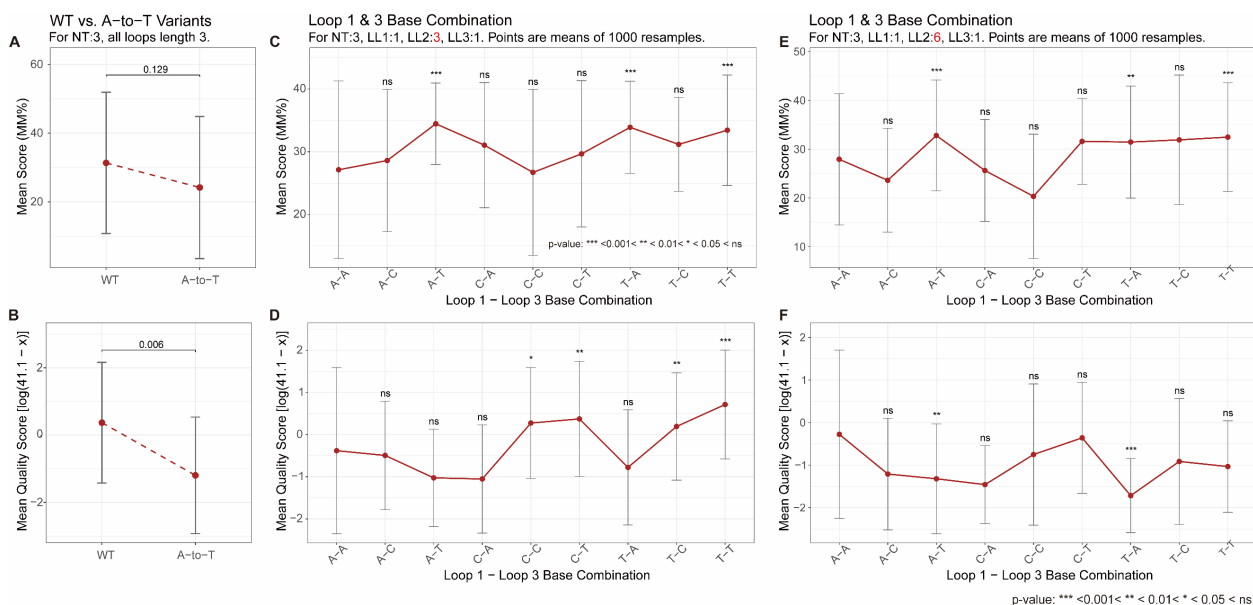


Figure 6. G4 stability score variation with loop composition. **(A,B)** Influence of adenine in human telomeric DNA repeats. Mean MM% from G4-Seq **(A)** and mean transformed quality score from G4-miner **(B)** are shown for wild-type (WT) telomeric G4 sequences ((GGGTTA)_n) and sequences containing A-to-T variations. **(C–F)** Influence of non-guanine base composition in 1-nt first and third loops. Mean MM% **(C,E)** and transformed quality scores **(D,F)** are shown for three-layer G4s with 1-3-1 **(C,D)** and 1-6-1 **(E,F)** loop length configurations. All nine possible single-base combinations (A, C, or T in first loop × A, C, or T in third loop) were tested. The original G4-miner quality score (x) was transformed as $\ln(41.1 - x)$ in panels **(B,D,F)**. Error bars represent standard deviations. Statistical significance was assessed using Wilcoxon rank-sum test: in **(A,B)**, between WT and variant groups; in **(C–F)**, between each combination and the reference A-A combination (indicated by symbols above data points). The terms LL1, LL2, LL3 represent lengths of loop 1, loop 2 and loop 3, respectively, and NT represents number of G-tetrads.

In contrast, both methods struggled to consistently recapitulate the base-specific effects observed in the 1-nt first and third loops by Guédin et al. [67]. When the central loop was 3-nt long (Figure 6C,D), G4-Seq detected the expected stability trend ($A < C < T$) only when the first loop was adenine, and correctly identified the destabilization by the A-A combination, but failed for other base combinations. G4-miner captured the $A < C < T$ trend when the first loop was cytosine or thymine, but was insensitive when the first loop was adenine. With a 6-nt central loop (Figure 6E,F), G4-miner detected the destabilizing effect of adenine in specific comparative sets (e.g., C-A vs. C-C vs. C-T; T-A vs. T-C vs. T-T). Although both sequencing methods exhibit sensitivity to loop base composition, the resulting patterns are only partially consistent with biophysical observations. Given the well-documented, often conflicting results about loop composition effects in controlled biophysical experiments [43,64,67,70–75], it is plausible that the influence of loop composition is inherently variable and strongly context-dependent, and lacks consistent governing rules. Therefore, it is unsurprising that high-throughput sequencing methods, which integrate myriad contextual variables, face challenges in robustly detecting and reproducing these subtle, sequence-specific trends.

4. The Influence of Flanking Regions on G4 Structure and Stability

Within the genomic context, G4-forming sequences are invariably embedded within longer DNA strands, flanked by upstream and downstream nucleotides. These flanking regions, frequently omitted from in vitro studies, play a critical and multifaceted role in modulating G4 structure and stability.

4.1. Flanking Nucleotides as Modulators of G4 Folding

The addition of flanking nucleotides to a core G4 sequence can impart stabilizing effects. For instance, in studies of the MYC promoter G4 under physiological potassium concentrations (20 mM K⁺), isoforms retaining their natural flanking sequences exhibited T_m 2–3 °C higher than truncated core-only constructs [61]. This stabilization is often attributed to the flanking bases forming protective “capping” structures via stacking interactions with the terminal G-tetrads [78], or through specific interactions with the loop regions [79,80]. Subsequent research further suggested that such capping structures may facilitate the coordination of potassium cations, thereby enhancing stability [61].

Conversely, other studies report a destabilizing influence. A systematic investigation by Chen et al. on 371 distinct intramolecular G4s revealed that the addition of short 2-thymine flanking segments generally lowered T_m values. While the effect was typically modest, a dramatic decrease of 16.5 °C was observed in one case, challenging the universality of a stabilizing role for short flanking sequences [81]. Similarly, Bugaut and Alberti found that TTA flanking sequences on human telomeric G4s were destabilizing, with a more pronounced effect from the 5′ flank [82]. Zhang et al. also reported that 5′-flanking residues could prevent the formation of certain two-layer G4 conformations [83].

