

REVIEW

Folding dynamics of polymorphic G-quadruplex structures

J. Tassilo Grün¹  | Harald Schwalbe^{2,3} 

¹Department of Chemical and Biological Physics, Weizmann Institute of Science, Rehovot, Israel

²Institute for Organic Chemistry and Chemical Biology, Johann Wolfgang Goethe-University, Frankfurt/M, Germany

³Center for Biomolecular Magnetic Resonance (BMRZ), Johann Wolfgang Goethe-University Frankfurt, Frankfurt/M, Germany

Correspondence

J. Tassilo Grün, Department of Chemical and Biological Physics, Weizmann Institute of Science, Herzl Str. 234, 7640001 Rehovot, Israel.

Email: tassilo.gruen@weizmann.ac.il

Harald Schwalbe, Institute for Organic Chemistry and Chemical Biology, Johann Wolfgang Goethe-University, Max-von-Laue-Strasse 7, 60438 Frankfurt/M, Germany.
Email: schwalbe@nmr.uni-frankfurt.de

Funding information

Deutsche Forschungsgemeinschaft; Hessisches Ministerium für Wissenschaft und Kunst; Council for Higher Education & Israel Academy of Sciences and Humanities

Abstract

G-quadruplexes (G4), found in numerous places within the human genome, are involved in essential processes of cell regulation. Chromosomal DNA G4s are involved for example, in replication and transcription as first steps of gene expression. Hence, they influence a plethora of downstream processes. G4s possess an intricate structure that differs from canonical B-form DNA. Identical DNA G4 sequences can adopt multiple long-lived conformations, a phenomenon known as G4 polymorphism. A detailed understanding of the molecular mechanisms that drive G4 folding is essential to understand their ambivalent regulatory roles. Disentangling the inherent dynamic and polymorphic nature of G4 structures thus is key to unravel their biological functions and make them amenable as molecular targets in novel therapeutic approaches. We here review recent experimental approaches to monitor G4 folding and discuss structural aspects for possible folding pathways. Substantial progress in the understanding of G4 folding within the recent years now allows drawing comprehensive models of the complex folding energy landscape of G4s that we herein evaluate based on computational and experimental evidence.

1 | INTRODUCTION**1.1 | Structural polymorphism in G4**

G-rich nucleic acid sequences are able to fold into non-canonical secondary structures known as G-quadruplexes (G4).^[1–4] Two essential parameters define the basic G4 architecture: (1) four strands (G-tracts) with G-residues form G-tetrads via Hoogsteen hydrogen bonding; (2) these G-tetrads stack on each other and recruit monovalent cations (preferably Na⁺ or K⁺). This broad and simple definition of G4s can be fulfilled by a variety of different G4 and G4-like^[5–7] structures (Figure 1I). Restricting and fine-tuning in specific sequences found in the genome, like:

$$\left[G^1_x - L^I_y - G^2_x - L^{II}_y - G^3_x - L^{III}_y - G^4_x \right], \text{ with } x = 3 - 5 \text{ and } y = 1 - 7$$

leads to a manifold of different conformations characterized by their relative strand orientation and the resulting intramolecular loop geometries (Figure 1II).^[9]

A sequence-based prediction for distinct conformations is already complex within this canonical set of G4 topologies.^[10] In addition to this *canonical* G4 structural polymorphism, new aspects of structural complexity have been described recently, which are referred to as *non-canonical* polymorphism.^[2,11–13] Aspects including bulges,^[14–16] exceptional loop arrangements^[17,18] and snap-back motifs^[19–23] can be observed in G4s from sequences that do not comply with the narrow definition given above (Figure 1, III).

Polymorphism is pronounced among G4s from different sequences, but is also observed within a given G4 forming sequence, leading to concurrent folding isomers in heterogeneous ensembles.^[24–26] In an enlarged conformational space, due to additional combinatorial

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Biopolymers* published by Wiley Periodicals LLC.

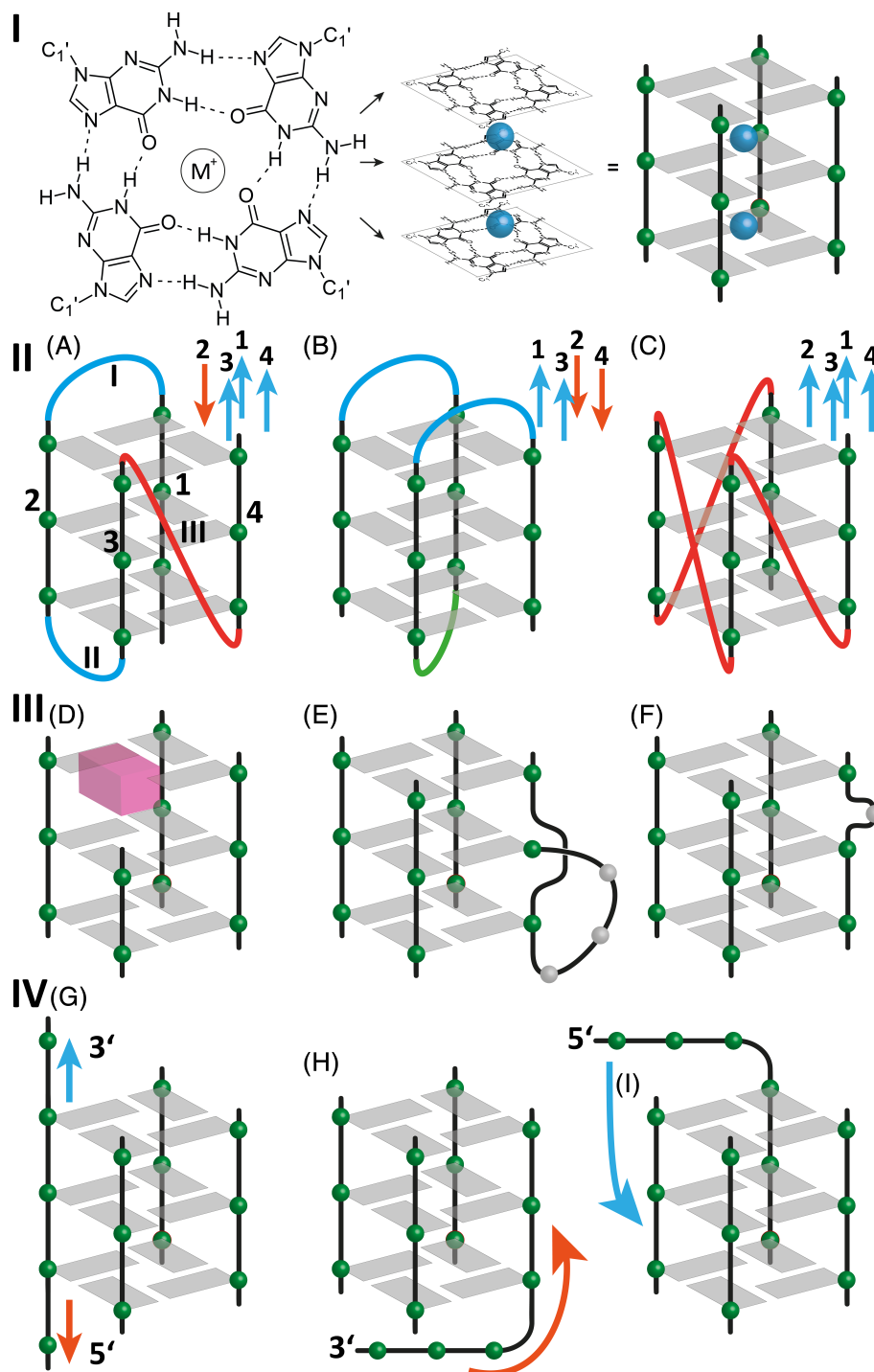


FIGURE 1 Basic principles of G4 architecture. (I) G-tetrad hydrogen bond pattern and tetrad stacking to form the G4 core. Right: Schematic representation of a G4 with G-residues shown in green. Monovalent cations (M^+) are indicated with blue spheres. (II) Canonical G4 structural polymorphism with different folding topologies. Loops I-III and G-tracts 1-4 are indicated in 5'-3' direction. Conformations: (A) *hybrid* (3 + 1) with edgewise (blue) and double chain-reversal (red) loops; (B) *anti-parallel* (2 + 2, basket) with diagonal (green) loop; (C) *parallel*. (III) Non-canonical polymorphism with (d) G-vacancy site, (e) bulged nucleotide and (f) snap-back motif. (IV) Forms of non-canonical folding isomerism with (g) G-register shifts and (h,i) spare-tire exchange (Figure parts adapted from Grün^[8])

possibilities, two particularly intriguing forms of folding isomerism arise (Figure 1, IV):

Firstly, if the number of subsequent G-residues exceeds (or undercuts^[27-29]) the number of G-tetrad layers in a quadruplex, an

exchange of a single G-register along the G4 core can be observed, leading to a conformational subset of shifted G-register isomers.^[30]

Secondly, if the number of subsequent G-tracts is greater than four, different isomers can be formed by incorporating different

G-tracts into the formation of the G4 core.^[31,32] For the latter, the term *spare-tire* isomerism has been newly coined.^[31,33]

Note that we here define both forms of folding isomerism with respect to the relation of the involved distinct G4 conformations arising from the same nucleotide sequence.

1.2 | Regulatory role of G4

The emerging role of G4 forming elements in the human genome has been reviewed extensively in the context of transcription regulation,^[34] replication,^[35,36] genomic instability,^[37,38] epigenetic modifications^[39–43] and telomere stabilization.^[44] Since the first reports on transcription regulation through small-molecule ligands that target gene promoter G4s,^[45] a plethora of approaches have evolved that focus on the development and characterization of G4 stabilizing agents.^[46,47] While these strategies aim at stabilizing G4s as molecular mechanism in novel anti-cancer therapies,^[48] more recently, converse strategies have been proposed to counteract genomic instability induced by the stable formation of G4s in chromosomal DNA.^[37] Both these general approaches, stabilization and destabilization, interfere with a misregulated G4 formation at a pathological stage. The ambivalent consequences of G4 structure formation in different contexts highlights the requirement for the dynamic regulation of G4s: transient folding and unfolding has to be maintained for balanced cellular homeostasis. It is thus not surprising that an inherently dynamic nature is also a key feature of G4 formation in RNAs.^[49–51]

The versatile potential to fold into a manifold of distinct G4 structures has thermodynamic, kinetic and also biological consequences. In particular for the latter, it is a crucial aspect to ensure adaptability of G4 elements in response to external stimuli. As a prime example for the structural adaptability, the functional role of spare-tire G-tracts has been proposed to maintain G4 formation after oxidative damage, since G-stretches with increasing lengths are especially prone to oxidize.^[31,32,52–55] This mechanism enables the subsequent recruitment of repair machineries, which can reset the damaged DNA stretch.^[55,56] In this simple picture, the maintenance of function for polymorphic G4s refers to an on/off switching for example, in transcriptional control, if a G4 structure is present or not.

However, there is now growing evidence that G4 polymorphism itself is a crucial aspect of G4 regulatory function. We exemplify these new findings with two structural aspects: Spare-tire isomers with different loop lengths have drastically different affinities towards G4 interacting proteins, even for G4s with the same folding topology (e.g., parallel loop isomers).^[57–59] Structural isomers with different topologies (e.g., hybrid and parallel) result in vastly different unwinding efficiencies for G4 specific helicases.^[38,60–65] These structural aspects might already arise from small changes. Thus, while formation and fold topology of a G4 might be maintained in principle, the result of modified sequences (due to oxidation or mutations^[66,67]) could be a completely altered regulation cascade.^[59]

In view of the consequences of G4 polymorphism, the simple model that G4s act as steric bulking structures, often anthropomorphically called roadblocks, is not suited to explain their regulatory

function.^[39] Hence, statements about a general functionality for G4 formation have to be taken with care, whenever the specific conformation is not considered. For G4 forming sequences that are able to fold into different stable conformations, the individual kinetics of concurrent folding pathways towards a specific conformation might be biologically even more relevant than the stabilities at thermal equilibrium.

2 | EXPERIMENTAL ASPECTS

2.1 | Preparation of non-equilibrium conformational states

A prerequisite to study coherent structural changes is the preparation of suitable starting points away from equilibrium (Figure 2). The experimental approaches to prepare any kind of trapped, retained or excited state can heavily influence the folding progression, possible pathways and kinetics.^[68–70] It is thus worthwhile to compare and evaluate the experimental premises to understand possible ambiguous results.

2.1.1 | Reversible folding (I)

Mechanical unwinding

Mechanical unwinding is an intriguing, but technically demanding possibility to investigate G4 folding as a measure of force under isothermal experimental conditions and at constant cation concentration (Figure 2A). Cheng *et al.* have used this method to study the folding/unfolding of a *BCL2* promoter G4 with single molecule force spectroscopy; the observed force changes are in the range of pN.^[71] Using a tethered oligonucleotide that has been fixed with magnetic beads, they were able to describe kinetic differences for spare-tire isomers of the *BCL2* G4 with this approach. While this method clearly ensures a pure unfolded state, it should be noted that this state will be characterized by inherently lower conformational entropy compared to other denatured states due to fewer translational degrees of freedom.^[72]

Thermal hysteresis

In reversible thermal melting and annealing experiments, G4 forming oligonucleotides can show pronounced hysteresis (Figure 2B).^[73–75] The complex behavior at thermal transitions can be used to gain insight into the folding process of G4s, typically with photometric detection (UV/CD). Mittermaier *et al.* have presented a sophisticated strategy to extract dynamic information from experiments under deliberately chosen conditions that provoke thermal hysteresis.^[30,69,73,76] Rapid temperature jumps (T-jump) have been used to induce G4 folding in circular dichroism (CD)-spectroscopic^[77] and mass spectrometric^[78] setups and new probe designs could potentially allow T-jump induced folding also for NMR spectroscopy.^[79–81] Thermal (un-)folding is typically limited to lower than physiological K^+ concentrations, due to the high thermal stability of many G4 structures.^[74,75]

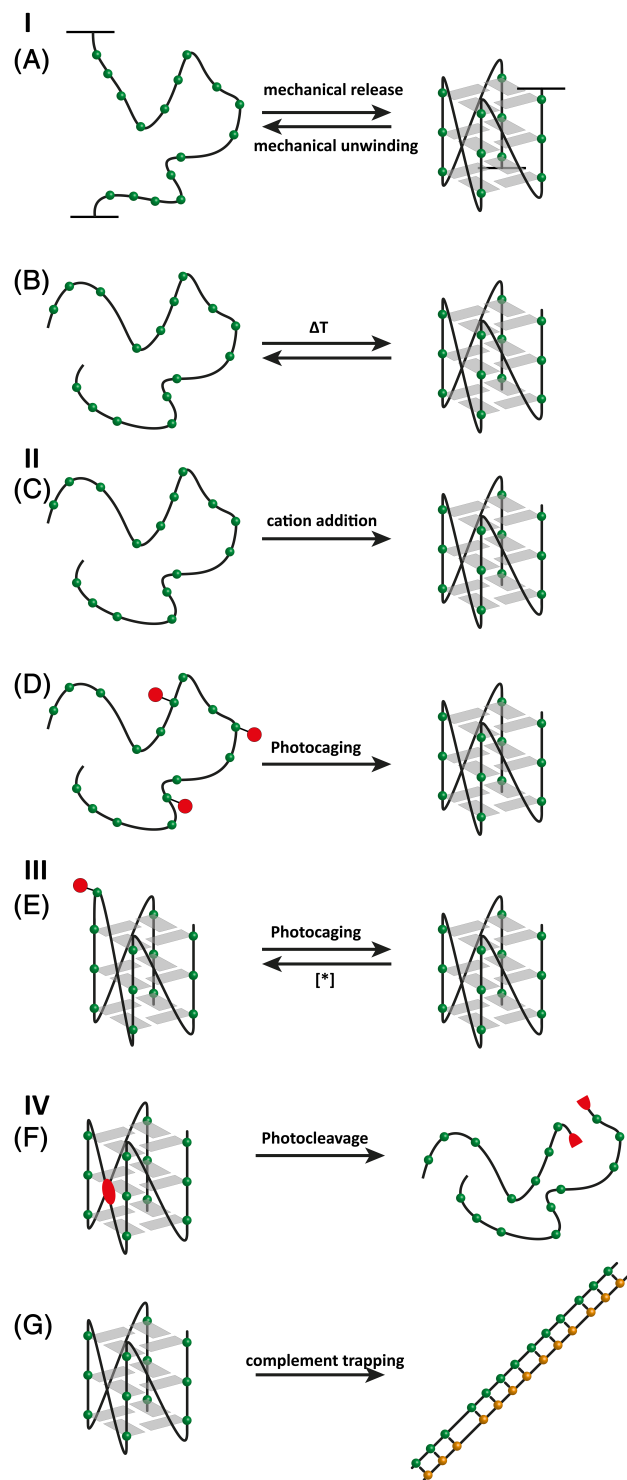


FIGURE 2 Non-equilibrium G4 dynamics. (I) Reversible methods: (A) mechanical unwinding with magnetic/optic beads, (B) thermal melting observed with T-jump or thermal hysteresis. (II) Irreversible folding methods: (C) rapid mixing with monovalent cations, (D) photolytic release of photolabile protecting groups (photocages). (III, E) conformational selection of folded G4s with photocages, after photolysis complete refolding or re-equilibration into a polymorphic ensemble [*] can be observed. (IV) Irreversible Unfolding methods: (F) Photocleavage of the DNA backbone to destabilize intramolecular G4s, (G) complement trapping to induce duplex formation

2.1.2 | Irreversible folding (II) and refolding (III)

Cation-induced folding

A widespread strategy of inducing G4 folding is to dissolve DNA with G4 forming sequences under buffer conditions that lack G4 stabilizing cations (Figure 2C). Thus, G4 folding is inhibited at ambient temperatures. Under careful experimental control (e.g., unwanted cation uptake from tubes), in particular DNA G4s can be prepared in an unfolded state. Folding at a specific temperature then can be induced by mixing with for example, Na^+ or K^+ which allows the application for any spectroscopical method as readout.^[82] In general, cation-induced folding is a very simple, reliable and broadly applicable method, yet the degree of denaturation has to be evaluated carefully with spectroscopic methods. For RNA G4s for example, preparation of an unfolded state following this procedure is more difficult. In a previous study pre-formation of an RNA G4 fold was observed even in the obvious absence of K^+ -ions and complete folding occurs at much lower K^+ equivalents as compared to the corresponding DNA sequence.^[83] CD spectra provide characteristic patterns for specific G4 architectures. Even more insightful, NMR spectra of the fingerprint region show hydrogen-bonded imino ^1H signals that allow a sensitive evaluation of residual structure formation.^[84] Indeed, in many cases the formation of pre-folded states (see below) has been observed or structure formation even at very low, sub stoichiometric K^+ concentrations.^[83,85] In many experiments, the concentration of K^+ is often below physiological concentrations, which greatly affects the thermal stability of G4s.^[75] The kinetics are accelerated with increasing K^+ concentration (even at unphysiological high concentrations of greater than 100 mM)^[86] but the main effects, in particular the branching of pathways are present already at very low K^+ concentrations (<3 mM).^[82,83]

Photocaging for conformational selection

Photocaging of RNA and DNA with photolabile protecting groups on their nucleobase moieties can be applied for studying G4 folding.^[87,88] Photocages can inhibit hydrogen bond interactions site specifically on distinct nucleobases in the oligonucleotide and can be removed upon light irradiation with a selective wavelength, thereby releasing the completely unmodified nucleobase. This concept was first exploited to study RNA refolding,^[89–92] and has been applied to G4s recently in two different ways, either by selecting single folded conformations out of a polymorphic conformational ensemble from a G4 forming sequence instead of using sequence mutations (Figure 2E)^[69], or to suppress completely the G4 folding to trap the unfolded state (Figure 2D).^[33,76] The method allows investigating the isothermal folding at constant experimental conditions and ensures a robust disruption of hydrogen bond interactions or preferential pre-orientations. While photocages are introduced at the nucleobases and act irreversible in the direction of *folding*, the incorporation of photosensitive scaffolds into the DNA backbone can be used to cleave and hence *unfold* G4s (Figure 2F),^[93] or enable a reversible *switching* between different G4 conformational states.^[94]

2.1.3 | Irreversible unfolding (IV)

Unfolding with complementary trapping

The folded state itself can also serve as starting point for investigating the reverse folding trajectory of an order–disorder transition (Figure 2G). Addition of the complementary strand (complementary trapping) has been used with PNA and DNA strands to trigger the irreversible unfolding of G4s towards a duplex fold.^[95–98] The observed kinetic rates for unfolding in these experiments are typically orders of magnitude slower than for folding, reflecting the high stability of the G4 fold. Klejevskaja *et al.* have studied a self-assembled double stranded DNA mini-circle with an embedded single stranded region that codes for the cMYC G4.^[99] Using a clever FRET-based strategy with two dyes placed at the brink of the inserted G4 element, they measured the G4 unfolding kinetics after adding the complementary strand. The observed unfolding is approximately 10 times slower than for the isolated single stranded G4 oligonucleotide, reflecting a higher kinetic stability under conditions with restricted flexibility that resemble the chromosomal context of genomic DNA. Since the complementary strand itself is typically able to form a folded i-motif secondary structure, the experimental conditions for complementary trapping have to be evaluated carefully to prevent convolution of the structural dynamics.

2.2 | Spectroscopic methods to study folding and refolding kinetics

2.2.1 | Circular dichroism

CD spectroscopy is an excellent method to monitor folding kinetics of G4s and provides an easy access to the basic structural constitution of G4s. Much of the pioneering work on G4 folding has been conducted or supported by CD spectroscopy.^[75,82,95,100–102] CD spectroscopy gives a simple and characteristic readout that can be used to distinguish between an unfolded state and different folded conformations [parallel: ~260–265 nm (+), anti-parallel: ~295 nm (+), hybrid ~265/295 nm (+)].^[103] CD spectroscopy is mainly indifferent to aspects of non-canonical polymorphism such as parallel G-register shifted or spare-tire isomers. However, there are now sophisticated analysis tools available that allow deconvolution of complex CD spectra from polymorphic G4s.^[103,104] Time-resolved CD spectroscopy was used in combination with an laser-induced temperature jump to monitor G4 folding down to a millisecond timescale.^[77] UV spectroscopy is suited as a readout in a similar way, but is restricted to a simple “folded/non-folded” monitoring with characteristic changes at 295 nm.^[105–108]

2.2.2 | Mass spectrometry

Mass spectrometry (MS) gives orthogonal insight into aspects of G4 folding that are not directly available with spectroscopy, while MS

itself does not provide direct structural information.^[109] MS has been used for the evaluation of G4-ligand binding and of the folding pathways of G4s by analysis of cation binding stoichiometry.^[110,111] A precise evaluation of K⁺ binding to the DNA strand is crucial to understand the enthalpic and entropic contributions that affect G4 folding in thermal experiments.^[111] In a recent fascinating paper, Gabelica *et al.* have presented an approach for the detection of mass-resolved CD spectra of G4 forming oligonucleotides.^[112] This powerful method in combination with advanced computational methods for the deconvolution of CD spectroscopical and mass spectrometric parameters will enable new perspectives on G4 folding.^[104,113–116]

2.2.3 | Single molecule spectroscopy

Force spectroscopy/microscopy can be used for directed force manipulations on G4 oligonucleotides accomplished with magnetic^[71,117–119] or optical beads/tweezers,^[120–122] also in combination with fluorescence detection.^[123–125] Sugiyama *et al.* have demonstrated the observation of G4 folding in DNA nanostructures using high-speed atomic force microscopy (AFM).^[72,126–128] Förster resonance electron transfer (FRET) yields a very specific readout for two-site distances, which requires, however, the incorporation of dye labels that potentially bias the G4 structural integrity.^[75,129] While FRET provides no direct information on the G4 conformation it allows a very selective observation of folding trajectories in single molecule experiments.^[99,130,131] The selective observation of different FRET states make this method suitable for the application in high molecular weight complexes, in particular for the investigation of G4 unwinding by helicases.^[64,65,132–134]

2.2.4 | Nuclear magnetic resonance

Nuclear magnetic resonance (NMR) spectroscopy is a powerful and versatile method to study G4 structural dynamics at atomic resolution.^[81,135] Substantial information on the number of states adopted by a given G4 can already be read-off in one-dimensional NMR spectra, as the spectral regions for imino hydrogen atoms involved in Watson-Crick, Hoogsteen or i-motif interactions are clearly distinct. Counting the number of resolved imino hydrogen atoms often already provides a direct readout of multiple, polymorphic states in slow conformational exchange, implying at least millisecond lifetimes of these states.^[136] The quantification of arising imino ¹H signals in time-resolved experiments was used to study the folding of RNA and DNA G4s^[33,69,83,85] and DNA i-motifs,^[137] respectively. To access the rich and complex structural information of NMR spectra, however, higher dimensional homo- or heteronuclear correlated spectra are required since signal resolution decreases with increasing molecular size.^[138–140] Especially nucleic acids show an inherently poor spectral dispersion due to only four different nucleobases that constitute the basic polymer building blocks.^[136]