Beyond modulating stability, flanking sequences exert a profound influence on G4 topology. Chen et al. demonstrated that adding nucleotides, particularly at the 5′ end, strongly favors the formation of parallel topologies, often overriding the intrinsic propensity of a core sequence to form hybrid or antiparallel folds [81]. They proposed a mechanistic explanation: a 5′-end flanking nucleotide can disrupt a specific hydrogen bond with a 5′-terminal *syn* guanine, thereby destabilizing all folds that require this conformation and effectively selecting for parallel structures.

In addition to these general effects, specific flanking sequences can engender unique structural consequences. For example, Pavc et al. showed that a 5′ GC sequence can mediate G4 dimerization through the formation of an inter-quadruplex GCGC-tetrad, leading to stable higher-order assemblies [84]. Conversely, Do and Phan (2012) found that certain 5′ flanking residues could disrupt the stacking interactions necessary for dimerization, favoring monomeric G4 formation instead [85].

The recognition of this “flanking effect” necessitates a critical re-evaluation of the vast body of biophysical data acquired using truncated, minimal G4 constructs. While numerous NMR and crystal structures of such core sequences have revealed a rich polymorphism encompassing hybrid and antiparallel folds [86,87], the presence of a free 5′-terminus may itself be an experimental artifact that biases against parallel topologies. Consequently, the full spectrum of G4 structural diversity observed *in vitro* may not accurately reflect the *in vivo* landscape, where G4s are perpetually flanked by genomic sequences. From this perspective, high-throughput genomic methods, which analyze long DNA fragments with native flanks, might provide a more biologically representative view of predominant G4 topologies than some high-resolution structural studies of isolated cores.

4.2. Aligning the Evidence: The Flanking Effect in a Genomic Context

Given that flanking segments are an inherent feature of most genomic G4s, the influence of their nucleotide composition on G4 stability must be considered. Integrating evidence from Hatzakis et al. [61] and Chen et al. [81] reveals that specific flanking sequences can exert distinct effects on stability; for instance, 5′-TGA and 3′-GAA flanking sequences appear to stabilize G4s, whereas 5′- and 3′-TT sequences tend to be destabilizing. Although derived from different experimental systems, this distinction is consistently reflected in high-throughput data. Both the MM% values from G4-Seq and the transformed quality scores from G4-miner show clear differences between these groups: G4s flanked

by TT sequences consistently exhibit significantly lower stability scores than those with stabilizing flanking sequences at both the 5' and 3' ends (Figure 7).