3 | FOLDING ENERGY LANDSCAPE AND FOLDING PATHWAYS

3.1 | Ensemble effects

From NMR and CD spectroscopic G4 folding experiments, we derive a conformational energy landscape that depicts the entire experimentally observable conformational space of G4 DNAs. To some extent, this landscape is simplified, compared to theoretical landscapes predicted by molecular dynamics (MD) simulations that potentially aim to represent the complete phase space.^[70,141] In spectroscopic experiments only macrostates, represented by a particular conformational state (or structural fold) of the DNA strand of a certain lifetime can be observed. The multiples of subordinated microstates that contribute to for example, the conformational entropy of a macrostate thus are rather a subject of MD simulations than experimental evidence. Nevertheless, also in experimentally derived kinetic models conformational states should be referred to as ensembles (e.g., transitory ensemble), when the range of involved microstates exceeds a structurally clearly defined macrostate.^[142]

The description of folding pathways can be experimentally achieved on a single molecule level or in an ensemble average. In NMR, for example, the observation of an ergodic ensemble^[143] (100 μ M NMR sample \approx 300 nmol DNA \approx 10¹⁷ folding events) leads to an extensive mapping of the energy landscape. In CD spectroscopy, this number is lower by a factor not smaller than 10⁻³. Different to other biomolecules, in the special case of chromosomal DNAs the folding of a particular G4 sequence is not an ensemble process, but a single event in each living cell. The assumption that G4 folding is an infrequent event is based on the fact that G4 folding does not happen spontaneously in double stranded DNA. A presumable requirement for G4 folding in any chromosomal region different from the single stranded telomeres is negative superhelicity, which is induced during transcription or replication.^[39,144-147] Interestingly, G4 folding is also associated to accessible chromatin states and therefore might even precede transcription.^[148] Regardless of other cellular triggers for G4 folding, we try to give a rough estimate for the relevant rates of G4 folding events with respect to transcription. Assuming a total intracellular concentration of \sim 10⁵ mRNAs per cell,^[149,150] with a median copy number per gene of \sim 17 per mRNA (in comparison: the protein concentration is \sim 10⁹ per cell,^[151] with \sim 50.000 copies per protein^[152,153]). The typical intracellular lifetimes for mRNAs are longer than several hours,^[152-154] but especially regulatory mRNAs have significantly shorter lifetimes,^[154] leading to a potentially higher transcription rate of certain G4 mediated genes. However, these calculations still lead to only a very few, approx. <100 potential folding events of a distinct promoter G4 per hour per cell. Hence, the total number of transcription initiated G4 folding events is hardly to call a dynamic ensemble. This situation changes, if tissues are considered: in a tumor tissue for example, \sim 10⁸ to \sim 10⁹ cells are observed per gram tissues,^[155] which adds up to a tremendous number of independent G4 folding processes *in vivo*.

3.2 | A view from computation

In stark contrast to funnel-like energy landscapes that describe folding trajectories of proteins, G4-forming oligonucleotides exhibit a rough conformational energy landscape (Figure 3, I).^[70,141] Instead of approaching a native folded state following a funnel-like, two-state folding transition, different competing basins of metastable states are observed. These basins, or flat wells, in many instances lead to concurrently, coexisting folded conformations (different macrostates) in thermal equilibrium. During folding, during which high-energy non-equilibrium states are initially formed, different conformational states with low energy barriers can be sampled. Thus, the routes along the folding energy landscapes are rather complex and involve stochastic sampling of different folding and misfolding pathways. The relative contributions of different pathways for folding reaction undergoing kinetics partitioning has been studied with MD simulations, and revealed possible involvements of hairpins,^[142,156] triplexes^[157] and strand slipped conformations.^[29,70,158] Stable G-hairpins,^[159] cross or parallel G-hairpins^[142] and newly discovered pseudocircular G-hairpins^[160] represent interesting possible waymarks along the folding pathways. Derived from the computational picture of the conformational energy landscape, the underlying molecular folding mechanisms with competing trajectories are described as kinetic partitioning mechanism, as opposed to a funnel-like mechanism.^[70]

3.3 | Folding kinetics and rate constants

The kinetic partitioning populating parallel folding pathways causes complex and multiphasic folding kinetics (Figure 3, II). The kinetic rates for G4 folding are vastly different^[86] and span a range between sub-seconds^[77] to minutes.^[161] The kinetics report on structural rearrangements that are driven by enthalpic and entropic contributions, leading to the observation of different temperature dependencies in an Arrhenius analysis. Linear Arrhenius behavior with either positive^[100] or negative^[162] activation energies, as well as non-linear^[86] Arrhenius behavior has been reported.^[33] Since the apparent activation energies report only on the observable rate constants, it is conceivable that the overall folding kinetics are determined by different rate limiting steps in the multiphasic folding mechanism. Noteworthy, the kinetics are highly influenced by the experimental setup, temperature-induced folding is typically observed to be orders of magnitude faster than for example, isothermal K⁺-induced folding.^[69,77] The reasons for the different kinetics can be explained by differences in the energetic nature of the unfolded states, that is, different entropy contributions, or the presence of pre-folded states.^[68,70,111] In general, the comparably slow folding kinetics of DNA G4s reflect the consequences of the kinetic partitioning mechanism, including the sampling of misfolded macrostates.^[70] The pronounced polymorphism found in DNA G4s is less abundant in RNA G4s and consequently the folding kinetics are remarkably faster.^[83,86] A main contribution for the diverging kinetics of DNA versus RNA G4s might be the slightly favored *anti*-glycosidic conformation of RNA.^[83] Giving a random distribution of *syn*- and *anti*-glycosidic conformations results in

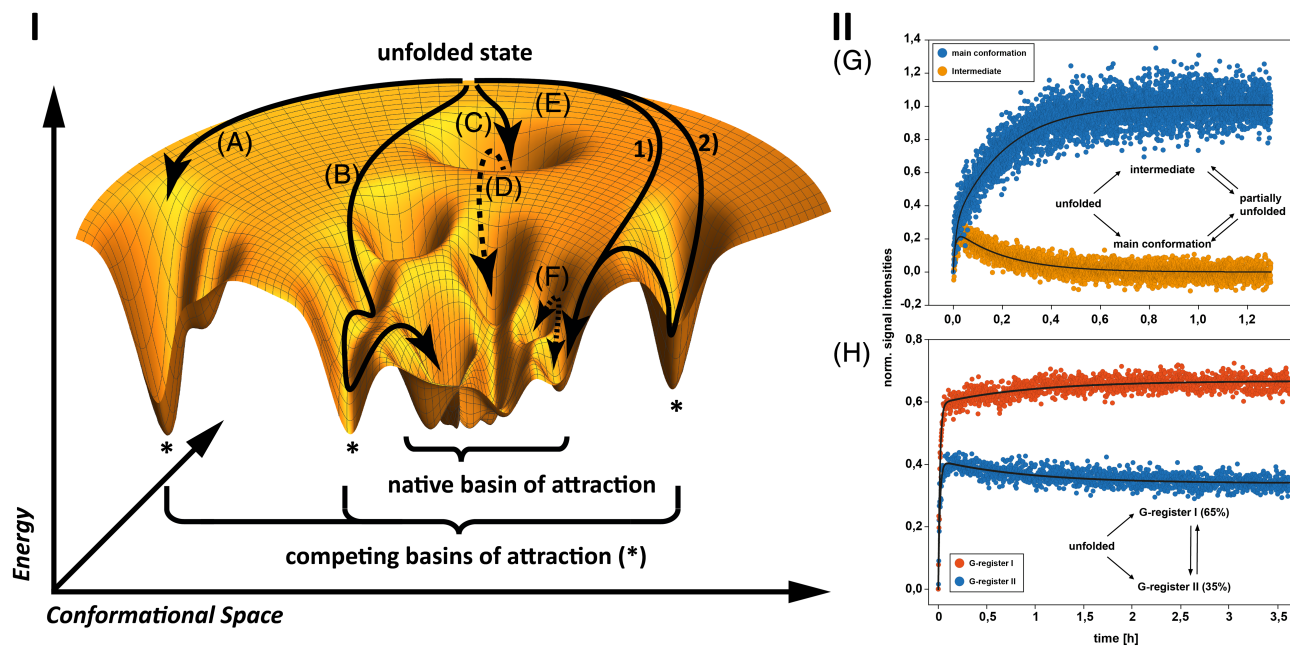


FIGURE 3 Folding Pathways along the rough conformational energy landscape. (I) A central native basin of attraction represents an ensemble of conformational states with similar stabilities and low relative energy barriers. Competing basins of attraction might trap conformations that are separated by high activation energy barriers. (A) Folding into a competing basin can result in isolated conformations (*coexisting or metastable*). (B) Multiphasic folding pathway with short-lived *intermediates*. (C) Folding into a competing basin that can refold into the native basin. (D) Refolding via transitory ensembles with high activation energy barriers. (E) Parallel folding pathways with direct folding (1) and folding via *off-pathway intermediates* that require refolding (2). (F) Low activation energy barrier refolding or minor rearrangements between coexisting states. (II) Examples for experimental observations of kinetic partitioning folding mechanisms; kinetic traces and proposed kinetic models for different conformations of the *cMYC* G4. (G) Folding with parallel pathways and *off-pathway* formation according to (E). (H) Folding in two coexisting G-register isomers with subsequent refolding according to (B,F) (Figure parts adapted from Grün^[8])

$2^{13} = 4096$ possible microstates in an unfolded DNA chain able to fold into a 3-tetrad G4, while an RNA chain does not require *syn/anti* flipping.^[70]

At this point, we wish to compare the folding kinetics with the biologically relevant timescale for DNA G4 folding. In the context of transcription regulation, the crucial temporal parameter is the transcription rate of RNA polymerase II (RNAPII). The reported RNAPII rates can be estimated with $\sim 1.1\text{--}4.3$ kilobases per minute (~ 15 ms per nt on average)^[163,164] in eukaryotic cells. This rate can be significantly slowed down to ~ 6 bases per second (~ 170 ms per nt),^[164] which is in line with observed dwell times of several seconds for molecules in transcription domains.^[165]

3.4 | Loop effects

It is self-evident that the constitution of the loops restrains the flexibility of any advanced G4 folding states and thus the loops have major contributions to the folding dynamics. The effects of loop lengths on stability and preferences for particular conformations has been dissected earlier.^[166–168] The stable formation of parallel conformations is guided by single nucleotide lateral loops ($L^I + L^{II}$), while the length of the proximal loop is less crucial.^[169–172] Interestingly, the proximal loop was found to be the main effector for the folding kinetics of parallel G4s. Nguyen *et al.* found a nearly linear dependence of the folding time with increasing length (1–25 nts) of the proximal loop (L^I).^[161]

However, if the loop sequence is redesigned to form a duplex stem loop (internal hairpin) the folding time re-accelerates as compared to an unstructured loop of the same length. In a recent study, we have also observed decelerated kinetics for a 6 nt long loop, as compared to a single nucleotide loop in a *cMYC* spare-tire isomer.^[33] This conformation showed non-Arrhenius kinetics that highlight the impact of the flexibility of the proximal loop for parallel G4s.

3.5 | Individual states along the folding energy landscape

3.5.1 | Unfolded ensembles

If the G4-forming DNA sequences are embedded in a native genomic context, it is not straightforward to define an unfolded state. Two general cases can be distinguished: a double stranded, helical and a single stranded, unwound situation. In every instance, the situation will be different from biophysical studies, where typically isolated oligonucleotides of a finite length are investigated. To define an unfolded state of a G4-forming oligonucleotide, we assume an idealized, undisturbed chain represented by a random coil like, unstructured conformational ensemble. Given such a state would physically exist under physiological conditions, we would expect that this chain samples the entire conformational space randomly, resulting in a zero

net-sum for the orientation vectors of all residues. Under real, non-ideal sample conditions the deviations from the ideal definition of unfolded states are highly determined by the experimental design. In any case, the resulting unfolded ensembles will likely feature a non-random distribution of conformational microstates resulting in a preferential pre-orientation or the formation of pre-folded states.

3.5.2 | Pre-folded states

How much the existence of different pre-folded states affects the folding pathways has been demonstrated by Frelüh *et al.*,^[68] examining the consequences of temperature and pH on the folding of human telomeric G4 sequences. They find different defined pre-folded structures even in the absence of cations that guide the folding towards either a hybrid-1 or a hybrid-2 final conformation. Similar observations for the formation of pre-folded states, presumably hairpin-like conformations are also reported for the bimolecular *Oxytrichia nova* telomeric G4^[173] as well as for the *EGFR*^[25] and *cKit*^[174] promoter G4s. The hydrogen bond interactions in pre-folded states mark energy barriers that need to be overcome if the pattern is non-productive for the final fold. While the beforementioned sequences show only partial folding patterns that are also retained in the final G4 folds, the G4-forming sequence of the *WNT1* promoter adopts a completely aligned, stable hairpin conformation under K^+ -free conditions.^[175] Though this hairpin cannot be defined as pre-folded state but rather as an alternative fold, it gives crucial insight into the consequences of starting the folding as a transition from a base-paired state to the G4 rather than from an unfolded or partially folded state. The folding kinetics in the reported study were dominated by the unfolding rate of the hairpin and hence are rather slow (82 min^{-1}).