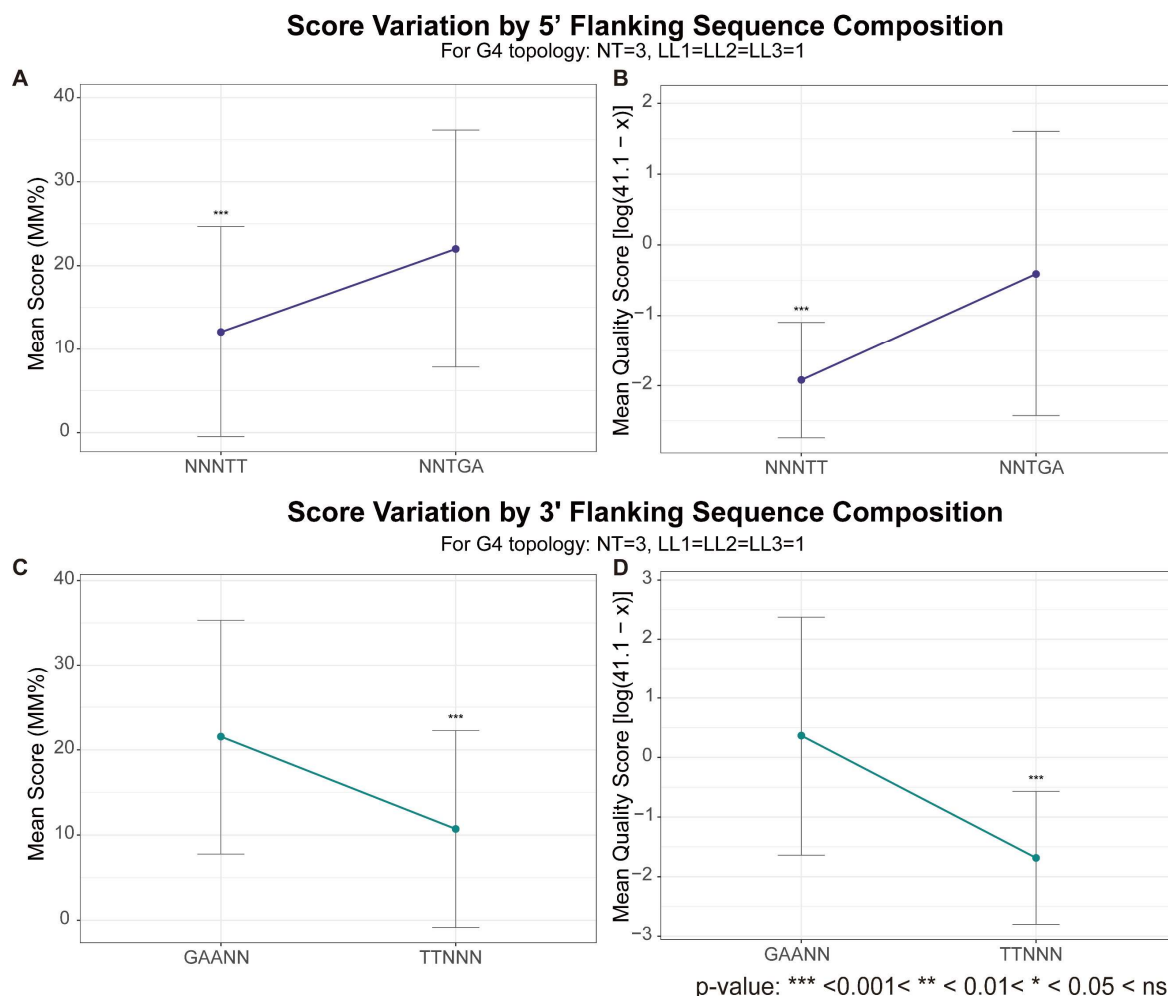


Figure 7. Influence of flanking sequences on G4 stability scores. (A,B) Effect of 5'-flanking sequences. Mean MM% from G4-Seq (A) and mean transformed quality score from G4-miner (B) are shown for three-layer G4s (all loops 1 nt) flanked at the 5' end by either NNNTT or NNTGA. (C,D) Effect of 3'-flanking sequences. Mean MM% (C) and transformed quality score (D) are shown for three-layer G4s (all loops 1 nt) flanked at the 3' end by either GAANN or TTNNN. The original G4-miner quality score (x) was transformed as $\ln(41.1 - x)$ in panels (B,D). Error bars represent standard deviations. Statistical significance between flanking sequence groups was assessed using the Wilcoxon rank-sum test. The terms LL1, LL2, LL3 represent lengths of loop 1, loop 2 and loop 3, respectively, and NT represents number of G-tetrads.

5. Conclusion and Outlook

5.1. Toward a Unified Understanding of Sequence-Stability Relationships

This review has systematically examined the intrinsic sequence features governing G-quadruplex (G4) thermostability and has endeavored to bridge observations from high-precision biophysical methods with those from high-throughput genomic assays. Overall, the two methodological paradigms yield consistent conclusions regarding the influence of most sequence features. For those trends showing incomplete agreement—such as the effects of loop permutation and base composition—the discrepancies primarily arise from their strongly context-dependent nature. In characterizing these particular features, both methodological paradigms exhibit instability in their conclusions, reflecting the inherent complexity and condition-specific variability of these structural determinants. The collec-

tive evidence allows for the synthesis of a set of generalized, albeit context-dependent, rules (Table 1).

Table 1. Alignment of evidence from high-precision biophysical experiments and high-throughput sequencing data.

Sequence Feature	Patterns from Biophysical Studies	Consistency with High-Throughput Data	Key Observations from High-Throughput Data
G-Tract Length	Stability generally increases with length, though non-monotonically; anomalous stability observed for 3-layer G4s with 1-nt loops [40]	Generally consistent (Both methods, Figure 2).	1. G4-miner capturing all trends accurately 2. G4-Seq showing discrepancies for 5-layer G4s with longer loops
Bulges	Bulges are generally destabilizing; the effect is size-dependent [45,48]; compensatory stabilization was reported [49].	Consistent (Both methods, Figure 3)	1. Both methods detecting destabilization by bulges 2. Negative correlation for bulges < 5 nt suggesting destabilizing effect of bulge size 3. Positive correlation for bulges > 5 nt suggesting compensatory structural mechanisms
Total Loop Length	Strong inverse correlation with stability was observed [53].	Consistent (Both methods, Figure 4)	1. Robustly captured by both G4-Seq and G4-miner
Loop Permutation	Significant impact on stability and topology was observed; central loop length plays dominant role [62,64]	Partially Consistent (Figure 5)	1. Partial agreement for unequal loop lengths $x \neq y \neq z$ 2. G4-miner better capturing loop combinations $x = y \neq z$ 3. Reflecting context-dependency of bio-physical observations and attenuated sensitivity of sequencing-based methods to loop order
Loop Base Composition	The effect is highly context-dependent; adenine can be stabilizing or destabilizing depending on structural context [67,70–73]	Partially Consistent (Figure 6)	1. Both methods exhibiting sensitivity to loop composition 2. Both methods detecting stabilizing effect of adenine in telomeric G4s 3. Limited consistency for G4s in more diverse sequence contexts
Flanking Regions	The effect is composition-sensitive; specific flanking sequences differentially modulate G4 stability, with 5'-TGA/3'-GAA conferring stabilization [61] and 5'-TT/3'-TT producing destabilization [81].	Consistent (Both methods, Figure 7)	1. Both methods confirming differential effects of specific flanking motifs

The length and integrity of the G-tracts establish the foundational stability of the G-quadruplex. Generally, the correlation between G-tract length and G4 stability is consistently observed across methodological paradigms, though it does not follow a simple monotonic function. Furthermore, the destabilizing effect of bulges—including the influence of bulge size—is also captured by both experimental approaches. However, the

influence is modulated by additional complex factors, such as the conformational strain imposed by short loops, which underlies the “anomalous stability” of certain configurations, and the capacity of long bulges to form stabilizing secondary structures.

The well-established inverse correlation between loop length and G4 stability is consistently observed across both biophysical and high-throughput methodological paradigms. In contrast, the effects of loop permutation show only partial consistency between these approaches, likely due to the intrinsic context-dependency of this structural feature combined with the attenuated sensitivity of sequencing-based methods to loop rearrangement. Furthermore, the influence of loop nucleotide composition demonstrates marked context-dependence. While both methodological frameworks detect the stabilizing role of adenine in canonical telomeric G4 structures, they fail to establish clear, generalizable rules regarding base-specific effects in more diverse sequence contexts.

Evidence suggests that the base composition of flanking sequences can modulate stability, with certain motifs showing consistent effects across experimental approaches. Critically, as an inevitable component of genomic G4s, flanking sequences appear to exert a structural influence that favors parallel topologies, which should not be ignored especially when discussing G4 conformation and stability under genomic context.

5.2. Toward Quantitative and Predictive Models

The rules summarized in this review remain largely qualitative or semi-quantitative. For instance, we can reasonably predict that a sequence featuring short loops, continuous G3-tracts, no bulges, and specific 5' flanking bases will form a highly stable parallel G4, yielding a strong signal in G4-Seq assays. Conversely, a sequence with long loops, a central bulge, and short G2-tracts is likely to adopt an unstable, transient structure associated with a weak G4-Seq signal.

However, translating these qualitative relationships into a fully quantitative and predictive framework represents a critical objective for future research. Achieving this will require innovative experimental strategies, such as the construction of a synthetic oligonucleotide library comprising thousands of rationally designed G4 sequences that systematically vary key structural features. Based on parallel analyses using both high-throughput biophysical profiling and next-generation sequencing-based methods, such a design would enable the direct mapping and calibration of sequencing-derived metrics—such as mismatch rates or quality scores—against precisely measured thermodynamic parameters. Such effort would pave the way for establishing a foundational “conversion table” or computational model capable of bridging observations across methodological scales.

5.3. Unresolved Questions and Concluding Remarks

Despite considerable progress in deciphering the principles of G4 stability, several fundamental questions remain open. What is the precise quantitative relationship between the kinetics of polymerase pausing—as detected in high-throughput assays—and the thermodynamic stability of the underlying G4 structure? To what extent do the numerous G4-binding proteins in the cellular milieu modulate or override these intrinsic sequence-stability rules? Furthermore, do RNA G4s and DNA–RNA hybrid G4s follow the same structural principles as their DNA counterparts?

From a methodological perspective, significant limitations persist. While high-throughput approaches can statistically capture the influence of major structural features on stability, their analytical power is constrained. For instance, when controlling for variables—such as examining perfect three-layer G4s with 1-nt loops—many specific sequence combinations are either absent from the human genome or occur too infrequently

to support robust statistical analysis. Consequently, this review has necessarily focused on more prevalent sequence features. Similarly, the biophysical studies we have integrated are predominantly those that systematically and batch-test the effects of specific sequence variations.

In conclusion, the thermostability of a G-quadruplex is not dictated by any single sequence element but emerges from the complex interplay between its G-tract core, loop architecture, and flanking environment. By systematically deconstructing these features and understanding their roles in analytical methods of different precision and throughput, we are beginning to align experimental evidence from complementary domains. This synthesis of knowledge is a crucial prerequisite for unlocking the ability to predict—on a genome-wide scale—where functionally significant G4s reside, how stable they are, and what roles they fulfill in cellular physiology and disease.

Author Contributions: Conceptualization, K.X. and J.T.; methodology, K.X.; validation, K.X. and J.F.; formal analysis, K.X.; investigation, K.X. and R.Z.; writing—original draft preparation, K.X.; writing—review and editing, J.F., R.Z. and J.T.; visualization, K.X.; supervision, J.T.; project administration, J.T.; funding acquisition, K.X. and J.T. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (No. 62002060) and the Fundamental Research Funds for the Central Universities of China (No. 2242023K5005).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available in Gene Expression Omnibus at <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE63874> (accessed on 1 July 2025) and <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE159307> (accessed on 1 July 2025), reference number 8 and 26 respectively.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Data Sources and Processing

The G4-Seq data for the Na⁺ vs. K⁺ conditions were obtained from Chambers et al. [8], and the G4-miner data were sourced from Tu et al. [30]. Both datasets were processed using the respective workflows described in the original publications. Potential G4 (pG4) loci were identified using pqsfinder [9], and downloaded from https://pqsfinder.fi.muni.cz/hub/hg19/pqsfinder_hg19_gff.tar.gz (accessed on 10 September 2025). For each pG4 locus, stability levels were assigned based on the maximum Mismatch Percentage (MM%) values from G4-Seq or averaged sequencing quality scores from G4-miner, calculated within a 50 bp window centered on the locus and restricted to the same strand.

References

1. Varshney, D.; Spiegel, J.; Zyner, K.; Tannahill, D.; Balasubramanian, S. The Regulation and Functions of DNA and RNA G-Quadruplexes. *Nat. Rev. Mol. Cell Biol.* **2020**, *21*, 459–474. [CrossRef]
2. Chaires, J.B. Human Telomeric G-quadruplex: Thermodynamic and Kinetic Studies of Telomeric Quadruplex Stability. *FEBS J.* **2010**, *277*, 1098–1106. [CrossRef] [PubMed]
3. Lane, A.N.; Chaires, J.B.; Gray, R.D.; Trent, J.O. Stability and Kinetics of G-Quadruplex Structures. *Nucleic Acids Res.* **2008**, *36*, 5482–5515. [CrossRef] [PubMed]
4. Bhattacharyya, D.; Mirihana Arachchilage, G.; Basu, S. Metal Cations in G-Quadruplex Folding and Stability. *Front. Chem.* **2016**, *4*, 38. [CrossRef]