3.5.3 | Intermediate states

We can imagine any kind of intermediate state being either along a folding pathway that leads to the final G4 conformation (*on*-pathway) or being along a second folding pathway that reaches an impasse or leads to an alternative G4 conformation (*off*-pathway). For *on*-pathway trajectories, several intermediate states during G4 folding have been proposed and predicted, most prominent among them are hairpins and triplexes.^[142,176] While the spectroscopical evidence for hairpin pre-folded states gives a good representation of a trapped intermediate state, the transient formation of intramolecular G-triplexes remains more elusive.^[177–179] The quest to detect triplex folding intermediates is mainly driven by the assumption of a consecutive, sequential assembly of G-tracts: single strand—hairpin—triplex—quadruplex. This kind of a stepwise strand recruitment folding mechanism has been discussed first for inter/tetramolecular G4s, where this stepwise assembly is intuitive.^[180–182] Predictions from molecular dynamics simulations give insight into the feasibility of intramolecular triplex formation,^[70,142,157] but experimental evidence for their *transient* formation remains ambiguous.^[127] Intramolecular parallel, anti-parallel and hybrid G4s have different sterical requirements for the orientation of the

strands and the resulting interconnecting loops. Even for a rigorous sequential strand recruitment, several options are possible that lead to different intermediates. While triplex formation for a hybrid G4 resulting from a hairpin and a third strand can be a reasonable arrangement, for an anti-parallel (*chair*) G4 the merging of two hairpins is an easy way to go around triplexes.^[183] The relative strand orientation does not necessarily imply a hydrogen bonded, rigid tetrad formation. The assembly of a quadruplex can also involve a *zipping* along tetrad formation of the G-tracts into the final G4 structure.^[110] This results in $n \times$ possible G-tetrad intermediates for a n -layered G4. Gabelica *et al.* demonstrated with mass spectrometry that the recruitment of cations occurs stepwise.^[110] Their detection of a single K^+ -bound state is thus indicative for a 2-tetrad (or tetrad + triad)^[111] intermediate state, where the cation is coordinated between the G-planes.

3.5.4 | Misfolded states

During folding, different kinds of misfolded states can be generated that are associated to the different aspects of structural polymorphism in G4s and it depends on definition what to call a misfolded state. Misfolding *during* folding can lead to *off*-pathway intermediates that need to unfold to form a stable G4 structure. Misfolding can also be defined as the formation of different stable or metastable G4 conformations. An interesting example for the latter can be seen in human telomeric tandem repeats. The uniformity of the repeating G-tracts can lead to a kinetic frustration, given a random, unpreferential nucleation of G4s.^[120,184,185] On the other hand, the formation of G-wires for example, in *Tetrahymena* telomeric repeats, can result in different species of higher-order G4 structures.^[186] These structures are polymorphic, but with a well-defined periodicity, which makes the term *misfolded* inappropriate.

3.5.5 | Metastable states

Misfolding can lead to the formation of fully folded intermediates that are long-lived (exceeding biological relevant timescales), which we define as metastable states in the following. Examples for kinetic products of G4 folding away from thermal equilibrium have been reported with direct NMR-spectroscopical evidence for topological isomers of the telomeric G4 (hybrid-1 and -2)^[85]; for G-register shifted,^[69] and for spare-tire^[33] isomers of the *cMYC* G4. In all instances, the subsequent refolding into thermal equilibrium is on a timescale of several hours at ambient temperatures.

3.5.6 | Coexisting states

The refolding dynamics between isomeric G4 conformations might be as well reversible, instead of (nearly) depleting an initial kinetic product. The G4 sequence from the human telomerase promoter (*hTERT*) adopts a hybrid and a parallel conformation in a 0.4:0.6 ratio.^[24]

Nußbaumer *et al.*^[187] have investigated the refolding dynamics with time-resolved heteronuclear 2D NMR after incorporation of ¹³C building blocks. We have also investigated the refolding dynamics between the hybrid and the parallel conformation and found consistently slow kinetics in ($\sim 0.7 \text{ h}^{-1}$) proceeding towards a thermal re-equilibration of the populations, irrespective of the starting point (100% hybrid or 100% parallel).^[69] This example shows that coexisting states are indeed cross talking and are in dynamical exchange, although on a slow timescale.

4 | PREVALENCE OF G4 ELEMENTS WITH NON-CANONICAL FOLDING ISOMERISM IN GENE PROMOTERS

The vast majority of G4 forming sequences that are found in gene promoters feature the possibility to undergo G-register or spare-tire exchanges.^[9,30,31,188] In the *cMYC* G4 both kinds of folding isomerism

are found in a compact sequence segment that can be deconvoluted in structural and kinetic aspects. However, in many cases the promoter regions are convoluted with multiple interdependent G4 elements, which complicates a full disentanglement of the folding dynamics.

4.1 | *cMYC* promoter

The G4 forming sequence (Figure 4, I) derived from the nuclease hypersensitive element III₁ (NHE-III₁) –142 to –115 nucleotides upstream the P1 promoter of the human *cMYC* oncogene is a prime example for a polymorphic G4 ensemble and its structure has been studied extensively in the recent two decades.^[20,45,145,192–196] The 27 nucleotide long sequence features five G-tracts (1–5), of which three G-tracts have four consecutive G residues. A total of three stable parallel spare-tire isomers involving strands 1234,^[145] 1245^[195] or 2345^[194] have been reported with a subset of four different

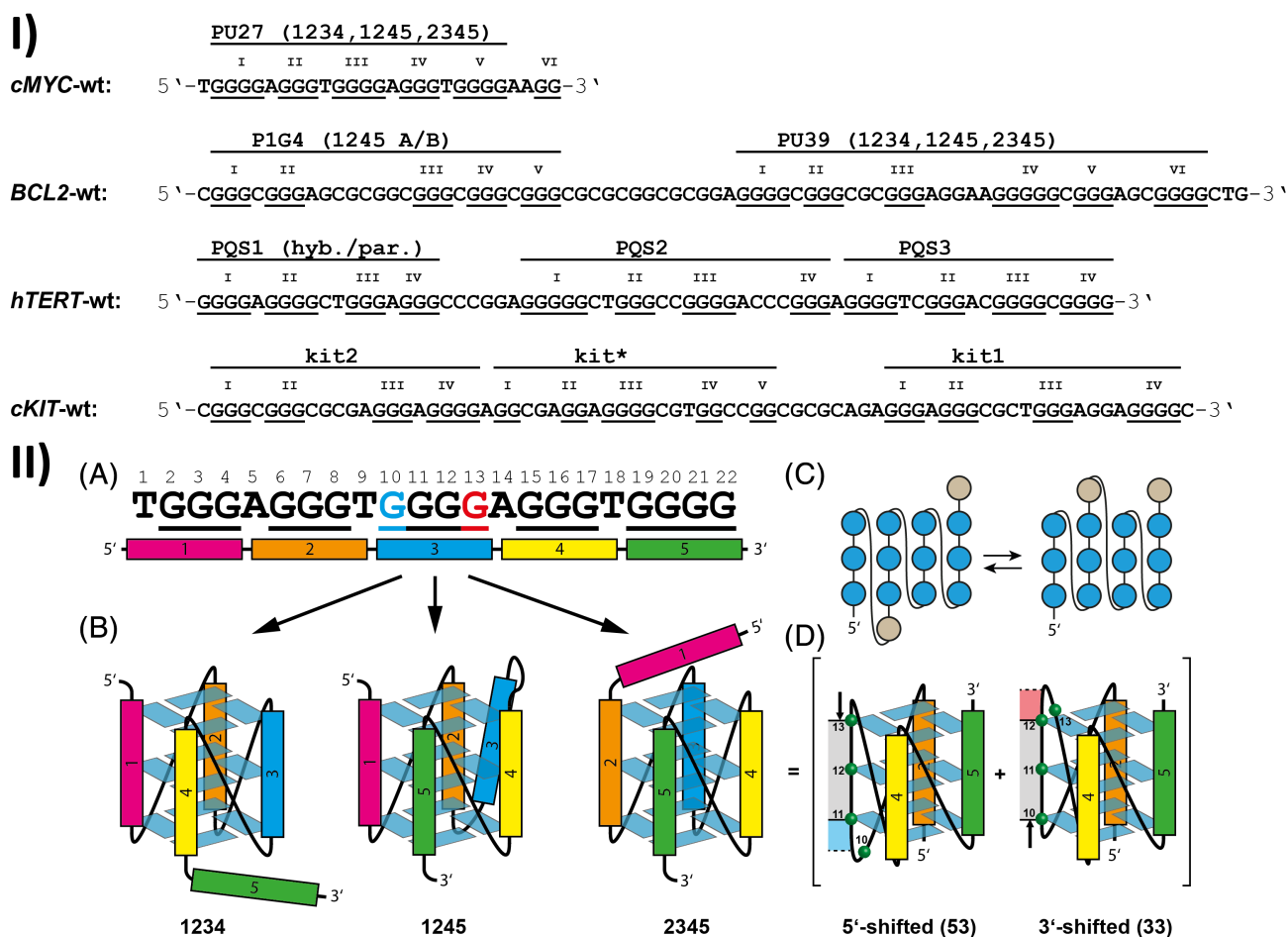


FIGURE 4 Non-canonical Polymorphism in promoter G4s. (I) wt-Promoter sequences^[189] of *cMYC*^[45] (PU27), *BCL2*^[169,190] (P1G4 and PU39), *hTERT*^[24,191] (putative quadruplex sequences 1, 2, 3) and *cKIT*^[174] (kit2, kit* and kit1). (II) *cMYC* G4: (A) 22-mer wildtype sequence spanning five G-tracts. (B) Schematic representation of spare-tire isomers incorporating different G-tracts into parallel G4 formation: 1234, 1245 and 2345. (C) Ball representation and (D) schematic representation of G-register isomers of the 2345 conformation (Figure parts adapted from Grün.^[8] Adapted with permission from Grün *et al.*^[69] Copyright 2020 American Chemical Society. Adapted with permission from Grün *et al.*^[33] Copyright 2021 The Authors. Published by American Chemical Society)

G-register isomers^[30,162,192] for the majorly populated spare-tire isomer 2345 (Figure 4, II).^[145,162,192,194,195] Even though the overall structure of the G4 core architecture is very much alike, the different lengths of the lateral and proximal loops cause diverging thermal stabilities for all conformations.^[162,197] The resulting conformational landscape is determined by two attributes with a vastly different impact on folding kinetics. While the separation into spare-tire isomers requires recruitment of differently distanced parts of the DNA chain, the division in G-register isomers is topologically more indifferent.

One thus may assume (Figure 5) three major wells representing the spare-tire isomers, where the well for conformation 2345 is fine structured in four competing basins of attraction that represent the G-register isomers. The possibility to fold into the direction of a four-fold basin with an increased conformational space thus compensates for the entropy penalty during folding resulting in an accelerated concurrent overall folding for 2345.^[76] We denote the G-register isomers of 2345 with respect to the G-tracts 3 (first digit) and 5 (second digit) in a two digit abbreviation depending if the registers are shifted in 5'- or 3'-direction, resulting in conformations 33, 35, 55 and 53.^[30,69,76] Out of the four possible G-register isomers, only two are majorly populated in the wildtype sequence at thermal equilibrium, namely conformations 33 and 53.^[30,198] While 33 is thermodynamically slightly preferred, we have observed a kinetic overshoot for 53 in K^+ -induced

folding. The subsequent refolding between these states proceeds on a timescale of hours, proceeding via transitory ensembles that do not require a complete unfolding-refolding mechanism.

It was shown that also the remaining two competing wells with spare-tire isomers 1234 and 1245 can be populated, if any of the randomly sampled initial contacts or geometries are trapped sufficiently long to result in macrostates that can undergo further kinetic steps towards other spare-tire isomers.^[33] The folding kinetics of each conformation are highly dependent on the lifetimes of potential *on/off*-pathway intermediates.^[33,95,110] The experimental data from our recent publication^[33] in line with observations in other experimental studies^[95,110] suggest that the stability of a major possible intermediate is determined by the length of the proximal loop (Figure 6). These experimental findings support the idea that an initial collapse into lateral hairpins might be a reasonable first step down the folding energy landscape. The findings of several pre-folded hairpins in other G4 forming sequences are in line with the conclusions for *cMYC* G4 folding.^[25,173,174] We have described this intermediate as misfolded state during folding of the spare-tire isomer 1234 based on CD- and NMR-spectroscopical observations.^[33] While the stable G4 conformation of this G4 forming sequence is parallel, we detected an anti-parallel intermediate conformation with a life-time of several minutes. Gray *et al.* have reported similar observations for another parallel spare-tire isomer (2345-33) of the *cMYC* G4, with a clear

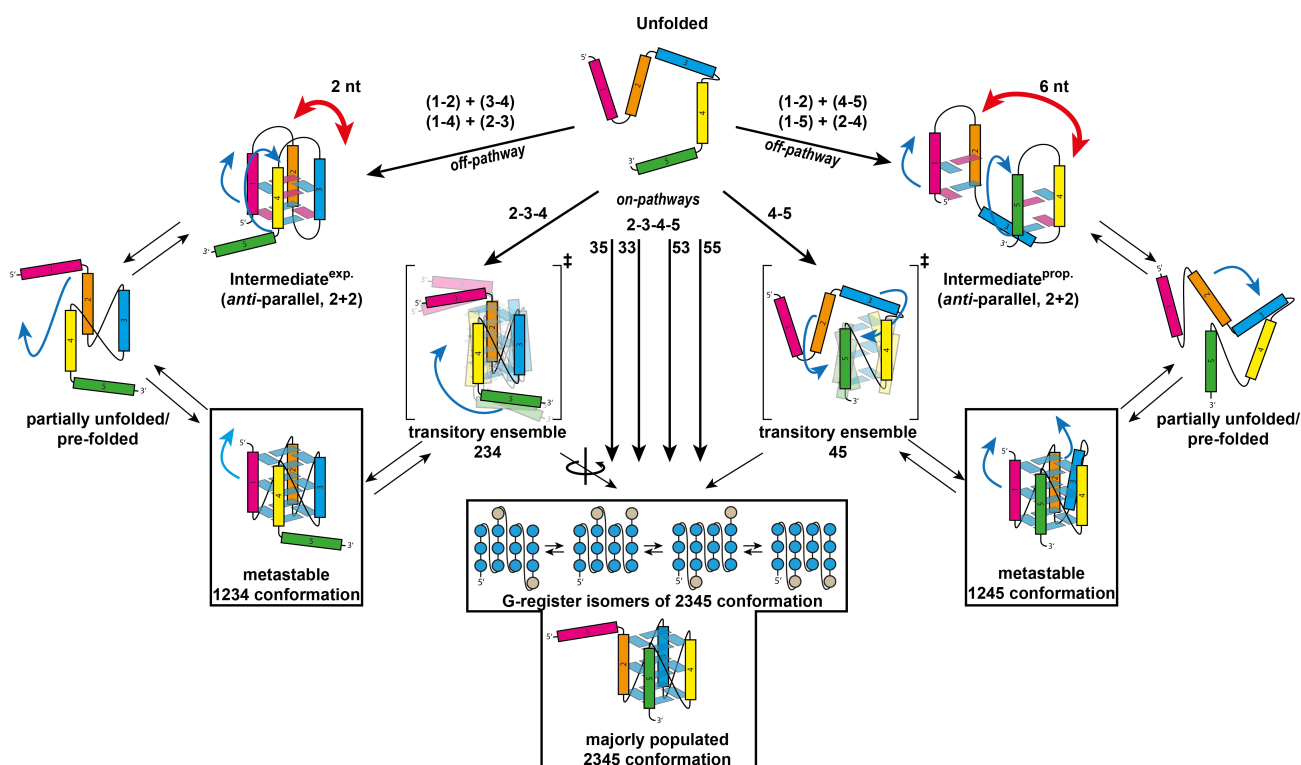


FIGURE 5 Dissection of the *cMYC* G4 folding energy landscape. Proposed model of the *cMYC* G4 structural ensemble including non-canonical folding isomerism. The three major wells are represented by stable parallel spare-tire isomers. The well for the majorly populated conformation 2345 is subdivided into different G-register isomers that share common pathways and are separated in a late stage during folding. Figure adapted from previous publication (Figure parts adapted from Grün.^[8] Adapted with permission from Grün *et al.*^[33] Copyright 2021 The Authors. Published by American Chemical Society)

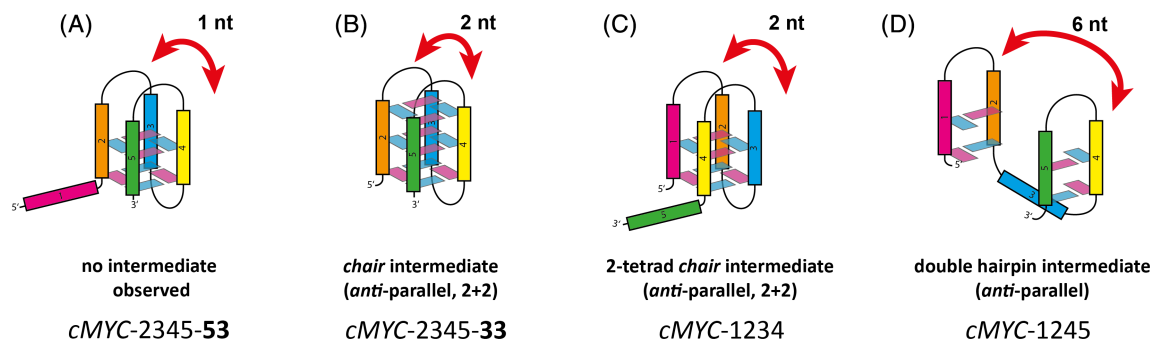


FIGURE 6 Proposed intermediates during *cMYC* G4 folding. The proximal loop determines the propensity and stability of an anti-parallel *chair* formation as kinetic intermediate. (a) Folding kinetics of the 2345 conformation with a 5'-shifted G-tract 3 and a 3'-shifted G-tract 5 (abbr. 53) are fast and no intermediate formation was experimentally observed. (b) When G-tract 3 is shifted in the 3'-direction the length of the proximal loop increases to 2 nts. Gray *et al.* have observed the formation of an anti-parallel intermediate with time-resolved circular dichroism spectroscopy.^[95] (c) The loop arrangement in the 1234 conformation is similar to (b) resulting in the formation of an intermediate with a comparable lifetime. (d) The proximal loop for a potential anti-parallel chair G4 in conformation 1245 is very long (6 nts). The high flexibility causes unprecedented kinetic behavior, which is determined by the lifetimes of the double hairpins (Figure parts adapted from Grün^[8])

anti-parallel CD-signature.^[95] While Gray *et al.* have concluded an *on-pathway* intermediate state that is able to rearrange into the parallel conformation, we have proposed an *off-pathway* state. The kinetic models thus are slightly different, but nevertheless: an anti-parallel conformation with mixed *syn-anti* patterns for the nucleobase orientations requires refolding in order to transform to the stable parallel conformation. This intermediate can be further linked to the single K^+ -conformation reported by Marchand *et al.*,^[110] which prompted the proposal of a 2-tetrad chair conformation. The low stability for the resulting single nucleotide lateral loops in case of a closed attical G-tetrad is predicted from simulations^[141,142,156] and further supports this structural model. Interestingly, although the anti-parallel intermediate formation was reported for 2345-53 we could not detect the stable formation of intermediates with relevant lifetimes for NMR spectroscopy (>ms) in the wildtype-like sequence that is able to adopt both the 53 and the 33 conformation.

4.2 | BCL2 promoter

In the *BCL2* promoter the situation is already more complex than in *cMYC*, since here a total of three G4 elements can be found, where each of the G4 elements shows inherent polymorphism (Figure 4, I; P32^[199]: -1906 to -1875, not shown; Pu39^[169,200]: -1489 to -1451; P1G4^[190]: -1439 to -1412).^[201,202] Even though each G4 element (Pu39 or P1G4, for P32 no structural data are available) shows an increased number of isomeric conformational substates the G4 elements are still well separated, which in principle allows an isolated investigation of each G4 ensemble. In particular, the Pu39 G4 element can adopt different spare-tire isomers incorporating either G-tracts 2345 (hybrid)^[200] or 1245 (parallel).^[190] The folding of Pu39 has been investigated with force spectroscopy and it was found that *BCL2-2345* (7 nt long proximal loop) is kinetically favored, despite *BCL2-1245* (13 nt long proximal loop) being thermally more stable.^[71,201] The findings further suggest the involvement of additional

states during folding, presumably *BCL2-1234* that had not been described structurally before.

4.3 | hTERT promoter

The G4 forming sequence in the human telomerase reverse transcriptase promoter (*hTERT*) consists of an extended G-rich sequence (-20 to -110 nt upstream of the TSS; Figure 4, I) that can adopt up to three concurrent G4s that adopt a subset of G-register isomers.^[24,191,203,204] It has been initially proposed that this G4 element consists of two stacked G4s with an unprecedented long hairpin loop formation (Figure 4, I: PQS2 III/IV + PQS3 I/II),^[204] newer models for the entire G4 element, however, presume a stacked triple parallel G4 arrangement.^[191,203] In the only available study on the folding dynamics of possible higher-order G4 structures published by Selvam *et al.* possible folding pathways in *hTERT* have been investigated with an mechanical unwinding approach.^[205] The presented findings for the complex unfolding patterns imply multiple pathways and associative interactions between neighboring G4s.

The isolated 5'-proximal G4 (Figure 4, I: PQS1) was shown to coexist in a hybrid and parallel conformation.^[24] While the hybrid conformation is thermodynamically slightly more favored, the folding kinetics for the parallel conformation are $\sim 1.6\times$ faster.^[69] The refolding dynamics between the hybrid and parallel conformation are slow ($\sim 0.7\text{ h}^{-1}$).^[69,187] These findings prompt the hypothesis of a potential cooperative kinetic mechanism that guides the G4 element towards the “correctly” (parallel) folded higher-order G4 structures in *hTERT*.

4.4 | cKIT promoter

A similar situation to the *hTERT* promoter can be found in the promoter region of *cKIT* (Figure 4, I),^[206-208] with three highly

polymorphic nearby G4 structures (*cKIT-1*,^[21] *cKIT**,^[25,209] *cKIT-2*,^[105,210,211] in this particular order in 3'-5' direction). The adjacent G4s *cKIT** and *cKIT-2* show a crosstalk and mutually influence the stability of each other.^[209] *cKIT-2* forms a parallel G4 and features complex folding kinetics with branched pathways, long lived intermediates and a clear evidence for the formation of pre-folded states even in the absence of monovalent cations.^[105] In contrast to that, *cKIT** forms a two-tetrad anti-parallel (*chair*) G4, which folds significantly faster than *cKIT-2* (and presumably *cKIT-1*).^[110,174] While it is reasonable to assume that the faster folding of *cKIT** might guide the folding of the adjacent G4s, it is worth noting that the folding of *cKIT** itself is highly influenced by the constitution of the 3'-tail and thus depends on a folding cooperativity with *cKIT-1*. Hence, a mutual dependence of the folding dynamics of each G4 in higher-order G4 structures seems very likely.

5 | DISCUSSION AND FUTURE CHALLENGES IN THE FIELD

We have herein presented an overview of experimental studies that shed light on different aspects of the complex folding dynamics of G4s, in particular DNA G4s from gene promoter sequences. To our understanding, it is not expedient to conclude on generalized folding pathways, or even recurrent mechanistic patterns in different G4 forming sequences. Attempts to draw simplified folding trajectories will always come short on the conformational diversity of possible macrostates during folding. There are plenty of ways to fold a G4 and it is likely that many of them contribute to a realm, or corridor towards folded G4s. However, not all routes lead to Rome: some are wrong tracks, irreversible pathways that lead to misfolded states or short tracks that lead to kinetically favored states. Every G4 folding event in chromosomal DNA is basically a single take - other than the ensemble average of folding events typically investigated in most spectroscopic experiments. There might be cellular regulation mechanisms that guide G4 folding and refolding. Up to now, to little is known about how the folding dynamics are affected in the presence of G4 interacting proteins and the molecular mechanisms of G4 folding in a genomic context. It is conceivable that G4 folding is variable during different stages of the cell cycle or that potential G4 chaperones mediate refolding of different G4 conformations. There is now growing evidence that the non-canonical polymorphism of G4 structures is involved into regulation cascades of proteins that differentiate between different distinct G4 conformations.^[59] Nucleolin is one well-established example for the discrimination of different loop isomers in parallel G4s^[57,58] and G4 unwinding helicases do as well show vastly different unwinding efficiencies on different topological isomers.^[38] Furthermore, the epigenetic modulation of chromatin states is now strongly linked to G4 formation and their influence on genetic stability.^[39,41] G4 folding in this context might be not influenced just locally but could be mechanically mediated from distant sites in gene promoters through the DNA polymer chain.^[212] Thus, biophysical studies on systems with increasing molecular size, more complex than single stranded

oligonucleotides will be required to elucidate the altered folding dynamics of G4 forming sequences.