5. Ambrus, A.; Chen, D.; Dai, J.; Bialis, T.; Jones, R.A.; Yang, D. Human Telomeric Sequence Forms a Hybrid-Type Intramolecular G-Quadruplex Structure with Mixed Parallel/Antiparallel Strands in Potassium Solution. *Nucleic Acids Res.* **2006**, *34*, 2723–2735. [[CrossRef](#)] [[PubMed](#)]
6. Qin, Y.; Fortin, J.S.; Tye, D.; Gleason-Guzman, M.; Brooks, T.A.; Hurley, L.H. Molecular Cloning of the Human Platelet-Derived Growth Factor Receptor β (PDGFR- β) Promoter and Drug Targeting of the G-Quadruplex-Forming Region to Repress PDGFR- β Expression. *Biochemistry* **2010**, *49*, 4208–4219. [[CrossRef](#)]
7. Šponer, J.; Bussi, G.; Stadlbauer, P.; Kührová, P.; Banáš, P.; Islam, B.; Haider, S.; Neidle, S.; Otyepka, M. Folding of Guanine Quadruplex Molecules—Funnel-like Mechanism or Kinetic Partitioning? An Overview from MD Simulation Studies. *Biochim. Biophys. Acta (BBA)-Gen. Subj.* **2017**, *1861*, 1246–1263. [[CrossRef](#)]
8. Chambers, V.S.; Marsico, G.; Boutell, J.M.; Di Antonio, M.; Smith, G.P.; Balasubramanian, S. High-Throughput Sequencing of DNA G-Quadruplex Structures in the Human Genome. *Nat. Biotechnol.* **2015**, *33*, 877–881. [[CrossRef](#)]
9. Hon, J.; Martínek, T.; Zendulka, J.; Lexa, M. Pqsfinder: An Exhaustive and Imperfection-Tolerant Search Tool for Potential Quadruplex-Forming Sequences in R. *Bioinformatics* **2017**, *33*, 3373–3379. [[CrossRef](#)]
10. Stegle, O.; Payet, L.; Mergny, J.-L.; MacKay, D.J.C.; Huppert, J.L. Predicting and Understanding the Stability of G-Quadruplexes. *Bioinformatics* **2009**, *25*, i374–i1382. [[CrossRef](#)]
11. Han, H.; Hurley, L.H.; Salazar, M. A DNA Polymerase Stop Assay for G-Quadruplex-Interactive Compounds. *Nucleic Acids Res.* **1999**, *27*, 537–542. [[CrossRef](#)]
12. Guiblet, W.M.; DeGiorgio, M.; Cheng, X.; Chiaromonte, F.; Eckert, K.A.; Huang, Y.-F.; Makova, K.D. Selection and Thermostability Suggest G-Quadruplexes Are Novel Functional Elements of the Human Genome. *Genome Res.* **2021**, *31*, 1136–1149. [[CrossRef](#)] [[PubMed](#)]
13. Prorok, P.; Artufel, M.; Aze, A.; Coulombe, P.; Peiffer, I.; Lacroix, L.; Guédin, A.; Mergny, J.-L.; Damaschke, J.; Schepers, A.; et al. Involvement of G-Quadruplex Regions in Mammalian Replication Origin Activity. *Nat. Commun.* **2019**, *10*, 3274. [[CrossRef](#)] [[PubMed](#)]
14. Ogasawara, S. Transcription Driven by Reversible Photocontrol of Hyperstable G-Quadruplexes. *ACS Synth. Biol.* **2018**, *7*, 2507–2513. [[CrossRef](#)] [[PubMed](#)]
15. Xiao, K.; Zhang, R.; Tao, T.; Shu, H.; Huang, H.; Sun, X.; Tu, J. Stability Matters: Revealing Causal Roles of G-Quadruplexes (G4s) in Regulation of Chromatin and Transcription. *Genes* **2025**, *16*, 1231. [[CrossRef](#)]
16. Zhang, Y.; Cheng, Y.; Chen, J.; Zheng, K.; You, H. Mechanical Diversity and Folding Intermediates of Parallel-Stranded G-Quadruplexes with a Bulge. *Nucleic Acids Res.* **2021**, *49*, 7179–7188. [[CrossRef](#)]
17. You, H.; Zeng, X.; Xu, Y.; Lim, C.J.; Efremov, A.K.; Phan, A.T.; Yan, J. Dynamics and Stability of Polymorphic Human Telomeric G-Quadruplex under Tension. *Nucleic Acids Res.* **2014**, *42*, 8789–8795. [[CrossRef](#)]
18. Balasubramanian, S.; Hurley, L.H.; Neidle, S. Targeting G-Quadruplexes in Gene Promoters: A Novel Anticancer Strategy? *Nat. Rev. Drug Discov.* **2011**, *10*, 261–275. [[CrossRef](#)]
19. Onel, B.; Lin, C.; Yang, D. DNA G-Quadruplex and Its Potential as Anticancer Drug Target. *Sci. China Chem.* **2014**, *57*, 1605–1614. [[CrossRef](#)]
20. Teng, F.-Y.; Jiang, Z.-Z.; Guo, M.; Tan, X.-Z.; Chen, F.; Xi, X.-G.; Xu, Y. G-Quadruplex DNA: A Novel Target for Drug Design. *Cell. Mol. Life Sci.* **2021**, *78*, 6557–6583. [[CrossRef](#)]
21. Brown, R.V.; Wang, T.; Chappeta, V.R.; Wu, G.; Onel, B.; Chawla, R.; Quijada, H.; Camp, S.M.; Chiang, E.T.; Lassiter, Q.R.; et al. The Consequences of Overlapping G-Quadruplexes and i-Motifs in the Platelet-Derived Growth Factor Receptor β Core Promoter Nuclease Hypersensitive Element Can Explain the Unexpected Effects of Mutations and Provide Opportunities for Selective Targeting of Both Structures by Small Molecules to Downregulate Gene Expression. *J. Am. Chem. Soc.* **2017**, *139*, 7456–7475. [[CrossRef](#)]
22. Puig Lombardi, E.; Holmes, A.; Verga, D.; Teulade-Fichou, M.-P.; Nicolas, A.; Londoño-Vallejo, A. Thermodynamically Stable and Genetically Unstable G-Quadruplexes are Depleted in Genomes across Species. *Nucleic Acids Res.* **2019**, *47*, 6098–6113. [[CrossRef](#)] [[PubMed](#)]
23. Mergny, J.-L.; Phan, A.-T.; Lacroix, L. Following G-quartet Formation by UV-spectroscopy. *FEBS Lett.* **1998**, *435*, 74–78. [[CrossRef](#)]
24. Mergny, J.; Lacroix, L. UV Melting of G-Quadruplexes. *CP Nucleic Acid. Chem.* **2009**, *37*, 17.1.1–17.1.15. [[CrossRef](#)]
25. Rachwal, P.A.; Fox, K.R. Quadruplex Melting. *Methods* **2007**, *43*, 291–301. [[CrossRef](#)]
26. Haq, I.; Chowdhry, B.Z.; Jenkins, T.C. Calorimetric Techniques in the Study of High-Order DNA-Drug Interactions. In *Methods in Enzymology; Drug-Nucleic Acid Interactions*; Academic Press: Cambridge, MA, USA, 2001; Volume 340, pp. 109–149.
27. De Cian, A.; Guittat, L.; Kaiser, M.; Saccà, B.; Amrane, S.; Bourdoncle, A.; Alberti, P.; Teulade-Fichou, M.-P.; Lacroix, L.; Mergny, J.-L. Fluorescence-Based Melting Assays for Studying Quadruplex Ligands. *Methods* **2007**, *42*, 183–195. [[CrossRef](#)] [[PubMed](#)]
28. Wang, S.; Saneyoshi, H.; Xu, P.; Oguri, N.; Yamashita, A.; Xu, Y. Manipulating DNA and RNA Structures via Click-to-Release Caged Nucleic Acids for Biological and Biomedical Applications. *Nucleic Acids Res.* **2025**, *53*, gkaf571. [[CrossRef](#)] [[PubMed](#)]