ACKNOWLEDGEMENTS

The authors thank Julia Wirmer-Bartoschek and Ines Burkhart for discussions and proof reading of the manuscript. J. Tassilo Grün is supported by the Israel Academy of Sciences and Humanities & Council for Higher Education Excellence Fellowship Program for International Postdoctoral Researchers. This work was further supported by the Deutsche Forschungsgemeinschaft (DFG) through GRK1986 "CLIC". Work at BMRZ is supported by the states of Hesse. Open access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

J. Tassilo Grün  <https://orcid.org/0000-0002-7571-9656>

Harald Schwalbe  <https://orcid.org/0000-0001-5693-7909>

REFERENCES

- [1] D. Varshney, J. Spiegel, K. Zyner, D. Tannahill, S. Balasubramanian, *Nat. Rev. Mol. Cell Biol.* **2020**, 21, 459.
- [2] J. Jana, S. Mohr, Y. M. Vianney, K. Weisz, *RSC Chem. Biol.* **2021**, 2, 338.
- [3] J. Spiegel, S. Adhikari, S. Balasubramanian, *Trends Chem.* **2020**, 2, 123.
- [4] S. Neidle, *Nat. Rev. Chem.* **2017**, 1, 41.
- [5] V. Kocman, J. Plavec, *Nat. Commun.* **2014**, 5, 5831.
- [6] V. Kocman, J. Plavec, *Nat. Commun.* **2017**, 8, 15355.
- [7] S. Hadži, V. Kocman, D. Oblak, J. Plavec, J. Lah, *Angew. Chem. Int. Ed. Engl.* **2019**, 58, 2387.
- [8] J. T. Grün Time-resolved NMR-spectroscopic studies of conformational dynamics in DNA G-quadruplexes, Goethe-Universität Frankfurt am Main (PhD Thesis), 2021
- [9] J. L. Huppert, *Nucleic Acids Res.* **2005**, 33, 2908.
- [10] S. A. Dvorkin, A. I. Karsiotis, M. Webba da Silva, *Sci. Adv.* **2018**, 4, eaat3007.
- [11] A. Varizhuk, D. Ischenko, V. Tsvetkov, R. Novikov, N. Kulemin, D. Kaluzhny, M. Vlasenok, V. Naumov, I. Smirnov, G. Pozmogova, *Biochimie* **2017**, 135, 54.
- [12] H. L. Lightfoot, T. Hagen, N. J. Tatum, J. Hall, *FEBS Lett.* **2019**, 593, 2083.
- [13] A. Piazza, X. Cui, M. Adrian, F. Samazan, B. Heddi, A.-T. Phan, A. G. Nicolas, *eLife* **2017**, 6, e26884.
- [14] V. T. Mukundan, A. T. Phan, *J. Am. Chem. Soc.* **2013**, 135, 5017.
- [15] P. Das, K. H. Ngo, F. R. Winnerdy, A. Maity, B. Bakalar, Y. Mechulam, E. Schmitt, A. T. Phan, *Nucleic Acids Res.* **2021**, 49, 1724.
- [16] T. Q. Ngoc Nguyen, K. W. Lim, A. T. Phan, *Nucleic Acids Res.* **2020**, 48, 10567.
- [17] T. H. A. Truong, F. R. Winnerdy, A. T. Phan, *Angew. Chemie Int. Ed.* **2019**, 58, 13834.
- [18] M. Marušič, R. N. Veedu, J. Wengel, J. Plavec, *Nucleic Acids Res.* **2013**, 41, 9524.

- [19] L. Schnarr, J. Jana, P. Preckwinkel, K. Weisz, *J. Phys. Chem. B* **2020**, *124*, 2778.
- [20] A. T. Phan, V. Kuryavyi, H. Y. Gaw, D. J. Patel, *Nat. Chem. Biol.* **2005**, *1*, 167.
- [21] A. T. Phan, V. Kuryavyi, S. Burge, S. Neidle, D. J. Patel, *J. Am. Chem. Soc.* **2007**, *129*, 4386.
- [22] Y. Chen, P. Agrawal, R. V. Brown, E. Hatzakis, L. Hurley, D. Yang, *J. Am. Chem. Soc.* **2012**, *134*, 13220.
- [23] M. Adrian, D. J. Ang, C. J. Lech, B. Heddi, A. Nicolas, A. T. Phan, *J. Am. Chem. Soc.* **2014**, *136*, 6297.
- [24] K. W. Lim, L. Lacroix, D. J. E. Yue, J. K. C. Lim, J. M. W. Lim, A. T. Phan, *J. Am. Chem. Soc.* **2010**, *132*, 12331.
- [25] M. L. Greco, A. Kotar, R. Rigo, C. Cristofari, J. Plavec, C. Sissi, *Nucleic Acids Res.* **2017**, *45*, 10132.
- [26] J. Marquevielle, C. Robert, O. Lagrabette, M. Wahid, A. Bourdoncle, L. E. Xodo, J.-L. Mergny, G. F. Salgado, *Nucleic Acids Res.* **2020**, *48*, 9336.
- [27] X. Li, K. Zheng, J. Zhang, H. Liu, Y. He, B. Yuan, Y. Hao, Z. Tan, *Proc. Natl. Acad. Sci. U. S. A* **2015**, *112*, 14581.
- [28] Y. He, K. Zheng, C. Wen, X. Li, J. Gong, Y. Hao, Y. Zhao, Z. Tan, *J. Am. Chem. Soc.* **2020**, *142*, 11394.
- [29] I. Kejnovská, P. Stadlbauer, L. Trantírek, D. Renčíuk, M. Gajarský, D. Krafčík, J. Palacký, K. Bednářová, J. Šponer, J. L. Mergny, M. Vorlíčková, *Chem. - A Eur. J.* **2021**, *27*, 12115.
- [30] R. W. Harkness, A. K. Mittermaier, *Nucleic Acids Res.* **2016**, *44*, 3481.
- [31] A. M. Fleming, J. Zhou, S. S. Wallace, C. J. Burrows, *ACS Cent. Sci.* **2015**, *1*, 226.
- [32] C. A. Omega, A. M. Fleming, C. J. Burrows, *Biochemistry* **2018**, *57*, 2958.
- [33] J. T. Grün, A. Blümler, I. Burkhart, J. Wirmer-Bartoschek, A. Heckel, H. Schwalbe, *J. Am. Chem. Soc.* **2021**, *143*, 6185.
- [34] N. Kim, *Curr. Med. Chem.* **2019**, *26*, 2898.
- [35] L. K. Lerner, J. E. Sale, *Genes (Basel)* **2019**, *10*, 95.
- [36] T. M. Bryan, *Molecules* **2019**, *24*, 3439.
- [37] P. Lejault, J. Mitteaux, F. R. Sperti, D. Monchaud, *Cell Chem. Biol.* **2021**, *28*, 436.
- [38] M. Sauer, K. Paeschke, *Biochem. Soc. Trans.* **2017**, *45*, 1173.
- [39] J. Robinson, F. Raguseo, S. P. Nuccio, D. Liano, M. Di Antonio, *Nucleic Acids Res.* **2021**, *49*, 8419.
- [40] G. Guilbaud, P. Murat, B. Recolin, B. C. Campbell, A. Maiter, J. E. Sale, S. Balasubramanian, *Nat. Chem.* **2017**, *9*, 1110.
- [41] R. Hänsel-Hertsch, D. Beraldi, S. V. Lensing, G. Marsico, K. Zyner, A. Parry, M. Di Antonio, J. Pike, H. Kimura, M. Narita, D. Tannahill, S. Balasubramanian, *Nat. Genet.* **2016**, *48*, 1267.
- [42] S.-Q. Mao, A. T. Ghanbarian, J. Spiegel, S. Martínez Cuesta, D. Beraldi, M. Di Antonio, G. Marsico, R. Hänsel-Hertsch, D. Tannahill, S. Balasubramanian, *Nat. Struct. Mol. Biol.* **2018**, *25*, 951.
- [43] R. Hänsel-Hertsch, M. Di Antonio, S. Balasubramanian, *Nat. Rev. Mol. Cell Biol.* **2017**, *18*, 279.
- [44] T. M. Bryan, *Molecules* **2020**, *25*, 3686.
- [45] A. Siddiqui-Jain, C. L. Grand, D. J. Bearss, L. H. Hurley, *Proc. Natl. Acad. Sci. U. S. A* **2002**, *99*, 11593.
- [46] I. Bessi, J. Wirmer-Bartoschek, J. Dash, H. Schwalbe, *Mod. Magn. Reson.* (Ed: G. Webb), 2nd ed., **2018**, 2189.
- [47] E. Ruggiero, S. N. Richter, *Nucleic Acids Res.* **2018**, *46*, 3270.
- [48] N. Kosiol, S. Juraneck, P. Brossart, A. Heine, K. Paeschke, *Mol. Cancer* **2021**, *20*, 40.
- [49] J. U. Guo, D. P. Bartel, *Science* **2016**, *353*, aaf5371.
- [50] G. Biffi, M. Di Antonio, D. Tannahill, S. Balasubramanian, *Nat. Chem.* **2014**, *6*, 75.
- [51] L. Dumas, P. Herviou, E. Dassi, A. Cammas, S. Millevoi, *Trends Biochem. Sci.* **2021**, *46*, 270.
- [52] N. An, A. M. Fleming, C. J. Burrows, *ACS Chem. Biol.* **2016**, *11*, 500.
- [53] A. M. Fleming, J. Zhu, Y. Ding, C. J. Burrows, *ACS Chem. Biol.* **2017**, *12*, 2417.
- [54] A. M. Fleming, J. Zhu, Y. Ding, J. A. Visser, J. Zhu, C. J. Burrows, *Biochemistry* **2018**, *57*, 991.
- [55] A. M. Fleming, C. J. Burrows, *J. Am. Chem. Soc.* **2020**, *142*, 1115.
- [56] J. Zhou, A. M. Fleming, A. M. Averill, C. J. Burrows, S. S. Wallace, *Nucleic Acids Res.* **2015**, *43*, 4039.
- [57] S. Lago, E. Tosoni, M. Nadai, M. Palumbo, S. N. Richter, *Biochim. Biophys. Acta - Gen. Subj.* **2017**, *1861*, 1371.
- [58] A. Saha, P. Duchambon, V. Masson, D. Loew, S. Bombard, M.-P. Teulade-Fichou, *Biochemistry* **2020**, *59*, 1261.
- [59] P. Sengupta, A. Bhattacharya, G. Sa, T. Das, S. Chatterjee, *Biochemistry* **2019**, *58*, 1975.
- [60] W.-Q. Wu, X.-M. Hou, M. Li, S.-X. Dou, X.-G. Xi, *Nucleic Acids Res.* **2015**, *43*, 4614.
- [61] V. Popuri, C. Z. Bachrati, L. Muzzolini, G. Mosedale, S. Costantini, E. Giacomini, I. D. Hickson, A. Vindigni, *J. Biol. Chem.* **2008**, *283*, 17766.
- [62] O. Mendoza, A. Bourdoncle, J. B. Boulé, R. M. Brosh, J. L. Mergny, *Nucleic Acids Res.* **2016**, *44*, 1989.
- [63] A.-L. Valton, M.-N. Prioleau, *Trends Genet.* **2016**, *32*, 697.
- [64] S. Chatterjee, J. Ziegelbaum, P. Savitsky, A. Sturzenegger, D. Huttner, P. Janscak, I. D. Hickson, O. Gileadi, E. Rothenberg, *Nat. Commun.* **2014**, *5*, 5556.
- [65] R. Tippiana, H. Hwang, P. L. Opreko, V. A. Bohr, S. Myong, *Proc. Natl. Acad. Sci. U. S. A* **2016**, *113*, 8448.
- [66] M. Volek, S. Kolesnikova, K. Svehlova, P. Srb, R. Sgallova, T. Streckerova, J. A. Redondo, V. Veverka, E. A. Curtis, *Nucleic Acids Res.* **2021**, *49*, 1816.
- [67] J. Gong, C. Wen, M. Tang, R. Duan, J. Chen, J. Zhang, K. Zheng, Y. He, Y. Hao, Q. Yu, S. Ren, Z. Tan, *Proc. Natl. Acad. Sci. U. S. A* **2021**, *118*, e2013230118.
- [68] T. Frelih, B. Wang, J. Plavec, P. Šket, *Nucleic Acids Res.* **2020**, *48*, 2189.
- [69] J. T. Grün, C. Hennecker, D.-P. Klötzner, R. W. Harkness, I. Bessi, A. Heckel, A. K. Mittermaier, H. Schwalbe, *J. Am. Chem. Soc.* **2020**, *142*, 264.
- [70] J. Šponer, G. Bussi, P. Stadlbauer, P. Kührová, P. Banáš, B. Islam, S. Haider, S. Neidle, M. Otyepka, *Biochim. Biophys. Acta. Gen. Subj.* **2017**, *1861*, 1246.
- [71] Y. Cheng, Q. Tang, Y. Li, Y. Zhang, C. Zhao, J. Yan, H. You, *J. Biol. Chem.* **2019**, *294*, 5890.
- [72] Y. Sannohe, M. Endo, Y. Katsuda, K. Hidaka, H. Sugiyama, *J. Am. Chem. Soc.* **2010**, *132*, 16311.
- [73] R. W. Harkness, V. N. Avakyan, H. F. Sleiman, A. K. Mittermaier, *Nat. Commun.* **2018**, *9*, 3152.
- [74] J.-L. Mergny, L. Lacroix, *Oligonucleotides* **2003**, *13*, 515.
- [75] A. N. Lane, J. B. Chaires, R. D. Gray, J. O. Trent, *Nucleic Acids Res.* **2008**, *36*, 5482.
- [76] R. W. Harkness, C. Hennecker, J. T. Grün, A. Blümler, A. Heckel, H. Schwalbe, A. K. Mittermaier, *Nucleic Acids Res.* **2021**, *49*, 1247.
- [77] K. Laouer, M. Schmid, F. Wien, P. Changenet, F. Hache, *J. Phys. Chem. B* **2021**, *125*, 8088.
- [78] A. Marchand, M. F. Czar, E. N. Eggel, J. Kaeslin, R. Zenobi, *Nat. Commun.* **2020**, *11*, 1.
- [79] J. Rinnenthal, D. Wagner, T. Marquardsen, A. Krahn, F. Engelke, H. Schwalbe, *J. Magn. Reson.* **2015**, *251*, 84.
- [80] G. Pintér, H. Schwalbe, *Angew. Chem. Int. Ed. Engl.* **2020**, *59*, 22086.
- [81] G. Pintér, K. F. Hohmann, J. T. Grün, J. Wirmer-Bartoschek, C. Glaubitz, B. Fürtig, H. Schwalbe, *Magn. Reson.* **2021**, *2*, 291.
- [82] R. D. Gray, J. B. Chaires, *Methods* **2012**, *57*, 47.
- [83] D. Müller, I. Bessi, C. Richter, H. Schwalbe, *Angew. Chemie* **2021**, *133*, 10990.
- [84] O. Binas, I. Bessi, H. Schwalbe, *ChemBioChem* **2020**, *21*, 1656.
- [85] I. Bessi, H. R. A. Jonker, C. Richter, H. Schwalbe, *Angew. Chemie - Int. Ed* **2015**, *54*, 8444.
- [86] A. Y. Q. Zhang, S. Balasubramanian, *J. Am. Chem. Soc.* **2012**, *134*, 19297.
- [87] A. Heckel, G. Mayer, *J. Am. Chem. Soc.* **2005**, *127*, 822.
- [88] G. Mayer, L. Kröck, V. Mikat, M. Engeser, A. Heckel, *ChemBioChem* **2005**, *6*, 1966.
- [89] P. Wenter, B. Fürtig, A. Hainard, H. Schwalbe, S. Pitsch, *Angew. Chemie - Int. Ed* **2005**, *44*, 2600.

- [90] P. Wenter, B. Fürtig, A. Hainard, H. Schwalbe, S. Pitsch, *ChemBioChem* **2006**, *7*, 417.
- [91] B. Fürtig, P. Wenter, L. Reymond, C. Richter, S. Pitsch, H. Schwalbe, *J. Am. Chem. Soc.* **2007**, *129*, 16222.
- [92] B. Fürtig, J. Buck, V. Manoharan, W. Bermel, A. Jäschke, P. Wenter, S. Pitsch, H. Schwalbe, *Biopolymers* **2007**, *86*, 360.
- [93] P. Müller, M. Sahlbach, S. Gasper, G. Mayer, J. Müller, B. Pötzsch, A. Heckel, *Angew. Chemie* **2021**, *60*, 22441.
- [94] J. Thevarpadam, I. Bessi, O. Binas, D. P. N. Gonçalves, C. Slavov, H. R. A. Jonker, C. Richter, J. Wachtveitl, H. Schwalbe, A. Heckel, *Angew. Chemie Int. Ed.* **2016**, *55*, 2738.
- [95] R. D. Gray, J. O. Trent, S. Arumugam, J. B. Chaires, *J. Phys. Chem. Lett.* **2019**, *10*, 1146.
- [96] J. J. Green, L. Ying, D. Klenerman, S. Balasubramanian, *J. Am. Chem. Soc.* **2003**, *125*, 3763.
- [97] A. Risitano, K. R. Fox, *Biochemistry* **2003**, *42*, 6507.
- [98] A. N. Lane, *Biochimie* **2012**, *94*, 277.
- [99] B. Klejevskaja, A. L. B. Pyne, M. Reynolds, A. Shivalingam, R. Thorogate, B. W. Hoogenboom, L. Ying, R. Vilar, *Chem. Commun.* **2016**, *52*, 12454.
- [100] R. D. Gray, J. O. Trent, J. B. Chaires, *J. Mol. Biol.* **2014**, *426*, 1629.
- [101] J. B. Chaires, *FEBS J.* **2010**, *277*, 1098.
- [102] R. D. Gray, J. B. Chaires, *Nucleic Acids Res.* **2008**, *36*, 4191.
- [103] R. Del Villar-Guerra, J. O. Trent, J. B. Chaires, *Angew. Chem. Int. Ed. Engl.* **2018**, *57*, 7171.
- [104] C. Sathyaseelan, V. Vijayakumar, T. Rathinavelan, *J. Mol. Biol.* **2021**, *433*, 166629.
- [105] R. Rigo, W. L. Dean, R. D. Gray, J. B. Chaires, C. Sissi, *Nucleic Acids Res.* **2017**, *45*, 13056.
- [106] J.-L. Mergny, A.-T. Phan, L. Lacroix, *FEBS Lett.* **1998**, *435*, 74.
- [107] J. L. Mergny, J. Li, L. Lacroix, S. Amrane, J. B. Chaires, *Nucleic Acids Res.* **2005**, *33*, 1.
- [108] A. I. Karsisiotis, N. M. A. Hessari, E. Novellino, G. P. Spada, A. Randazzo, M. Webba Da Silva, *Angew. Chemie - Int. Ed.* **2011**, *50*, 10645.
- [109] A. Ghosh, E. Largy, V. Gabelica, *Nucleic Acids Res.* **2021**, *49*, 2333.
- [110] A. Marchand, V. Erie Gabelica, *Nucleic Acids Res.* **2016**, *44*, 10999.
- [111] A. Marchand, F. Rosu, R. Zenobi, V. Gabelica, *J. Am. Chem. Soc.* **2018**, *140*, 12553.
- [112] S. Daly, F. Rosu, V. Gabelica, *Science* **2020**, *368*, 1465.
- [113] T. M. Allison, P. Barran, J. L. P. Benesch, S. Cianferani, M. T. Degiacomi, V. Gabelica, R. Grandori, E. G. Marklund, T. Menneteau, L. G. Migas, A. Politis, M. Sharon, F. Sobott, K. Thalassinou, *Anal. Chem.* **2020**, *92*, 10881.
- [114] T. M. Allison, P. Barran, S. Cianferani, M. T. Degiacomi, V. Gabelica, R. Grandori, E. G. Marklund, T. Menneteau, L. G. Migas, A. Politis, M. Sharon, F. Sobott, K. Thalassinou, J. L. P. Benesch, *Anal. Chem.* **2020**, *92*, 10872.
- [115] V. D'Atri, V. Gabelica, *Analyst* **2019**, *144*, 6074.
- [116] J. Chen, A. D. Gill, B. L. Hickey, Z. Gao, X. Cui, R. J. Hooley, W. Zhong, *J. Am. Chem. Soc.* **2021**, *143*, 12791.
- [117] H. You, J. Wu, F. Shao, J. Yan, *J. Am. Chem. Soc.* **2015**, *137*, 2424.
- [118] W. Li, X.-M. Hou, P.-Y. Wang, X.-G. Xi, M. Li, *J. Am. Chem. Soc.* **2013**, *135*, 6423.
- [119] H.-P. Ju, Y.-Z. Wang, J. You, X.-M. Hou, X.-G. Xi, S.-X. Dou, W. Li, P.-Y. Wang, *ACS Omega* **2016**, *1*, 244.
- [120] D. Koirala, C. Ghimire, C. Bohrer, Y. Sannohe, H. Sugiyama, H. Mao, *J. Am. Chem. Soc.* **2013**, *135*, 2235.
- [121] Y. Cui, D. Kong, C. Ghimire, C. Xu, H. Mao, *Biochemistry* **2016**, *55*, 2291.
- [122] S. Dhakal, Z. Yu, R. Konik, Y. Cui, D. Koirala, H. Mao, *Biophys. J.* **2012**, *102*, 2575.
- [123] X. Long, M. D. Stone, *PLoS One* **2013**, *8*, e83420.
- [124] X. Long, J. W. Parks, C. R. Bagshaw, M. D. Stone, *Nucleic Acids Res.* **2013**, *41*, 2746.
- [125] J. Mitra, M. A. Makurath, T. T. M. Ngo, A. Troitskaia, Y. R. Chemla, T. Ha, *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 8350.
- [126] D. Koirala, T. Mashimo, Y. Sannohe, Z. Yu, H. Mao, H. Sugiyama, *Chem. Commun.* **2012**, *48*, 2006.
- [127] T. Mashimo, H. Yagi, Y. Sannohe, A. Rajendran, H. Sugiyama, *J. Am. Chem. Soc.* **2010**, *132*, 14910.
- [128] M. Endo, X. Xing, X. Zhou, T. Emura, K. Hidaka, B. Tuesuwan, H. Sugiyama, *ACS Nano* **2015**, *9*, 9922.
- [129] S. Søndergaard, M. Aznauryan, E. K. Hastrup, B. Schiøtt, V. Birkedal, B. Corry, *ChemPhysChem* **2015**, *16*, 2562.
- [130] M. Aznauryan, S. Søndergaard, S. L. Noer, B. Schiøtt, V. Birkedal, *Nucleic Acids Res.* **2016**, *44*, 11024.
- [131] S. L. Noer, S. Preus, D. Gudnason, M. Aznauryan, J.-L. Mergny, V. Birkedal, *Nucleic Acids Res.* **2016**, *44*, 464.
- [132] A. F. Voter, Y. Qiu, R. Tippana, S. Myong, J. L. Keck, *Nat. Commun.* **2018**, *9*, 4201.
- [133] F.-Y. Teng, T.-T. Wang, H.-L. Guo, B.-G. Xin, B. Sun, S.-X. Dou, X.-G. Xi, X.-M. Hou, *J. Biol. Chem.* **2020**, *295*, 17646.
- [134] X.-M. Hou, W.-Q. Wu, X.-L. Duan, N.-N. Liu, H.-H. Li, J. Fu, S.-X. Dou, M. Li, X.-G. Xi, *Biochem. J.* **2015**, *466*, 189.
- [135] K.-I. Oh, J. Kim, C.-J. Park, J.-H. Lee, *Int. J. Mol. Sci.* **2020**, *21*, 2673.
- [136] R. K. R. Sannapureddi, M. K. Mohanty, A. K. Gautam, B. Sathyamoorthy, *J. Phys. Chem. Lett.* **2020**, *11*, 48.
- [137] A. L. Lieblein, J. Buck, K. Schlepckow, B. Fürtig, H. Schwalbe, *Angew. Chemie Int. Ed.* **2012**, *51*, 250.
- [138] B. Fürtig, C. Richter, J. Wöhnert, H. Schwalbe, *ChemBioChem* **2003**, *4*, 936.
- [139] M. Adrian, B. Heddi, A. T. Phan, *Methods* **2012**, *57*, 11.
- [140] B. Liu, H. Shi, H. M. Al-Hashimi, *Curr. Opin. Struct. Biol.* **2021**, *70*, 16.
- [141] J. Šponer, B. Islam, P. Stadlbauer, S. Haider, *Annu. Rep. Med. Chem.* **2020**, *54*, 197.
- [142] P. Stadlbauer, P. Kührová, L. Vicherek, P. Banáš, M. Otyepka, L. Trantírek, J. Šponer, *Nucleic Acids Res.* **2019**, *47*, 7276.
- [143] F. Feil, S. Naumov, J. Michaelis, R. Valiullin, D. Enke, J. Kärger, C. Bräuchle, *Angew. Chemie Int. Ed.* **2012**, *51*, 1152.
- [144] T. A. Brooks, L. H. Hurley, *Nat. Rev. Cancer* **2009**, *9*, 849.
- [145] R. I. Mathad, E. Hatzakis, J. Dai, D. Yang, *Nucleic Acids Res.* **2011**, *39*, 9023.
- [146] D. A. T. Sekibo, K. R. Fox, *Nucleic Acids Res.* **2017**, *45*, 12069.
- [147] D. Li, P. Peng, Z. Yang, B. Lv, *RSC Adv.* **2019**, *9*, 26248.
- [148] J. Shen, D. Varshney, A. Simeone, X. Zhang, S. Adhikari, D. Tannahill, S. Balasubramanian, *Genome Biol.* **2021**, *22*, 1.
- [149] E. Shapiro, T. Biezuner, S. Linnarsson, *Nat. Rev. Genet.* **2013**, *14*, 618.
- [150] G. K. Marinov, B. A. Williams, K. McCue, G. P. Schroth, J. Gertz, R. M. Myers, B. J. Wold, *Genome Res.* **2014**, *24*, 496.
- [151] R. Milo, *BioEssays* **2013**, *35*, 1050.
- [152] B. Schwanhüusser, D. Busse, N. Li, G. Dittmar, J. Schuchhardt, J. Wolf, W. Chen, M. Selbach, *Nature* **2011**, *473*, 337.
- [153] B. Schwanhüusser, D. Busse, N. Li, G. Dittmar, J. Schuchhardt, J. Wolf, W. Chen, M. Selbach, *Nature* **2013**, *495*, 126.
- [154] H. Tani, R. Mizutani, K. A. Salam, K. Tano, K. Ijiri, A. Wakamatsu, T. Isogai, Y. Suzuki, N. Akimitsu, *Genome Res.* **2012**, *22*, 947.
- [155] U. Del Monte, *Cell Cycle* **2009**, *8*, 505.
- [156] P. Stadlbauer, P. Kührová, P. Banáš, J. Koča, G. Bussi, L. Trantírek, M. Otyepka, J. Šponer, *Nucleic Acids Res.* **2015**, *43*, 9626.
- [157] P. Stadlbauer, L. Trantírek, T. E. Cheatham, J. Koča, J. Šponer, *Biochimie* **2014**, *105*, 22.
- [158] P. Stadlbauer, M. Krepl, T. E. Cheatham, J. Koca, J. Šponer, *Nucleic Acids Res.* **2013**, *41*, 7128.
- [159] M. Gajarský, M. L. Živković, P. Stadlbauer, B. Pagano, R. Fiala, J. Amato, L. Tomáška, J. Šponer, J. Plavec, L. Trantírek, *J. Am. Chem. Soc.* **2017**, *139*, 3591.
- [160] M. L. Živković, M. Gajarský, K. Beková, P. Stadlbauer, L. Vicherek, M. Petrová, R. Fiala, I. Rosenberg, J. Šponer, J. Plavec, L. Trantírek, *Nucleic Acids Res.* **2021**, *49*, 2317.