29. Marsico, G.; Chambers, V.S.; Sahakyan, A.B.; McCauley, P.; Boutell, J.M.; Antonio, M.D.; Balasubramanian, S. Whole Genome Experimental Maps of DNA G-Quadruplexes in Multiple Species. *Nucleic Acids Res.* **2019**, *47*, 3862–3874. [\[CrossRef\]](#)
30. Tu, J.; Duan, M.; Liu, W.; Lu, N.; Zhou, Y.; Sun, X.; Lu, Z. Direct Genome-Wide Identification of G-Quadruplex Structures by Whole-Genome Resequencing. *Nat. Commun.* **2021**, *12*, 6014. [\[CrossRef\]](#)
31. Wang, L.; Wang, Q.-M.; Wang, Y.-R.; Xi, X.-G.; Hou, X.-M. DNA-Unwinding Activity of *Saccharomyces Cerevisiae* Pif1 is Modulated by Thermal Stability, Folding Conformation, and Loop Lengths of G-Quadruplex DNA. *J. Biol. Chem.* **2018**, *293*, 18504–18513. [\[CrossRef\]](#)
32. Sahakyan, A.B.; Chambers, V.S.; Marsico, G.; Santner, T.; Di Antonio, M.; Balasubramanian, S. Machine Learning Model for Sequence-Driven DNA G-Quadruplex Formation. *Sci. Rep.* **2017**, *7*, 14535. [\[CrossRef\]](#)
33. Arora, A.; Maiti, S. Stability and Molecular Recognition of Quadruplexes with Different Loop Length in the Absence and Presence of Molecular Crowding Agents. *J. Phys. Chem. B* **2009**, *113*, 8784–8792. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Verdian Doghaei, A.; Housaindokht, M.R.; Bozorgmehr, M.R. Molecular Crowding Effects on Conformation and Stability of G-Quadruplex DNA Structure: Insights from Molecular Dynamics Simulation. *J. Theor. Biol.* **2015**, *364*, 103–112. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Aznauryan, M.; Birkedal, V. Dynamics of G-Quadruplex Formation under Molecular Crowding. *J. Phys. Chem. Lett.* **2023**, *14*, 10354–10360. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Jana, J.; Weisz, K. Thermodynamic Stability of G-Quadruplexes: Impact of Sequence and Environment. *ChemBioChem* **2021**, *22*, 2848–2856. [\[CrossRef\]](#)
37. Nishio, M.; Tsukakoshi, K.; Ikebukuro, K. G-Quadruplex: Flexible Conformational Changes by Cations, pH, Crowding and Its Applications to Biosensing. *Biosens. Bioelectron.* **2021**, *178*, 113030. [\[CrossRef\]](#)
38. Zhong, H.-S.; Dong, M.-J.; Gao, F. G4Bank: A Database of Experimentally Identified DNA G-Quadruplex Sequences. *Interdiscip. Sci.-Comput. Life Sci.* **2023**, *15*, 515–523. [\[CrossRef\]](#)
39. Zok, T.; Kraszewska, N.; Miskiewicz, J.; Pielacinska, P.; Zurkowski, M.; Szachniuk, M. ONQUADRO: A Database of Experimentally Determined Quadruplex Structures. *Nucleic Acids Res.* **2022**, *50*, D253–D258. [\[CrossRef\]](#)
40. Rachwal, P.A.; Brown, T.; Fox, K.R. Effect of G-Tract Length on the Topology and Stability of Intramolecular DNA Quadruplexes. *Biochemistry* **2007**, *46*, 3036–3044. [\[CrossRef\]](#)
41. Bugaut, A.; Balasubramanian, S. A Sequence-Independent Study of the Influence of Short Loop Lengths on the Stability and Topology of Intramolecular DNA G-Quadruplexes. *Biochemistry* **2007**, *47*, 689–697. [\[CrossRef\]](#)
42. Cang, X.; Šponer, J.; Cheatham, T.E. III Insight into G-DNA Structural Polymorphism and Folding from Sequence and Loop Connectivity through Free Energy Analysis. *J. Am. Chem. Soc.* **2011**, *133*, 14270–14279. [\[CrossRef\]](#)
43. Tippiana, R.; Xiao, W.; Myong, S. G-Quadruplex Conformation and Dynamics Are Determined by Loop Length and Sequence. *Nucleic Acids Res.* **2014**, *42*, 8106–8114. [\[CrossRef\]](#)
44. Papp, C.; Mukundan, V.T.; Jenjaroenpun, P.; Winnerdy, F.R.; Ow, G.S.; Phan, A.T.; Kuznetsov, V.A. Stable Bulged G-Quadruplexes in the Human Genome: Identification, Experimental Validation and Functionalization. *Nucleic Acids Res.* **2023**, *51*, 4148–4177. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Sarkar, S.; Tateishi-Karimata, H.; Ohyama, T.; Sugimoto, N. Imperfect G-Quadruplex as an Emerging Candidate for Transcriptional Regulation. *Nucleic Acids Res.* **2025**, *53*, gkaf164. [\[CrossRef\]](#) [\[PubMed\]](#)
46. Li, X.; Zheng, K.; Zhang, J.; Liu, H.; He, Y.; Yuan, B.; Hao, Y.; Tan, Z. Guanine-Vacancy-Bearing G-Quadruplexes Responsive to Guanine Derivatives. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 14581–14586. [\[CrossRef\]](#)
47. Grün, J.T.; Hennecker, C.; Klötzner, D.-P.; Harkness, R.W.; Bessi, I.; Heckel, A.; Mittermaier, A.K.; Schwalbe, H. Conformational Dynamics of Strand Register Shifts in DNA G-Quadruplexes. *J. Am. Chem. Soc.* **2019**, *142*, 264–273. [\[CrossRef\]](#) [\[PubMed\]](#)
48. Mukundan, V.T.; Phan, A.T. Bulges in G-Quadruplexes: Broadening the Definition of G-Quadruplex-Forming Sequences. *J. Am. Chem. Soc.* **2013**, *135*, 5017–5028. [\[CrossRef\]](#)
49. Ngoc Nguyen, T.Q.; Lim, K.W.; Phan, A.T. Duplex Formation in a G-Quadruplex Bulge. *Nucleic Acids Res.* **2020**, *48*, 10567–10575. [\[CrossRef\]](#)
50. Phan, A.T.; Modi, Y.S.; Patel, D.J. Propeller-Type Parallel-Stranded G-Quadruplexes in the Human *c-Myc* Promoter. *J. Am. Chem. Soc.* **2004**, *126*, 8710–8716. [\[CrossRef\]](#)
51. Risitano, A.; Fox, K. Influence of Loop Size on the Stability of Intramolecular DNA Quadruplexes. *Nucleic Acids Res.* **2004**, *32*, 2598–2606. [\[CrossRef\]](#)
52. Fernando, H.; Reszka, A.P.; Huppert, J.; Ladame, S.; Rankin, S.; Venkitaraman, A.R.; Neidle, S.; Balasubramanian, S. A Conserved Quadruplex Motif Located in a Transcription Activation Site of the Human C-Kit Oncogene. *Biochemistry* **2006**, *45*, 7854–7860. [\[CrossRef\]](#)
53. Guédin, A.; Gros, J.; Alberti, P.; Mergny, J.-L. How Long Is Too Long? Effects of Loop Size on G-Quadruplex Stability. *Nucleic Acids Res.* **2010**, *38*, 7858–7868. [\[CrossRef\]](#)