- [161] T. Q. N. Nguyen, K. W. Lim, A. T. Phan, *J. Phys. Chem. B* **2020**, *124*, 5122.
- [162] E. Hatzakis, K. Okamoto, D. Yang, *Biochemistry* **2010**, *49*, 9152.
- [163] J. Singh, R. A. Padgett, *Nat. Struct. Mol. Biol.* **2009**, *16*, 1128.
- [164] X. Darzacq, Y. Shav-Tal, V. de Turris, Y. Brody, S. M. Shenoy, R. D. Phair, R. H. Singer, *Nat. Struct. Mol. Biol.* **2007**, *14*, 796.
- [165] T. Morisaki, W. G. Müller, N. Golob, D. Mazza, J. G. McNally, *Nat. Commun.* **2014**, *5*, 1.
- [166] P. Hazel, J. Huppert, S. Balasubramanian, S. Neidle, *J. Am. Chem. Soc.* **2004**, *126*, 16405.
- [167] P. A. Rachwal, I. S. Findlow, J. M. Werner, T. Brown, K. R. Fox, *Nucleic Acids Res.* **2007**, *35*, 4214.
- [168] J. Jana, K. Weisz, *ChemBioChem* **2021**, *22*, 2848.
- [169] P. Agrawal, C. Lin, R. I. Mathad, M. Carver, D. Yang, *J. Am. Chem. Soc.* **2014**, *136*, 1750.
- [170] Y. Chen, D. Yang, *Curr. Protoc. Nucleic Acid Chem* **2012**, *50*, 17.5.1.
- [171] A. Bugaut, S. Balasubramanian, *Biochemistry* **2008**, *47*, 689.
- [172] P. A. Rachwal, T. Brown, K. R. Fox, *FEBS Lett.* **2007**, *581*, 1657.
- [173] S. Čeru, P. Šket, I. Prislán, J. Lah, J. Plavec, *Angew. Chemie - Int. Ed* **2014**, *53*, 4881.
- [174] A. Kotar, R. Rigo, C. Sissi, J. Plavec, *Nucleic Acids Res.* **2019**, *47*, 2641.
- [175] M. H.-J. Kuo, Z.-F. Wang, T.-Y. Tseng, M.-H. Li, S.-T. D. Hsu, J.-J. Lin, T.-C. Chang, *J. Am. Chem. Soc.* **2015**, *137*, 210.
- [176] X.-M. Hou, Y.-B. Fu, W.-Q. Wu, L. Wang, F.-Y. Teng, P. Xie, P.-Y. Wang, X.-G. Xi, *Nucleic Acids Res.* **2017**, *45*, 11401.
- [177] L. Cerofolini, J. Amato, A. Giachetti, V. Limongelli, E. Novellino, M. Parrinello, M. Fragai, A. Randazzo, C. Luchinat, *Nucleic Acids Res.* **2014**, *42*, 13393.
- [178] H.-X. Jiang, Y. Cui, T. Zhao, H.-W. Fu, D. Koirala, J. A. Punnoose, D.-M. Kong, H. Mao, *Sci. Rep* **2015**, *5*, 1.
- [179] V. Limongelli, S. De Tito, L. Cerofolini, M. Fragai, B. Pagano, R. Trotta, S. Cosconati, L. Marinelli, E. Novellino, I. Bertini, A. Randazzo, C. Luchinat, M. Parrinello, *Angew. Chemie Int. Ed.* **2013**, *52*, 2269.
- [180] J.-L. Mergny, *Nucleic Acids Res.* **2005**, *33*, 81.
- [181] O. Mendoza, M. Porrini, G. F. Salgado, V. Gabelica, J.-L. Mergny, *Chem. - A Eur. J.* **2015**, *21*, 6732.
- [182] C. Bardin, J. L. Leroy, *Nucleic Acids Res.* **2008**, *36*, 477.
- [183] V. Gabelica, *Biochimie* **2014**, *105*, 1.
- [184] J. A. Punnoose, Y. Ma, M. E. Hoque, Y. Cui, S. Sasaki, A. H. Guo, K. Nagasawa, H. Mao, *Biochemistry* **2018**, *57*, 6946.
- [185] S. Carrino, C. D. Henneker, A. C. Murrieta, A. Mittermaier, *Nucleic Acids Res.* **2021**, *49*, 3063.
- [186] K. Bose, C. J. Lech, B. Heddi, A. T. Phan, *Nat. Commun.* **2018**, *9*, 1959.
- [187] F. Nußbaumer, M. A. Juen, C. Gasser, J. Kremser, T. Müller, M. Tollinger, C. Kreutz, *Nucleic Acids Res.* **2017**, *45*, 9178.
- [188] A. K. Todd, M. Johnston, S. Neidle, *Nucleic Acids Res.* **2005**, *33*, 2901.
- [189] T. A. Brooks, S. Kendrick, L. Hurley, *FEBS J.* **2010**, *277*, 3459.
- [190] B. Onel, M. Carver, G. Wu, D. Timonina, S. Kalarn, M. Larriva, D. Yang, *J. Am. Chem. Soc.* **2016**, *138*, 2563.
- [191] R. C. Monsen, L. DeLeeuw, W. L. Dean, R. D. Gray, T. M. Sabo, S. Chakravarthy, J. B. Chaires, J. O. Trent, *Nucleic Acids Res.* **2020**, *48*, 5720.
- [192] J. Seenisamy, E. M. Rezler, T. J. Powell, D. Tye, V. Gokhale, C. S. Joshi, A. Siddiqui-Jain, L. H. Hurley, *J. Am. Chem. Soc.* **2004**, *126*, 8702.
- [193] A. T. Phan, Y. S. Modi, D. J. Patel, *J. Am. Chem. Soc.* **2004**, *126*, 8710.
- [194] A. Ambrus, D. Chen, J. Dai, R. A. Jones, D. Yang, *Biochemistry* **2005**, *44*, 2048.
- [195] J. Dickerhoff, B. Onel, L. Chen, Y. Chen, D. Yang, *ACS Omega* **2019**, *4*, 2533.
- [196] S. Stump, T.-C. Mou, S. R. Sprang, N. R. Natale, H. D. Beall, *PLoS One* **2018**, *13*, e0205584.
- [197] J. Jana, K. Weisz, *Chem. - A Eur. J.* **2020**, *26*, 17242.
- [198] R. W. Harkness, A. K. Mittermaier, *Biochim. Biophys. Acta - Prot Proteom* **2017**, *1865*, 1544.
- [199] H. Sun, J. Xiang, Y. Shi, Q. Yang, A. Guan, Q. Li, L. Yu, Q. Shang, H. Zhang, Y. Tang, G. Xu, *Biochim. Biophys. Acta - Gen. Subj.* **2014**, *1840*, 3052.
- [200] J. Dai, D. Chen, R. A. Jones, L. H. Hurley, D. Yang, *Nucleic Acids Res.* **2006**, *34*, 5133.
- [201] P. Sengupta, S. Chattopadhyay, S. Chatterjee, *Drug Discovery Today* **2017**, *22*, 1165.
- [202] T. S. Dexheimer, D. Sun, L. H. Hurley, *J. Am. Chem. Soc.* **2006**, *128*, 5404.
- [203] J. B. Chaires, J. O. Trent, R. D. Gray, W. L. Dean, R. Buscaglia, S. D. Thomas, D. M. Miller, *PLoS One* **2014**, *9*, e115580.
- [204] S. L. Palumbo, S. W. Ebbinghaus, L. H. Hurley, *J. Am. Chem. Soc.* **2009**, *131*, 10878.
- [205] S. Selvam, Z. Yu, H. Mao, *Nucleic Acids Res.* **2016**, *44*, 45.
- [206] S. Da Ros, E. Zorzan, M. Giantin, L. Zorro Shahidian, M. Palumbo, M. Dacasto, C. Sissi, *PLoS One* **2014**, *9*, e103876.
- [207] M. C. Miller, H. T. Le, W. L. Dean, P. A. Holt, J. B. Chaires, J. O. Trent, *Org. Biomol. Chem.* **2011**, *9*, 7633.
- [208] S. Rankin, A. P. Reszka, J. Huppert, M. Zloh, G. N. Parkinson, A. K. Todd, S. Ladame, S. Balasubramanian, S. Neidle, *J. Am. Chem. Soc.* **2005**, *127*, 10584.
- [209] R. Rigo, C. Sissi, *Biochemistry* **2017**, *56*, 4309.
- [210] S.-T. D. Hsu, P. Varnai, A. Bugaut, A. P. Reszka, S. Neidle, S. Balasubramanian, *J. Am. Chem. Soc.* **2009**, *131*, 13399.
- [211] V. Kuryavii, A. T. Phan, D. J. Patel, *Nucleic Acids Res.* **2010**, *38*, 6757.
- [212] G. Rosenblum, N. Elad, H. Rozenberg, F. Wiggers, J. Jungwirth, H. Hofmann, *Nat. Commun.* **2021**, *12*, 1.

AUTHOR BIOGRAPHIES



J. TASSILO GRÜN obtained his PhD in Chemistry in 2021 from Goethe University (Frankfurt, Germany) under the supervision of Professor Harald Schwalbe. His graduate research focused on real-time NMR investigation of G-quadruplex dynamics as part of the research program “Complex Light Control.” He is now a Postdoctoral Fellow at the Weizmann Institute of Science (Rehovot, Israel) with Professor Lucio Frydman, where his research interests center on the development of biomolecular NMR methods.



HARALD SCHWALBE obtained his PhD in Chemistry in 1995 from Goethe University (Frankfurt, Germany) under the supervision of Professor Christian Griesinger. His graduate research was on NMR methods development in protein NMR spectroscopy. He was postdoctoral fellow with Professor Sir Chris Dobson, FRS, at Oxford University (Oxford, U.K.). From 1995 to 1999, he worked on his Habilitation. He accepted an offer to become assistant professor at the Massachusetts Institute for Technology (Cambridge, U.S.A.) in 1999. In 2002, he became full professor of Organic Chemistry at Goethe University. His research focusses on the investigation of structure and dynamics of proteins, DNA, RNA and their complexes, particular with NMR spectroscopy.

How to cite this article: J. T. Grün, H. Schwalbe, *Biopolymers* **2022**, *113*(1), e23477. <https://doi.org/10.1002/bip.23477>