54. Hao, F.; Ma, Y.; Guan, Y. Effects of Central Loop Length and Metal Ions on the Thermal Stability of G-Quadruplexes. *Molecules* **2019**, *24*, 1863. [\[CrossRef\]](#)
55. Smargiasso, N.; Rosu, F.; Hsia, W.; Colson, P.; Baker, E.S.; Bowers, M.T.; De Pauw, E.; Gabelica, V. G-Quadruplex DNA Assemblies: Loop Length, Cation Identity, and Multimer Formation. *J. Am. Chem. Soc.* **2008**, *130*, 10208–10216. [\[CrossRef\]](#)
56. Lim, K.W.; Khong, Z.J.; Phan, A.T. Thermal Stability of DNA Quadruplex–Duplex Hybrids. *Biochemistry* **2013**, *53*, 247–257. [\[CrossRef\]](#)
57. Lim, K.W.; Jenjaroenpun, P.; Low, Z.J.; Khong, Z.J.; Ng, Y.S.; Kuznetsov, V.A.; Phan, A.T. Duplex Stem-Loop-Containing Quadruplex Motifs in the Human Genome: A Combined Genomic and Structural Study. *Nucleic Acids Res.* **2015**, *43*, 5630–5646. [\[CrossRef\]](#)
58. Ravichandran, S.; Razzaq, M.; Parveen, N.; Ghosh, A.; Kim, K.K. The Effect of Hairpin Loop on the Structure and Gene Expression Activity of the Long-Loop G-Quadruplex. *Nucleic Acids Res.* **2021**, *49*, 10689–10706. [\[CrossRef\]](#)
59. Li, W.; Wu, P.; Ohmichi, T.; Sugimoto, N. Characterization and Thermodynamic Properties of Quadruplex/Duplex Competition. *FEBS Lett.* **2002**, *526*, 77–81. [\[CrossRef\]](#)
60. Fotticchia, I.; Amato, J.; Pagano, B.; Novellino, E.; Petraccone, L.; Giancola, C. How are Thermodynamically Stable G-Quadruplex–Duplex Hybrids? *J. Therm. Anal. Calorim.* **2015**, *121*, 1121–1127. [\[CrossRef\]](#)
61. Hatzakis, E.; Okamoto, K.; Yang, D. Thermodynamic Stability and Folding Kinetics of the Major G-Quadruplex and Its Loop Isomers Formed in the Nuclease Hypersensitive Element in the Human c-Myc Promoter: Effect of Loops and Flanking Segments on the Stability of Parallel-Stranded Intramolecular G-Quadruplexes. *Biochemistry* **2010**, *49*, 9152–9160. [\[CrossRef\]](#)
62. Cheng, M.; Cheng, Y.; Hao, J.; Jia, G.; Zhou, J.; Mergny, J.-L.; Li, C. Loop Permutation Affects the Topology and Stability of G-Quadruplexes. *Nucleic Acids Res.* **2018**, *46*, 9264–9275. [\[CrossRef\]](#)
63. Chen, J.; Cheng, M.; Stadlbauer, P.; Šponer, J.; Mergny, J.-L.; Ju, H.; Zhou, J. Exploring Sequence Space to Design Controllable G-Quadruplex Topology Switches. *CCS Chem.* **2022**, *4*, 3036–3050. [\[CrossRef\]](#)
64. Agrawal, P.; Hatzakis, E.; Guo, K.; Carver, M.; Yang, D. Solution Structure of the Major G-Quadruplex Formed in the Human VEGF Promoter in K⁺: Insights into Loop Interactions of the Parallel G-Quadruplexes. *Nucleic Acids Res.* **2013**, *41*, 10584–10592. [\[CrossRef\]](#)
65. Beniaminov, A.; Shchyolkina, A.; Kaluzhny, D. Conformational Features of Intramolecular G4-DNA Constrained by Single-Nucleotide Loops. *Biochimie* **2019**, *160*, 122–128. [\[CrossRef\]](#)
66. Kumar, N.; Maiti, S. A Thermodynamic Overview of Naturally Occurring Intramolecular DNA Quadruplexes. *Nucleic Acids Res.* **2008**, *36*, 5610–5622. [\[CrossRef\]](#)
67. Guédin, A.; De Cian, A.; Gros, J.; Lacroix, L.; Mergny, J.-L. Sequence Effects in Single-Base Loops for Quadruplexes. *Biochimie* **2008**, *90*, 686–696. [\[CrossRef\]](#)
68. Olsen, C.M.; Lee, H.-T.; Marky, L.A. Unfolding Thermodynamics of Intramolecular G-Quadruplexes: Base Sequence Contributions of the Loops. *J. Phys. Chem. B* **2008**, *113*, 2587–2595. [\[CrossRef\]](#)
69. Li, J.; Chu, I.-T.; Yeh, T.-A.; Chen, D.-Y.; Wang, C.-L.; Chang, T.-C. Effects of Length and Loop Composition on Structural Diversity and Similarity of (G3TG3NmG3TG3) G-Quadruplexes. *Molecules* **2020**, *25*, 1779. [\[CrossRef\]](#)
70. Nagatoishi, S.; Isono, N.; Tsumoto, K.; Sugimoto, N. Loop Residues of Thrombin-Binding DNA Aptamer Impact G-Quadruplex Stability and Thrombin Binding. *Biochimie* **2011**, *93*, 1231–1238. [\[CrossRef\]](#)
71. Guédin, A.; Alberti, P.; Mergny, J.-L. Stability of Intramolecular Quadruplexes: Sequence Effects in the Central Loop. *Nucleic Acids Res.* **2009**, *37*, 5559–5567. [\[CrossRef\]](#)
72. Li, Y.Y.; Macgregor, R.B., Jr. A Thermodynamic Study of Adenine and Thymine Substitutions in the Loops of the Oligodeoxyribonucleotide HTel. *J. Phys. Chem. B* **2016**, *120*, 8830–8836. [\[CrossRef\]](#)
73. Tucker, B.A.; Hudson, J.S.; Ding, L.; Lewis, E.; Sheardy, R.D.; Kharlampieva, E.; Graves, D. Stability of the Na⁺ Form of the Human Telomeric G-Quadruplex: Role of Adenines in Stabilizing G-Quadruplex Structure. *ACS Omega* **2018**, *3*, 844–855. [\[CrossRef\]](#)
74. Zhu, H.; Xiao, S.; Liang, H. Structural Dynamics of Human Telomeric G-Quadruplex Loops Studied by Molecular Dynamics Simulations. *PLoS ONE* **2013**, *8*, e71380. [\[CrossRef\]](#)
75. Lenarčič Živković, M.; Rozman, J.; Plavec, J. Adenine-Driven Structural Switch from a Two- to Three-Quartet DNA G-Quadruplex. *Angew. Chem.* **2018**, *130*, 15621–15625. [\[CrossRef\]](#)
76. Nakata, M.; Kosaka, N.; Kawauchi, K.; Miyoshi, D. Quantitative Effects of the Loop Region on Topology, Thermodynamics, and Cation Binding of DNA G-Quadruplexes. *ACS Omega* **2024**, *9*, 35028–35036. [\[CrossRef\]](#)
77. Fu, J.; Chen, T.; Lu, N.; Pan, X.; Tu, J. Decoding G-Quadruplex Stability: The Role of Loop Architecture and Sequence Context in the Human Genome. *Biochimie* **2025**. [\[CrossRef\]](#)
78. Ambrus, A.; Chen, D.; Dai, J.; Jones, R.A.; Yang, D. Solution Structure of the Biologically Relevant G-Quadruplex Element in the Human c-MYC Promoter. Implications for G-Quadruplex Stabilization. *Biochimie* **2005**, *44*, 2048–2058. [\[CrossRef\]](#)
79. Lim, K.W.; Lacroix, L.; Yue, D.J.E.; Lim, J.K.C.; Lim, J.M.W.; Phan, A.T. Coexistence of Two Distinct G-Quadruplex Conformations in the hTERT Promoter. *J. Am. Chem. Soc.* **2010**, *132*, 12331–12342. [\[CrossRef\]](#)

80. Amrane, S.; Adrian, M.; Heddi, B.; Serero, A.; Nicolas, A.; Mergny, J.-L.; Phan, A.T. Formation of Pearl-Necklace Monomorphic G-Quadruplexes in the Human CEB25 Minisatellite. *J. Am. Chem. Soc.* **2012**, *134*, 5807–5816. [[CrossRef](#)]
81. Chen, J.; Cheng, M.; Salgado, G.F.; Stadlbauer, P.; Zhang, X.; Amrane, S.; Guédin, A.; He, F.; Šponer, J.; Ju, H.; et al. The Beginning and the End: Flanking Nucleotides Induce a Parallel G-Quadruplex Topology. *Nucleic Acids Res.* **2021**, *49*, 9548–9559. [[CrossRef](#)]
82. Bugaut, A.; Alberti, P. Understanding the Stability of DNA G-Quadruplex Units in Long Human Telomeric Strands. *Biochimie* **2015**, *113*, 125–133. [[CrossRef](#)] [[PubMed](#)]
83. Zhang, Z.; Dai, J.; Veliath, E.; Jones, R.A.; Yang, D. Structure of a Two-G-Tetrad Intramolecular G-Quadruplex Formed by a Variant Human Telomeric Sequence in K⁺ Solution: Insights into the Interconversion of Human Telomeric G-Quadruplex Structures. *Nucleic Acids Res.* **2009**, *38*, 1009–1021. [[CrossRef](#)] [[PubMed](#)]
84. Pavc, D.; Wang, B.; Spindler, L.; Drevenšek-Olenik, I.; Plavec, J.; Šket, P. GC Ends Control Topology of DNA G-Quadruplexes and Their Cation-Dependent Assembly. *Nucleic Acids Res.* **2020**, *48*, 2749–2761. [[CrossRef](#)]
85. Do, N.Q.; Phan, A.T. Monomer–Dimer Equilibrium for the 5′–5′ Stacking of Propeller-Type Parallel-Stranded G-Quadruplexes: NMR Structural Study. *Chem. Eur. J.* **2012**, *18*, 14752–14759. [[CrossRef](#)]
86. Phan, A.T. Human Telomeric G-quadruplex: Structures of DNA and RNA Sequences. *FEBS J.* **2010**, *277*, 1107–1117. [[CrossRef](#)]
87. Palacký, J.; Vorlíčková, M.; Kejnovská, I.; Mojžeš, P. Polymorphism of Human Telomeric Quadruplex Structure Controlled by DNA Concentration: A Raman Study. *Nucleic Acids Res.* **2012**, *41*, 1005–1016. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.